https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions

Himawari-8-derived diurnal variations of ground-level $PM_{2.5}$ pollution across China using a fast space-time Light Gradient Boosting Machine

Jing Wei[1,2*], Zhanqing Li[2**], Rachel T. Pinker[2], Lin Sun[3], Wenhao Xue[1], Runze Li[1]

1. State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China
2. Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
3. College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China

Correspondence to: Jing Wei (weijing_rs@163.com), Zhanqing Li (zli@atmos.umd.edu)

Abstract

$PM_{2.5}$ has been used as an important atmospheric environmental parameter primarily due to its impact on human health. $PM_{2.5}$ is affected by both natural and anthropogenic factors that usually have strong diurnal variations. Monitoring it does not only help understand the causes of air pollution but also our adaptation to it. Most existing $PM_{2.5}$ products have been derived from polar-orbiting satellites. This study exploits the usage of the next-generation geostationary meteorological satellite Himawari-8/AHI in revealing its diurnal variations. Given the huge volume of the satellite data, a highly efficient tree-based Light Gradient Boosting Machine (LightGBM) learning approach, which is based on the idea of gradient boosting, is applied by involving the spatiotemporal characteristics of air pollution, named the space-time LightGBM (STLG) model. Hourly $PM_{2.5}$ data set in China (i.e., ChinaHigh$PM_{2.5}$) at a 5 km spatial resolution is derived based on the Himawari-8/AHI aerosol products together with other variables. The hourly $PM_{2.5}$ estimates (N = 1,415,188) are well correlated with ground measurements (R^2 = 0.85) with a RMSE and MAE of 13.62 and 8.49 μg/m^3 respectively in China. Our model can capture well the $PM_{2.5}$ diurnal variations, where the pollution increases gradually in the morning, and

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

reaches a peak at about 10:00 a.m. local time, then decreases steadily until sunset. The proposed approach outperforms most traditional statistical regression and tree-based machine learning models with a much lower computation burden in terms of speed and memory, making it most suitable for routine pollution monitoring.

1. Introduction

China has faced severe environmental problems during the last two decades, especially air pollution (An et al., 2019; Chan & Yao, 2008; Li et al., 2017; Zhang et al., 2017). The sources of air pollution are complicated and come from both natural changes (e.g., forest fires, biomass burning) and human activity (e.g., industrial production, transportation) (Huang et al., 2014; Sun et al., 2014; Wei et al., 2019a). $PM_{2.5}$ has a greater impact than other air pollutants (e.g., PM_{10}, NO_2, SO_2, etc.) on the atmospheric environment and climate change (Li et al., 2017, 2019; Jacob & Winner, 2009; Ramanathan & Feng, 2009). Moreover, they can cause great harm to human health due to the smaller particle size (Delfino et al., 2005; Kim et al., 2015; Lelieveld et al., 2015). China has established and operates multiple ground-based observation networks to monitor air pollution in real-time across mainland China including information on $PM_{2.5}$ pollution.

For near-surface concentrations, the networks provide high-quality $PM_{2.5}$ measurements every hour (even every few minutes), but with rather non-uniform coverage. In recent years, an increased effort has been made in $PM_{2.5}$ estimates with products generated from multiple sun-synchronous satellites, e.g., MISR (Liu et al., 2005; van Donkelaar et al., 2006), MODIS (Liu et al., 2007; Ma et al., 2014; Wei et al., 2019a, 2020, 2021), and VIIRS (Wu et al., 2016; Yao et al., 2019). However, due to their low revisit cycles (one or two overpasses per day), they are unable to monitor the diurnal variation of pollution. Therefore, most currently available $PM_{2.5}$ dataset are at low temporal resolutions which cannot meet the acquirements of air pollution real-time monitoring, in particular, the people's daily life, e.g., if people know the time of heavy pollution during a day, they may change the times being outside, say for sports or others. However, following the launch of the Himawari-8/Advanced Himawari Imager (AHI) on 7 October 2014 (Bessho et al., 2016; Letu et al., 2020), the near-surface $PM_{2.5}$ concentrations in the eastern hemisphere can be estimated to address the diurnal cycle.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

Over the years, Wang et al. (2017) used the linear mixed-effect (LME) model, and Sun et al. (2019) applied the geographically weighted regression (GWR) and support vector regression (SVR) models to estimate hourly $PM_{2.5}$ data from the Himawari-8 aerosol optical depth (AOD) products in the Beijing–Tianjin–Hebei (BTH) region. Zhang et al. (2019) developed an improved LME model, and Xue et al. (2020) proposed an improved geographically and temporally weighted regression (IGTWR) model to derive the hourly $PM_{2.5}$ maps based on Himawari-8 AOD products over central and eastern China. In addition to the traditional statistical regression models, several artificial intelligence models, including the random forest (RF), eXtreme Gradient Boosting (XGBoost), and deep neural network (DNN), have been recently successfully adopted to Himawari-8 data to obtain hourly $PM_{2.5}$ concentrations from local regions to the whole of China (Chen et al., 2019; Liu et al., 2019; Sun et al., 2019). Nevertheless, due to its poor data mining ability, the traditional statistical regression methods usually suffer from large uncertainties. While artificial intelligence methods can achieve higher accuracy, they are often highly demanding on computational power and are thus often slow. In addition, the spatiotemporal variations of $PM_{2.5}$ were often neglected in previous studies (Chen et al., 2019; Liu et al., 2019; Sun et al., 2019; Wang et al., 2017; Zhang et al., 2019), resulting in relatively low accuracy.

Therefore, here, we develop a new high-efficiency and high-precision method for near-surface $PM_{2.5}$ estimation relying on the tree-based light gradient boosting machine (LightGBM) learning approach by including spatial and temporal information, namely, space-time LightGBM (STLG) model. This results in a high-temporal-resolution (hourly) $PM_{2.5}$ dataset over eastern China (5 km) of high quality from Himawari-8/AHI hourly AOD products. Section 2 provides the detailed information of the natural and human data used, and introduces the development of the STLG model. Section 3 validates the hourly $PM_{2.5}$ estimates, and shows the diurnal $PM_{2.5}$ variations across China, and further compares it with traditional models and previous studies. Section 4 summarizes the study.

2. Material and methods

2.1 Data source

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

2.1.1 $PM_{2.5}$ and AOD data

Individual-site $PM_{2.5}$ hourly measurements are collected at a total of 1583 monitoring stations across China during 2018 (Figure S1). Corresponding Himawari-8 hourly 5 km AOD products (500 nm) across mainland China are also collected. They are synthesized from the Level 2 instantaneous AOD products of 10 minutes, which are generated through a newly defined Lambert-assumed aerosol retrieval algorithm (Yoshida et al., 2018; Letu et al., 2018). The Himawari-8 AOD retrievals have been preliminarily evaluated against the AOD in-situ observations provided by the Aerosol Robotic (Giles et al., 2019) and Sun–Sky Radiometer (Li et al., 2018) Observation Networks, and they are well consistent (R = 0.75), and the average RMSE and MAE are 0.39 and 0.21 (Wei et al., 2019b). Here, only the high-quality AOD retrievals (500 nm) are selected for $PM_{2.5}$ estimation.

2.1.2 Meteorological condition

$PM_{2.5}$ can be significantly affected by meteorological conditions (Su et al., 2018), but most of the current available reanalysis meteorological products have low temporal resolutions (~3–6 hours). Recently (14 June 2018), the 5th generation ECMWF global atmospheric reanalysis (ERA5) at a horizontal resolution of 0.25°×0.25° has been released as well as the land version (12 July 2019) at a horizontal resolution of 0.1°×0.1° both at hourly time scale (1979 to present). Here we use seven ERA5 hourly meteorological parameters, including the 2-m temperature, total evaporation, relative humidity, 10-m u- and v-components of winds, surface pressure, and boundary layer height,

2.1.3 Human influence

Human activity is a key factor affecting $PM_{2.5}$ pollution. The global annual LandScan^{TM} product at 1 km spatial resolution in 2018 is selected to illustrate population distribution (Dobson et al., 2000). In addition, MEIC monthly anthropogenic source emission data is also employed (Li et al., 2017; Zhang et al., 2007). It is generated from agriculture, industry, power, residential, and transportation at more than 700 anthropogenic sources, including a total of 10 atmospheric pollutants and greenhouse gases. Here, four main precursors including the NH_3, NO_x, SO_2, and VOCs, and the direct emission to PM are selected.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

2.1.4 Ancillary data

Two additional ancillary data, including the MODIS monthly NDVI at a horizontal resolution of 0.05° × 0.05°, and the SRTM 90 m DEM product are selected to reflect the land cover and change and topographical conditions in China, respectively. All the selected variables (Table S1) with potential impact on $PM_{2.5}$ concentrations were resampled to 0.05°×0.05° as the Himawari-8 aerosol product.

2.2 Space-Time LightGBM model

The Light Gradient Boosting Machine (LightGBM) model, which is a newly developed tree-based machine learning approach, was introduced in 2017 (Ke et al., 2017). It is based on the gradient boosting framework to construct the decision tree and can address both regression and classification tasks. LightGBM is a fast, distributed, and highly efficient method to tackle the main challenges faced in traditional tree-based machine learning approaches, i.e., computational complexities, by reducing the number of data samples (M) and features (N). The LightGBM model includes three main steps when constructing the decision tree as follows:

1) Histogram-based algorithm. The continuous features are first converted to different bins, which are used to construct the feature index histograms without the need to sort during training. It follows by going through all the data bins to find the best split point from the feature histograms, which can significantly reduce the computation cost of the split gain. The overall complexity is $O(M \times N)$.
2) Gradient-based One-Side Sampling. The data samples are first sorted in descending order according to their absolute gradients, and the top a% of them are selected as a subset sample with large gradients. Then the b% samples are chosen from the remaining data as a subset sample with small gradients randomly. The sampled data with small gradients are multiplied by a weight coefficient ($\frac{1-a}{b}$). Last, a new classifier is learned and established using above-sampled data until convergence.
3) Exclusive Feature Bundling. A graph with weighted edges is first constructed, and each weight corresponds to the total number of conflicts between two features. Then the features are sorted in descending order according to the degree of each feature (the greater the degree, the greater the

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

conflict with other points)). Last, each feature is checked in the sorted sequence and either assigned to a combination with small conflicts or created a new combination.

Besides the main technologies mentioned above, there are also other details of the optimization, such as the leaf-wise tree growth strategy with depth restriction (Shi, 2007), histogram difference acceleration, sequential access gradient, and the support of category feature and parallel learning. All of the above-advanced methodologies make it reach high accuracy and efficiency (Ke et al., 2017).

It is well known that air pollution has spatiotemporal heterogeneities, leading to large differences in $PM_{2.5}$ concentrations from space to space on different days. However, such characteristics have always been ignored in most of the traditional statistical regression or artificial intelligence methods used in previous studies. To tackle this issue, we integrated the spatiotemporal information into the LightGBM model, and then a newly space-time LightGBM (i.e., STLG) model is developed in this study. The spatial feature is represented by the geographical distances of one pixel to other points in the circumscribed rectangle of the study region (Behrens et al., 2018; Baez-Villanueva et al., 2020). The distance is calculated using the Haversine method (Equation 1) to reflect the spherical distance between two points in the sphere space (Wei et al., 2021). The temporal feature is represented by the day of the year (DOY), which is used to distinguish each data record of different days throughout the year during the model training.

$$DIS = 2 * r * \text{asin}\left(\sqrt{sin^2\left(\frac{\varphi_2-\varphi_1}{2}\right) + \cos(\varphi_1)\cos(\varphi_2)\, sin^2\left(\frac{\gamma_2-\gamma_1}{2}\right)}\right) \qquad (1)$$

where φ and γ represent the latitude and longitude of a point in the sphere; r denotes the mean earth radius (≈ 6371 km). Figure 1 illustrates the flowchart of the new STLG model. Two independent ten-fold cross-validation methods (10-CV, Rodriguez et al., 2010) based on all the data samples (i.e., out-of-sample) and the $PM_{2.5}$ monitoring stations (i.e., out-of-station) are selected to validate the model performance and the spatial prediction ability, respectively.

[Please insert Figure 1 here]

3. Results and discussion

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

3.1 Model fitting and validation

3.1.1 Spatial-scale performance

The STLG model can also largely minimize the overfitting issue and show a strong data mining ability (Figure S2), which can more accurately establish the relationships between hourly $PM_{2.5}$ observations and its influence variables (i.e., R^2 = 0.97–0.98, RMSE = 4.18–7.31 μg/m^3). Figure 2 illustrates the out-of-sample evaluation results of estimated hourly $PM_{2.5}$ values over China from 08:00 to 17:00 local time in 2018. Results show that the STLG model yields a high accuracy in estimating hourly $PM_{2.5}$ concentrations with high sample-based CV-R^2 values ranging from 0.81 to 0.85, strong slopes of ~0.81–0.84, and small intercepts of ~5.52–7.84 μg/m^3. The uncertainties are overall small, with RMSEs (MAEs) ranging from 11.24 (6.82) μg/m^3 to 15.56 (9.79) μg/m^3. However, our model performs slightly different with small differences in main evaluation indicators throughout the day. The main reasons are that, on the one hand, the number of training samples is reduced due to sunrise (Figure 2a-b) and sunset (Figure 2i-j) in optical remote sensing, which affects the model training; on the other hand, air pollution shows obvious diurnal variations with different $PM_{2.5}$ pollution levels due to different intensity of human activities and natural conditions.

In general, our model is stable and robust, with an equal out-of-sample CV-R^2 of 0.85 and an equal regression slope of 0.81 in most hours during the daytime in China (Figure 2c-h). Furthermore, the out-of-station CV-R^2 values range from 0.76 to 0.81, and the RMSE (MAE) values range from 12.49 (7.85) μg/m^3 to 17.61 (11.33) μg/m^3 (Figure S3), indicating that our model has a strong spatial prediction ability and can well predict $PM_{2.5}$ values in those areas without surface observing stations. In addition, the station-based accuracy is slightly decreased with reference to the sample-based accuracy, which further illustrates the robustness of our model. However, two cross-validation results (e.g., slopes = 0.78–0.84) indicate that the hourly $PM_{2.5}$ concentrations are underestimated overall (Figures 2 and S3), which is a common issue in fine particle remote sensing (Wei et al., 2020). This can be explained by the large aerosol retrieval uncertainty as well as the small number of data samples under highly polluted conditions (Wei et al., 2019b, c).

[Please insert Figure 2 here]

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

In addition, the regional performance of the STLG model for hourly $PM_{2.5}$ estimates has also been evaluated (Figure 3). The hourly $PM_{2.5}$ estimates (N = 1,151,595) are highly consistent with the ground measurements with a high sample-based CV-R^2 of 0.87 and a strong regressed slope of 0.86, showing small estimation uncertainties (i.e., RMSE = 12.77 μg/m^3, MAE = 8.12 μg/m^3) over Eastern China. The STLG model shows good performance (e.g., CV-R^2 = 0.88, Slope = 0.87) in two typical urban agglomerations of public concern in China (Figure 3b, c). By contrast, our model performs relatively poorly in the Pearl River Delta (PRD, Figure 3d) region possibly due to the significant reduction in the number of data samples caused by high-frequency and long-term cloud cover in southern China. It should be noted that there are some differences in the uncertainty of hourly $PM_{2.5}$ estimates particularly because of varying levels of air pollution; the pollution level in the BTH region is about three times higher than that in the PRD region.

[Please insert Figure 3 here]

Figure 4 shows the accuracy of our STLG model at each monitoring station across China. At the individual-site scale, the number of data samples gradually decreases from Northern China to Southern China, mainly due to increasing cloud contaminations with an average of 997 in China. Except for several scattered monitoring stations in western China, the STLG model exhibits high performance and adaptability and can well estimate the hourly $PM_{2.5}$ concentrations at most monitoring stations (e.g., average CV-R^2 0 0.78, RMSE = 12.21, and MAE = 8.17 μg/m^3). In general, approximately 76%, 79%, and 82% of monitoring stations show high accuracy with out-of-sample CV-R^2 values > 0.7, RMSE values < 15 μg/m^3, and MAE values < 10 μg/m^3 in hourly $PM_{2.5}$ estimates, especially for those located in the Central and North China

[Please insert Figure 4 here]

3.1.2 Temporal-scale performance

Time series of daily performance of the STLG model in hourly $PM_{2.5}$ estimates in China have also been investigated. The data samples vary on a daily basis with an average of 3975 per day, yet more than 83% of days have a large number of data samples (> 2000) (Figure 5). The STLG model capture well

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

the hourly $PM_{2.5}$ values at most days with an average out-of-sample R^2 of 0.73, and the average RMSE and MAE are 13.06 and 8.53 μg/m^3, respectively. In general, the hourly $PM_{2.5}$ estimates are more reliable at approximately 79%, 70%, and 74% of the days with CV-R^2 values > 0.7, and RMSE and MAE values < 15 and 10 μg/m^3 in a year, respectively. The model performance also varies greatly at the seasonal level with average CV-R^2 values of 0.82, 0.71, 0.87, and 0.86, and the average RMSE values are 14.55, 9.63, 11.83, and 17.57 μg/m^3 for four seasons, respectively (Figure S4). In general, the overall uncertainty of $PM_{2.5}$ estimates increases at the beginning and end of the year, which may be caused by a harsher environment and more intense human activities in spring and winter.

[Please insert Figure 5 here]

We have also evaluated the temporally synthesized $PM_{2.5}$ data from the hourly data samples at each monitoring station in 2018 (Figure 6). The daily mean $PM_{2.5}$ estimates are highly correlated to those calculated from the surface observations (R^2 = 0.91), and the average RMSE (MAE) value is 10.11 (6.39) μg/m^3. This suggests that the STLG model can also capture the daily $PM_{2.5}$ variations accurately. Moreover, the daily synthetic $PM_{2.5}$ data derived from the geostationary satellites are based on higher temporal frequency than those derived from the sun-synchronous satellites. In general, the $PM_{2.5}$ synthetic values also show high accuracies and low estimation uncertainties (e.g., R^2 = 0.98, RMSE = 1.6–3.3 μg/m^3, MAE = 1.1–2.3 μg/m^3) from monthly to annual scales, allowing to better describe the spatiotemporal distributions and variations of $PM_{2.5}$ pollution across China.

[Please insert Figure 6 here]

3.2 Spatiotemporal characteristics

3.2.1 Diurnal variations

Figure 7 shows the Himawari-8-derived hourly mean near-surface $PM_{2.5}$ concentrations from 8:00 a.m. to 5:00 p.m. local time in 2018 across China. Our generated $PM_{2.5}$ maps can cover most areas of the Chinese mainland except for western Xinjiang and Tibet due to the limitation of satellite scanning, showing missing $PM_{2.5}$ predictions. $PM_{2.5}$ pollution shows significant diurnal variations across China, being at an overall low level at sunrise (~29.94±10.91 μg/m^3). With the increase in human activities, air

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

pollution got severe over time, reaching a peak at around 10:00–11:00 a.m. local time in China (~36±13 μg/m^3). The high levels of pollution can be maintained for several hours. With the decrease in human activities and the deposition of atmospheric fine particles, $PM_{2.5}$ decreases towards sunset in most areas across China (~23.21±9.73 μg/m^3). In general, air pollution in the morning (i.e., 08:00–12:00) is much more severe with approximately 1.3 times higher mean $PM_{2.5}$ concentrations than those in the afternoon (i.e., 13:00–17:00) in China, related to the influences of varying BLHs (Li et al., 2017; Su et al., 2018).

[Please insert Figure 7 here]

Table 1 shows the diurnal $PM_{2.5}$ variations are also investigated over Eastern China and three typical urban agglomerations. It is clear that $PM_{2.5}$ pollution is generally higher in eastern China than the national level at each hour in a day due to more intensive human distribution and activities. In the BTH region, $PM_{2.5}$ pollution varies greatly with hourly $PM_{2.5}$ concentrations ranging from 28.88±10.16 μg/m^3 (10:00 a.m.) to 49.31±15.03 μg/m^3 (16:00 p.m.) at the diurnal level with a large difference exceeding 20 μg/m^3. In addition, $PM_{2.5}$ pollution always remained at a high level > 42 μg/m^3 before 12:00 pm, yet dropped to a low level < 29 μg/m^3 after 16:00 p.m. This is closely related to people's normal life and production and the life cycle of $PM_{2.5}$ in a day. Similar conclusions can be obtained in the YRD region. In general, the PRD region is less polluted in the morning but more severe in the afternoon than the BTH region in a day. By contrast, $PM_{2.5}$ pollution is much lower and shows a smaller diurnal difference with hourly $PM_{2.5}$ values ranging from 29.49±5.97 μg/m^3 (11:00 a.m.) to 36.36±5.76 μg/m^3 (08:00 a.m.) in the PRD region than other two key regions. This is mainly contributed to its better natural conditions and fewer pollutant emissions (Su et al., 2018).

[Please insert Table 1 here]

3.2.2 Seasonal and annual variations

Seasonal and annual $PM_{2.5}$ maps are synthesized from daily $PM_{2.5}$ maps in 2018 across China according to our previous approach (Wei et al., 2019a). Our results show that $PM_{2.5}$ pollution varies greatly on the seasonal scale, where the pollution levels are generally low and show similar spatial patterns in summer (~22.86±7.05 μg/m^3) and autumn (~23.76±10.97 μg/m^3) across China (Table S2). By contrast, it is

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

much more severe in spring (~32.84±11.49 μg/m^3) and winter (~39.04±16.32 μg/m^3) across China, especially for the BTH and YRD regions in winter. The main reasons are that the high-frequency sandstorm and the long-distance transmission of sand dust in spring, and the burning of coal and fossil fuels for heating in winter (Fang et al., 2016), leading to a lot of pollutant emissions in Northern China. In addition, $PM_{2.5}$ pollution shows significant spatial heterogeneities across China, and the annual mean $PM_{2.5}$ concentration is 28.99±10.31 μg/m^3 in 2018. The high pollution levels are always observed in the Hebei, Shandong, Jiangsu, Anhui, Henan, Hubei, and Sichuan provinces in China. These areas are mainly caused by the interactions of intensive human activities, adverse stagnant weather (e.g., low BLH and small winds) as well as special terrains (e.g., basin), which can increase the anthropogenic aerosols (Chen et al., 2008; Wang et al., 2018). By contrast, $PM_{2.5}$ pollution is relatively light in the northeast (e.g., Heilongjiang and Jilin), southwest (e.g., Tibet and Yunan), and eastern coastal (e.g., Zhejiang and Fujian) areas of China benefiting from sparse human distributions and activities or superior meteorological conditions for strong pollutant dispersion (Su et al., 2018).

[Please insert Figure 8 here]

3.3 Discussion

3.3.1 Comparison with traditional models

First, we compared the STLG model with five widely-used statistical regression models for estimating $PM_{2.5}$ in China with the same input data set (Table S3). The multivariate linear regression (MLR) model performs worst as demonstrated by the statistical indicators; this, due to the complex nonlinear $PM_{2.5}$-AOD relationships; while the GWR model shows a better performance because it takes into account the spatial characteristics of $PM_{2.5}$. The generalized additive model (GAM) and the LME model show overall improved performance with decreasing estimation uncertainties because of the nonlinear characteristics and stronger data regression abilities. By contrast, the two-stage model outperforms with higher CV-R^2 values and smaller estimation uncertainties by combining the advantages of the GWR and LME models. Nevertheless, our model shows the best performance than all of the above traditional statistical regression models, mainly due to its stronger data mining ability.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

Table 2 shows the comparisons of the accuracy and efficiency among six tree-based machine learning models to hourly $PM_{2.5}$ estimates in China using the same input data set. Decision Tree (DT, Quinlan, 1986) is a traditional supervised learning classification method that has been used frequently and shows the worst performance because of the simple single classifier, although the training speed is the fastest and the memory consumption is the least. The model performance of ensemble learning approaches, i.e., Gradient Boosting Decision Tree (GBDT, Friedman, 2001), RF (Breiman, 2001), Extremely randomized trees (ERT, Geurts, et al., 2006), and eXtreme Gradient Boosting (XGBoost, Chen & Guestrin, 2016), can be significantly improved by combining several weak classifiers into a strong classifier. Among them, the ERT model yields a higher estimation accuracy and a stronger spatial prediction ability than other ensemble learning models. By contrast, the LightGBM model (Ke et al., 2017) performs best with the highest accuracy and smallest uncertainty among all the tree-based machine learning approaches.

However, the model efficiency is different among these models due to the large differences in the algorithm design frameworks. These tree-based machine learning models can be divided into two categories: one is “Bagging”, including the DT, RF, and ERT models, which synthesizes multiple independent and unrelated weak classifiers into a strong classifier. It allows work in parallel, which can save a lot of time but may need more computer memory. The other is “Boosting”, including the GBDT, XGBoost, and LightGBM models, which synthesizes multiple interdependent and related weak classifiers into a strong classifier. They can only work in serial, which may take a lot of time but not too much memory. In general, the STGB model is the most time-consuming, while the STET model is the most memory-consuming approach. By contrast, the LightGBM model runs very fast and consumes very small computer memory benefiting from a series of algorithm optimizations (Ke et al., 2017). Furthermore, after considering the spatiotemporal variation characteristics, all the newly defined space-time tree-based machine learning approaches (i.e., STET, STGB, STXB, STRF, STET, and STLG) show significant improvements in both overall estimation accuracy and spatial prediction ability in hourly $PM_{2.5}$ concentrations with reference to their original models. This further illustrates the importance of the introduction of spatiotemporal information in constructing the $PM_{2.5}$–AOD relationships. More importantly, the training speed of these models did not decrease much; in addition,

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

the memory consumption did not increase much either. In general, the STLG model shows the best performance with much high efficiency (i.e., training speed = 46s, memory usage = 0.60 GB) among all the space-time tree-based machine learning models. Therefore, our newly STLG model is highly valuable for accurate and fast air pollution monitoring, in particular for our future extended study to the global scale.

[Please insert Table 2 here]

3.3.2 Comparison with related studies

Last, we also compared with previous related studies in Himawari-8-based hourly $PM_{2.5}$ estimates at regional and national scales in China (Table 3). The comparison results show that the local hourly $PM_{2.5}$ concentrations retrieved from our national-scale model are more accurate than those derived from the models developed separately in local areas, e.g., the LME model (Wang et al., 2017), and the GWR, SVR, RF, and DNN models in the BTH region (Sun et al., 2019); the two-stage RF and DNN models in the YRD region (Fan et al., 2020; Tang et al., 2019). In addition, our model outperforms most of the statistical regression models, machine learning models focusing on entire China, e.g., the I-LME, and IGTWR, RF, Adaboost, XGBoost, and their stacked models in China (Chen et al., 2019; Liu et al., 2019; Xue et al., 2020; Zhang et al., 2019). The main reasons include the stronger data mining ability of our model, and the consideration of the key spatial and temporal information of air pollution that ignored in previous studies, and the introduction of more comprehensive factors (e.g., emission inventory) that affect $PM_{2.5}$ pollution.

[Please insert Table 3 here]

4. Summary and conclusion

$PM_{2.5}$ has great impacts on the atmospheric environment, and is also used as a key indicator in environmental health studies. It shows strong diurnal variations affected by both natural and human factors; however, previous studies are based on sun-synchronous satellites, which can only monitor air

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

pollution at coarser temporal (daily) scales. Thus, high-temporal-resolution $PM_{2.5}$ dataset with high precision are urgently needed. In this study, the Himawari-8/AHI hourly AOD products are employed to tackle this issue. In addition, considering the large volume of the input data and the large estimation errors of $PM_{2.5}$ for traditional methods, a new efficient and accurate space-time Light gradient boosting machine (i.e., STLG) model is developed, and a 5 km resolution hourly $PM_{2.5}$ dataset in China is generated together with the meteorological, human, land use, and topographical factors. The hourly $PM_{2.5}$ estimates are evaluated against surface observations, and the $PM_{2.5}$ spatiotemporal variations are also investigated.

The STLG model derives and predicts the hourly $PM_{2.5}$ values accurately with high out-of-sample (out-of-station) CV-R^2 values of ~0.81–0.85 (~0.76–0.81), and low RMSE values of ~11.24–15.56 (~12.49–17.61) μg/m^3 throughout the daytime. In addition, it can well capture the daily (e.g., R^2 = 0.91, RMSE = 10.11 μg/m^3), as well as monthly, seasonal and annual $PM_{2.5}$ values (e.g., R^2 = 0.98, RMSE = 1.6–3.3 μg/m^3). $PM_{2.5}$ shows significant diurnal variabilities in most areas across mainland China, where the $PM_{2.5}$ concentrations reach the maximum at 10 a.m. but are generally low at sunrise or sunset in a day. Moreover, $PM_{2.5}$ also varies greatly on a seasonal basis, where winter and summer experience the highest and lowest air pollution, respectively. Comparison results suggest that our model is more accurate than the traditional statistical regression, other tree-based machine learning, as well as those models developed by previous studies. In addition, it is more efficient with faster training speed and less memory consumption. These results illustrate that our algorithm has a strong applicability value to the real-time monitoring of $PM_{2.5}$ pollution in China.

Data availability

The $PM_{2.5}$ measurements are available at http://www.cnemc.cn, and the Himawari-8 AOD product is available at ftp.ptree.jaxa.jp, and the ERA5 reanalysis products are available at https://cds.climate.copernicus.eu/, and the MODIS product is available at https://search.earthdata.nasa.gov/, and the LandScan™ product is available at https://landscan.ornl.gov/.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

Author contribution

JW designed the research and wrote the initial draft of this manuscript. ZL, RP, and LS helped review and edit the paper. RL and WX helped process the data. All authors made substantial contributions to this work.

385

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

390 We would like to thank Dr. Qiang Zhang at Tsinghua University for providing MEIC pollution emission data in China.

Financial support

This research has been supported by the National Key R&D Program of China (2017YFC1501702), and
395 the National Natural Science Foundation of China (42030606).

References

An, Z., Huang, R. J., Zhang, R., Tie, X., Li, G., Cao, J., et al. Severe haze in northern China: a synergy of anthropogenic emissions and atmospheric processes. Proceedings of the National Academy of
400 Sciences, 116(18), 8657-8666, 2019.

Bessho, K., Date, K., Hayashi, M., Ikeda, A., & Yoshida, R. An introduction to Himawari-8/9— Japan's new-generation geostationary meteorological satellites. Journal of the Meteorological Society of Japan, 2016, 94(2), 151-183, 2016.

Behrens, T., Schmidt, K., Viscarra, R., Gries, P., Scholten, T., & Macmillan, R. Spatial modelling with
405 Euclidean distance fields and machine learning. European Journal of Soil Science, 69, 2018.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

Baez-Villanueva, O., Zambrano-Bigiarini, M., Beck, H., Mcnamara, I., & Thinh, N. RF-MEP: A novel random forest method for merging gridded precipitation products and ground-based measurements. Remote Sensing of Environment, 239(111606), 2020.

Breiman, L. Random forests. Machine Learning, 45, 5–32, 2001.

Chan, C., & Yao, X. Air pollution in megacities in China. Atmospheric Environment, 42(1), 1-42, 2008.

Chen, J., Yin, J., Zang, L., Zhang, T., & Zhao, M. Stacking machine learning model for estimating hourly $PM_{2.5}$ in China based on Himawari 8 aerosol optical depth data. Science of the Total Environment, 697, 134021, 2019.

Chen, T., & Guestrin, C. XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794, 2016.

Chen, Z., Cheng, S., Li, J., Guo, X., Wang, W., & Chen, D. Relationship between atmospheric pollution processes and synoptic pressure patterns in northern China. Atmospheric Environment, 42(24), 6078–6087, 2008.

Delfino, R.J., Sioutas, C., & Malik, S. Potential role of ultrafine particles in associations between airborne particle mass and cardiovascular health. Environmental Health Perspectives, 113(8), 934-946, 2005.

Dobson, J., Bright, E., Coleman, P., Durfee, R., & Worley, B. A Global Population database for Estimating Populations at Risk. Photogrammetric Engineering & Remote Sensing, 66(7), 2000.

Fan, W., Qin, K., Cui, Y., Li, D., & Bilal, M. Estimation of hourly ground-level $PM_{2.5}$ concentration based on Himawari-8 apparent reflectance. IEEE Transactions on Geoscience and Remote Sensing, https://doi.org/10.1109/TGRS.2020.2990791, 2020.

Friedman, J. Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232, 2001.

Geurts, P., Ernst, D., & Wehenkel, L. Extremely randomized trees. Machine Learning, 63, 3−42, 2006.

Giles, D., Sinyuk, A., Sorokin, M., Schafer, J., Smirnov, A., Slutsker, I., Eck, T., Holben, B., Lewis, J., Campbell, J., Welton, E., Korkin, S., & Lyapustin, A. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmos. Meas. Tech. 12, 169–209, 2019.

Guo, J., Zhang, X., Che, H., Gong, S., An, X., Cao, C., et al. Correlation between PM concentrations and aerosol optical depth in eastern China. Atmospheric Environment, 43(37), 5876–5886, 2009.

Huang, R., Zhang, Y., …, & Prévôt, André S. High secondary aerosol contribution to particulate pollution during haze events in China. Nature, 514, 218–222, 2014.

Jacob, D., & Winner, D. Effect of climate change on air quality. Atmospheric Environment, 43(1), 51-63, 2009.

Kampa, M., & Castanas, E. Human health effects of air pollution. Environmental Pollution, 151(2), 362-367, 2008.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. LightGBM: A highly efficient gradient boosting decision tree. In Advances in Neural Information Processing Systems, 3149–3157, 2017.

Kim, K., Kabir, E., & Kabir, S. A review on the human health impact of airborne particulate matter. Environment International, 74, 136-143, 2015.

Lelieveld, J., Evans, J., Fnais, M., Giannadaki, D., & Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature, 525, 367–371, 2015.

Letu, H., Yang, K., Nakajima, T. Y., Ishimoto, H., Nagao, T. M., Riedi, J., ... & Khatri, P. High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/AHI next-generation geostationary satellite. Remote Sensing of Environment, 239, 111583, 2020.

Li, M., Zhang, Q., Kurokawa, J., Woo, J., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D., Carmichael, G., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., & Zheng, B. MIX. a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. Atmospheric Chemistry and Physics, 17, 935-963, 2017.

Li, Z., Xu, H., Li, K., Li, D., Xie, Y., Li, L., Zhang, Y., Gu, X., Zhao, W., Tian, Q., Deng, R., Su, X., Huang, B., Qiao, Y., Cui, W., Hu, Y., Gong, C., Wang, Y., Wang, X., Wang, J., Du, W., Pan, Z., Li, Z., & Bu, D. Comprehensive study of optical, physical, chemical, and radiative properties of

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

total columnar atmospheric aerosols over China: an overview of Sun–Sky Radiometer Observation Network (SONET) measurements. Bull. Amer. Meteorol. Soc. 99, 739–755, 2018.

Li, Z., Guo, J., Ding, A., Liao, H., Liu, J., Sun, Y., Wang, T., Xue, H., Zhang, H., & Zhu, B. Aerosols and boundary-layer interactions and impact on air quality, National Science Review, 4, 810-833, 2017.

Li, Z., Wang, Y., Guo, J., Zhao, C., Cribb, M., Dong, X., et al. East Asian Study of Tropospheric Aerosols and their Impact on Regional Clouds, Precipitation, and Climate (EAST-AIRCPC), Journal of Geophysical Research: Atmospheres, 124, 2019.

Liu, J., Weng, F., Li, Z., & Cribb, M., Hourly $PM_{2.5}$ Estimates from a geostationary satellite based on an ensemble learning algorithm and their spatiotemporal patterns over Central East China. Remote Sensing, 2019, 11(18), 2120.

Liu, Y., Franklin, M, Kahn, R., & Koutrakis, P. Using aerosol optical thickness to predict ground-level $PM_{2.5}$ concentrations in the St. Louis area: a comparison between MISR and MODIS. Remote Sensing of Environment, 107(1-2), 33-44, 2007.

Liu, Y., Sarnat, J., Kilaru, V., Jacob, D., & Koutrakis, P. Estimating ground-level $PM_{2.5}$ in the eastern united states using satellite remote sensing. Environmental Science & Technology, 39(9), 3269-78, 2005.

Ma, Z., Hu, X., Huang, L., Bi, J., & Liu, Y. Estimating ground-level $PM_{2.5}$ in China using satellite remote sensing. Environmental Science and Technology, 48(13), 7436–7444, 2014.

Quinlan, J. Induction on decision tree. Machine Learning, 1, 81–106, 1986.

Ramanathan, V., & Feng, Y. Air pollution, greenhouse gases and climate change: global and regional perspectives. Atmospheric Environment, 43(1), 37-50, 2009.

Shi, H. Best-first decision tree learning. PhD thesis, The University of Waikato, 2007.

Su, T., Li, Z., & Kahn, R. Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: regional pattern and influencing factors. Atmospheric Chemistry and Physics, 18, 15921–15935, 2018.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021


Atmospheric Chemistry and Physics Discussions

Sun, Y., Zeng, Q., Geng, B., Lin, X., Sude, B., & Chen, L. Deep learning architecture for estimating hourly ground-level $PM_{2.5}$ using satellite remote sensing. IEEE Geoscience and Remote Sensing Letters, 16(9), 1343–1347, 2019.

Sun, Y., Zhuang, G., Wang, Y., Han, L., Guo, J., Dan, M., et al. The air-borne particulate pollution in Beijing—concentration, composition, distribution and sources. Atmospheric Environment, 38(35), 5991-6004, 2004.

Tang, D., Liu, D., Tang, Y., Seyler, B., Deng, X., & Zhan, Y. Comparison of GOCI and Himawari-8 aerosol optical depth for deriving full coverage hourly $PM_{2.5}$ across the Yangtze River Delta. Atmospheric Environment, 217, 116973, 2019.

van Donkelaar, A., Martin, R., & Park, R. Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing. Journal of Geophysical Research: Atmospheres, 111(D21), 2006.

Wang, W., Mao, F., Du, L., Pan, Z., Gong, W., and Fang, S. Deriving hourly $PM_{2.5}$ concentrations from Himawari-8 AODs over Beijing-Tianjin-Hebei in China. Remote Sensing, 9(8), 858, 2017.

Wang, X., Dickinson, R., Su, L., Zhou, C., and Wang, K. $PM_{2.5}$ pollution in China and how it has been exacerbated by terrain and meteorological conditions. Bulletin of the American Meteorological Society, 99 (1), 105–119, 2018.

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 Sensing of Environment, 231, 111221, 2019a.

Wei, J., Li, Z., Sun, L., Peng, Y., Zhang, Z., Li, Z., Su, T., Feng, L., Cai, Z., & Wu, H. Evaluation and uncertainty estimate of the next-generation geostationary meteorological Himawari-8/AHI aerosol products. Science of the Total Environment, 692, 879–891, 2019b.

Wei, J., Li, Z., Peng, Y., & Sun, L. MODIS Collection 6.1 aerosol optical depth products over land and ocean: validation and comparison. Atmospheric Environment, 201, 428–440, 2019c.

Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., & Song, Y. Improved 1 km resolution $PM_{2.5}$ estimates across China using enhanced space-time extremely randomized trees, Atmospheric Chemistry and Physics, 20(6), 3273–3289, 2020.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M. Reconstructing 1-km-resolution high-quality $PM_{2.5}$ data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. Remote Sensing of Environment, 252, 112136, 2021.

Wu, J., Yao, F., Si, M., & Li, W. VIIRS-based remote sensing estimation of ground-level $PM_{2.5}$ concentrations in Beijing–Tianjin–Hebei: a spatiotemporal statistical model. Remote Sensing of Environment, 184, 316–328, 2016.

Xin, J., Wang, Y., Pan, Y., Ji, D., Liu, Z., et al. The campaign on atmospheric aerosol research network of China. Bulletin of the American Meteorological Society, 96 (7), 1137–1155, 2015.

Xue, Y., Li, Y., Guang, J., Tugui, A., …, and Wang, Z. Hourly $PM_{2.5}$ estimation over central and eastern China based on himawari-8 data. Remote Sensing, 12(5), 855, 2020.

Yao, F., Wu, J., Li, W., & Peng, J. A spatially structured adaptive two-stage model for retrieving ground-level $PM_{2.5}$ concentrations from VIIRS AOD in China. ISPRS Journal of Photogrammetry and Remote Sensing, 151, 263–276, 2019.

Yoshida, M, Kikuchi, M., Nagao, T., Murakami, H., Nomaki, T., & Higurashi, A. Common retrieval of aerosol properties for imaging satellite sensors. Journal of the Physical Society of Japan, 96b, 193–209, 2018.

Zhang, Q., Streets, D., He, K., & Klimont, Z. Major components of China's anthropogenic primary particulate emissions. Environmental Research Letters, 2, No. 045027, 2007.

Zhang, Q., Zheng, Y., Tong, D., Shao, M., & Hao, J. Drivers of improved $PM_{2.5}$ air quality in china from 2013 to 2017. Proceedings of the National Academy of Sciences, 201907956, 2019.

Zhang, T., Zang, L., Wan, Y., Wang, W., and Zhang, Y. Ground-level $PM_{2.5}$ estimation over urban agglomerations in China with high spatiotemporal resolution based on Himawari-8. Science of the total environment. Science of the Total Environment, 676, 535–544, 2019.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions

Table 1. Hourly mean $PM_{2.5}$ concentrations ($\mu g/m^3$) in 2018 in China, and each region.

Time	China	ECHN	BTH	YRD	PRD
08:00	29.94±10.91	31.97±11.55	42.46±12.97	38.60±10.57	29.34±5.01
09:00	33.37±12.59	36.29±13.52	47.32±49.31	43.55±11.27	34.81±5.46
10:00	35.67±13.53	38.56±14.05	49.31±15.03	44.72±11.17	35.48±5.47
11:00	35.63±13.05	38.72±13.53	49.10±13.77	44.27±10.55	36.36±5.76
12:00	31.23±11.74	35.10±12.47	42.38±12.86	41.37±9.77	34.56±5.72
13:00	28.45±11.40	32.23±11.73	37.70±11.55	39.36±9.22	33.33±5.48
14:00	26.36±11.18	30.14±11.09	34.32±11.81	37.31±8.59	32.05±5.50
15:00	24.25±10.06	28.67±10.21	31.95±11.26	36.77±8.13	30.34±5.43
16:00	23.63±9.26	27.38±9.15	29.82±10.13	32.84±6.30	29.49±5.97
17:00	23.21±9.73	26.63±8.93	28.88±10.16	27.59±4.39	31.56±6.17
Morning	33.29±11.59	36.15±12.41	46.12±13.29	42.50±10.22	34.52±4.63
Afternoon	25.11±9.78	29.01±9.70	32.53±10.53	34.76±6.66	31.42±4.85

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions

Table 2. Comparison in model performance between different tree-based machine learning models using the same input data. Data are from 14:00 local time in 2018 in China (N = 162,840).

Model	Sample-based validation			Station-based validation			Speed (s)	Memory (GB)
	R^2	RMSE	MAE	R^2	RMSE	MAE		
DT	0.52	25.53	14.80	0.48	27.03	15.57	6	0.58
GBDT	0.65	20.03	13.17	0.61	21.20	14.10	94	0.59
XGBoost	0.73	17.94	10.78	0.68	19.59	11.93	456	0.69
RF	0.72	17.86	11.33	0.69	18.80	11.95	165	2.59
ERT	0.74	17.12	10.87	0.72	18.01	11.49	54	3.69
LightGBM	0.78	15.79	9.84	0.73	17.59	11.21	34	0.60
STDT	0.65	21.09	12.33	0.63	22.00	12.85	8	0.60
STGB	0.75	16.82	10.93	0.73	17.61	11.54	503	0.61
STXB	0.82	14.73	8.76	0.78	15.92	9.62	456	0.68
STRF	0.81	14.62	9.17	0.79	15.44	9.69	219	2.75
STET	0.82	14.42	8.95	0.80	15.30	9.55	77	4.25
STLG	0.85	13.09	8.11	0.81	14.63	9.29	46	0.60

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions

545 **Table 3.** Comparison in model performance with previous studies in hourly $PM_{2.5}$ estimates in China.

Model	Model validation			Region	Reference
	R^2	RMSE	MAE		
LME	0.86	24.5	14.2	BTH	Wang et al., 2017
LME	0.63	29.0	18.1	BTH	Sun et al., 2019
GWR	0.76	23.3	16.7		
SVR	0.77	21.5	12.3		
RF	0.82	20.3	12.1		
DNN	0.84	19.9	11.9		
two-stage RF	0.86	12.4	-	YRD	Tang et al., 2019
DNN	0.86	14.3	-	YRD	Fan et al., 2020
RF	0.82	19.6	12.2	China	Chen et al., 2019
Adaboost	0.84	18.3	10.7		
XGBoost	0.84	18.1	11.4		
Stacked model	0.85	17.3	10.5		
RF	0.86	17.3	10.3	China	Liu et al., 2019
I-LME	0.84	-	-	BTH	Zhang et al., 2019
	0.80	-	-	YRD	
	0.74	-	-	PRD	
	0.82	-	-	China	
IGTWR	0.78	21.1	-	China	Xue et al., 2020

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions

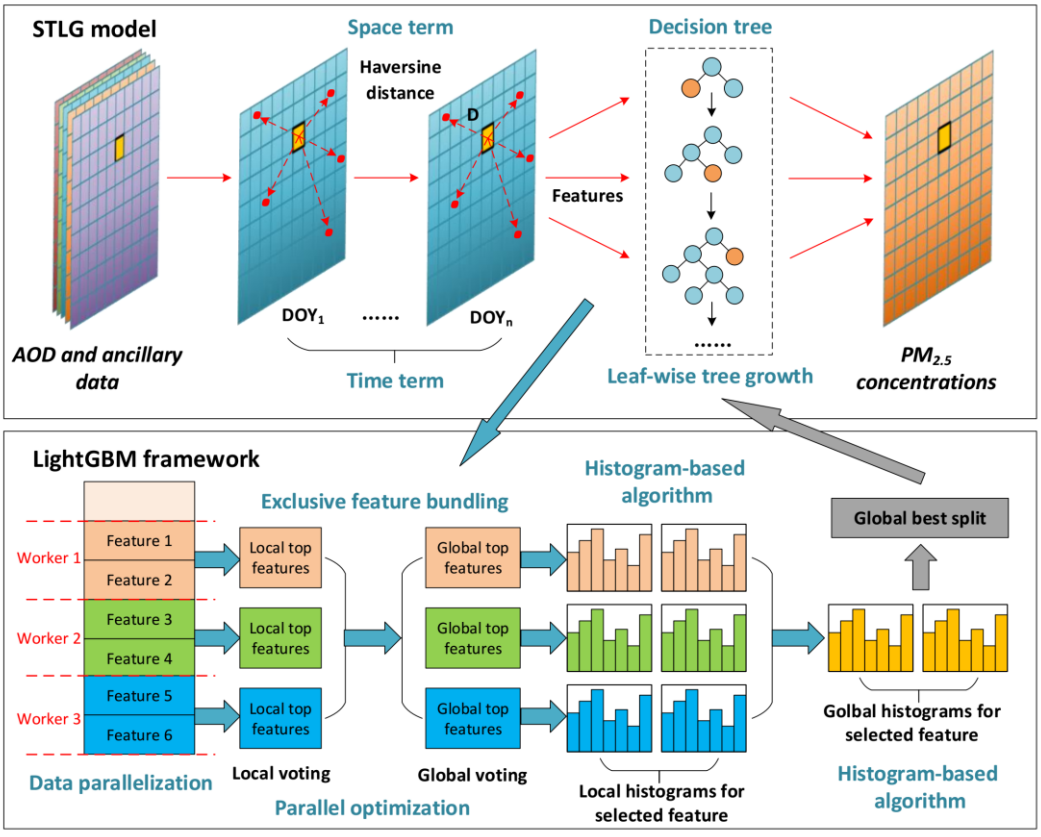


Figure 1. Schematics of the space-time LightGBM (STLG) model developed in this study.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

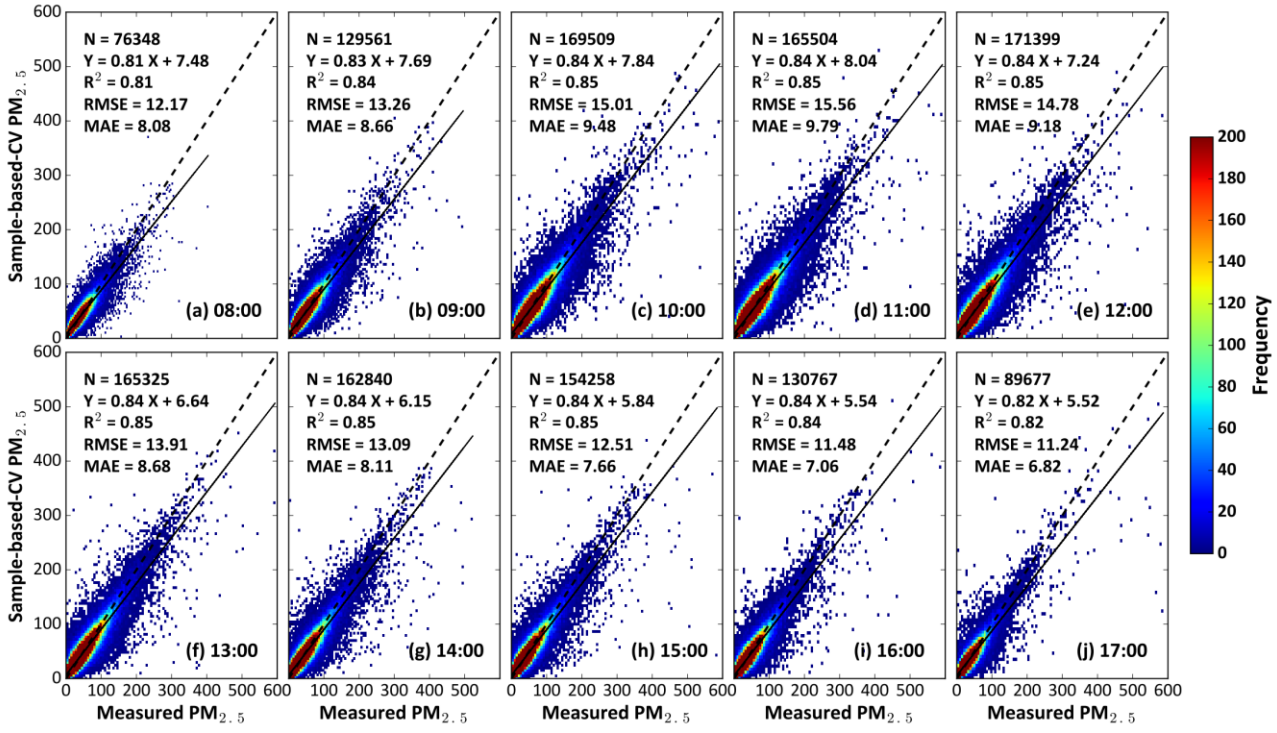
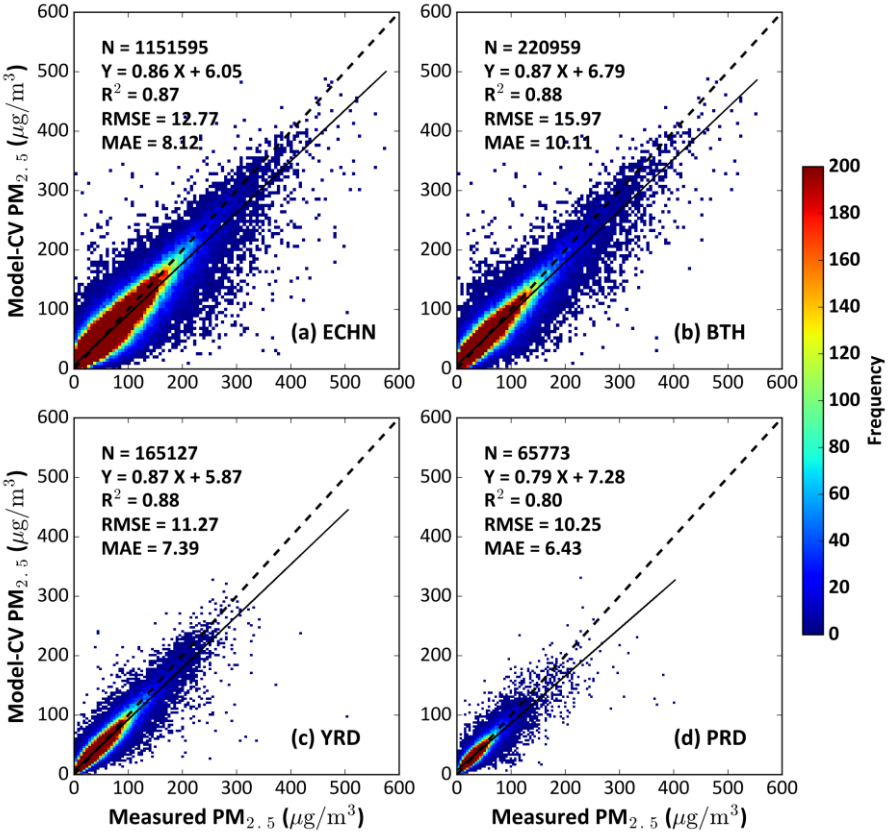


Figure 2. Out-of-sample cross-validation results of hourly $PM_{2.5}$ estimates ($\mu g/m^3$) from 08:00 to 17:00 local time in 2018 across China.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions



555 **Figure 3.** Out-of-sample cross-validation results of regional hourly $PM_{2.5}$ estimates (μg/m^3) in 2018 in China.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions

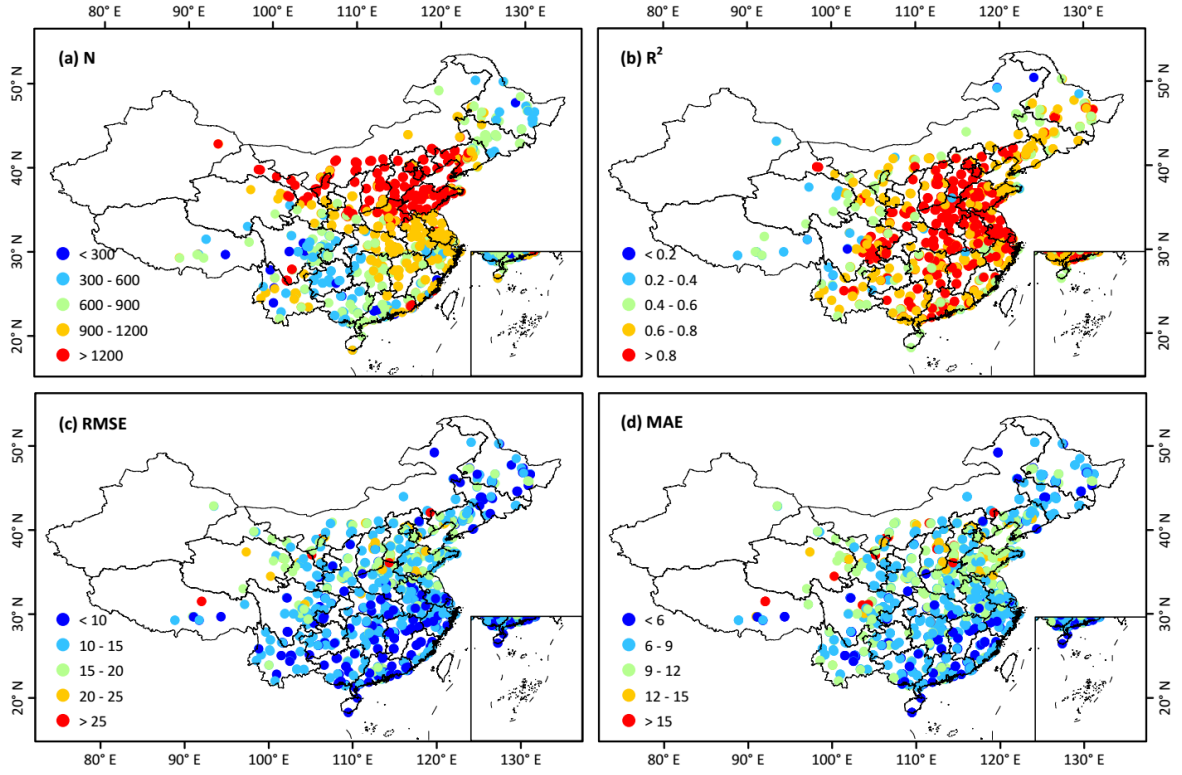


Figure 4. Individual-site-scale validation of hourly $PM_{2.5}$ estimates ($\mu g/m^3$) in 2018 in China.

560

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

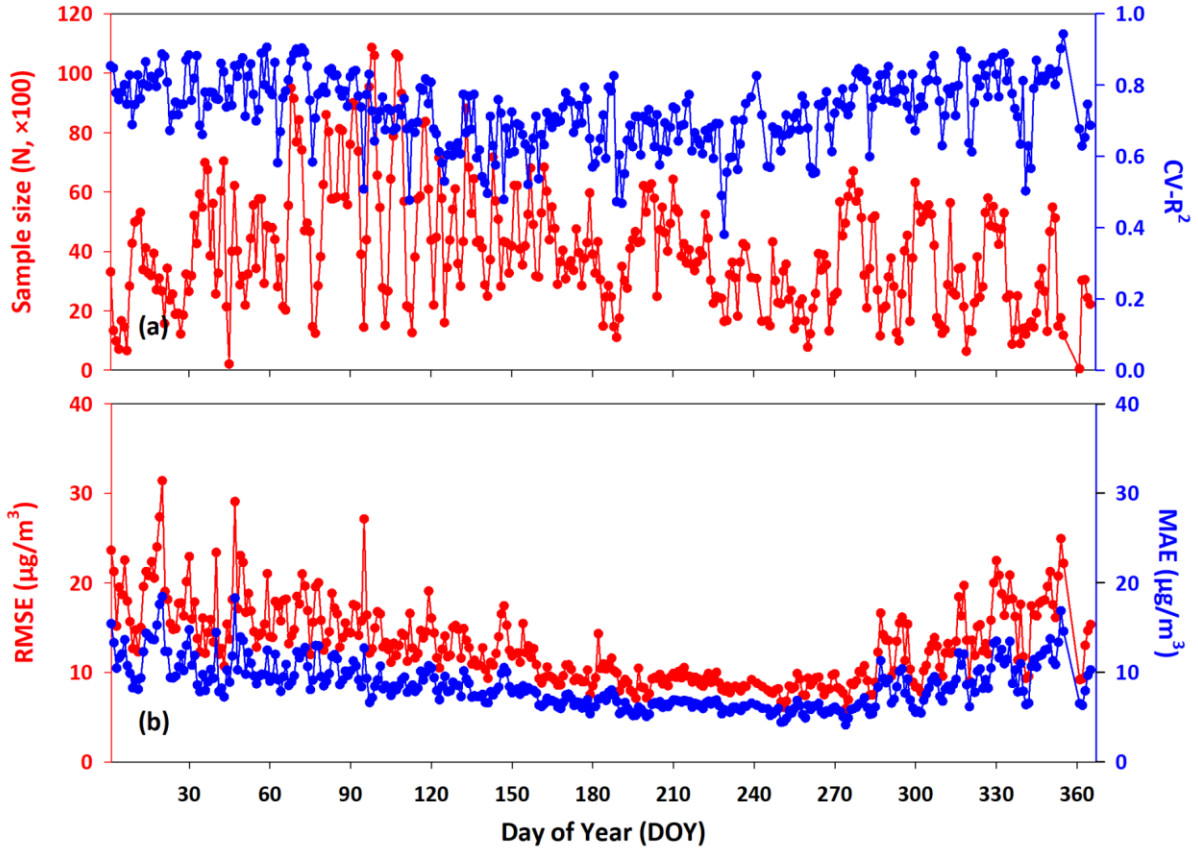
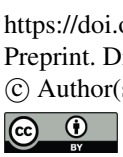
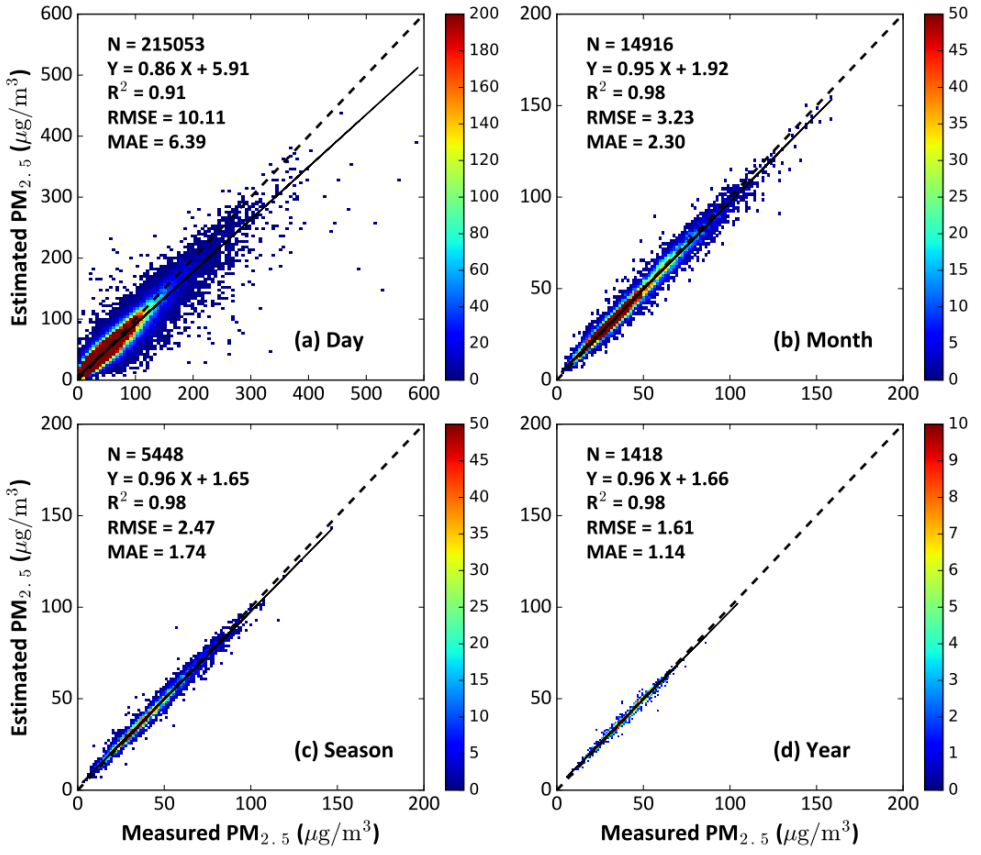


Figure 5. Time series of validation of hourly $PM_{2.5}$ estimates ($\mu g/m^3$) in 2018 across China.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021



Atmospheric Chemistry and Physics Discussions



565 **Figure 6.** Density scatter plots of validation results of (a) daily, (b) monthly, (c) seasonal, and (d) annual mean $PM_{2.5}$ estimates in 2018 across China.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

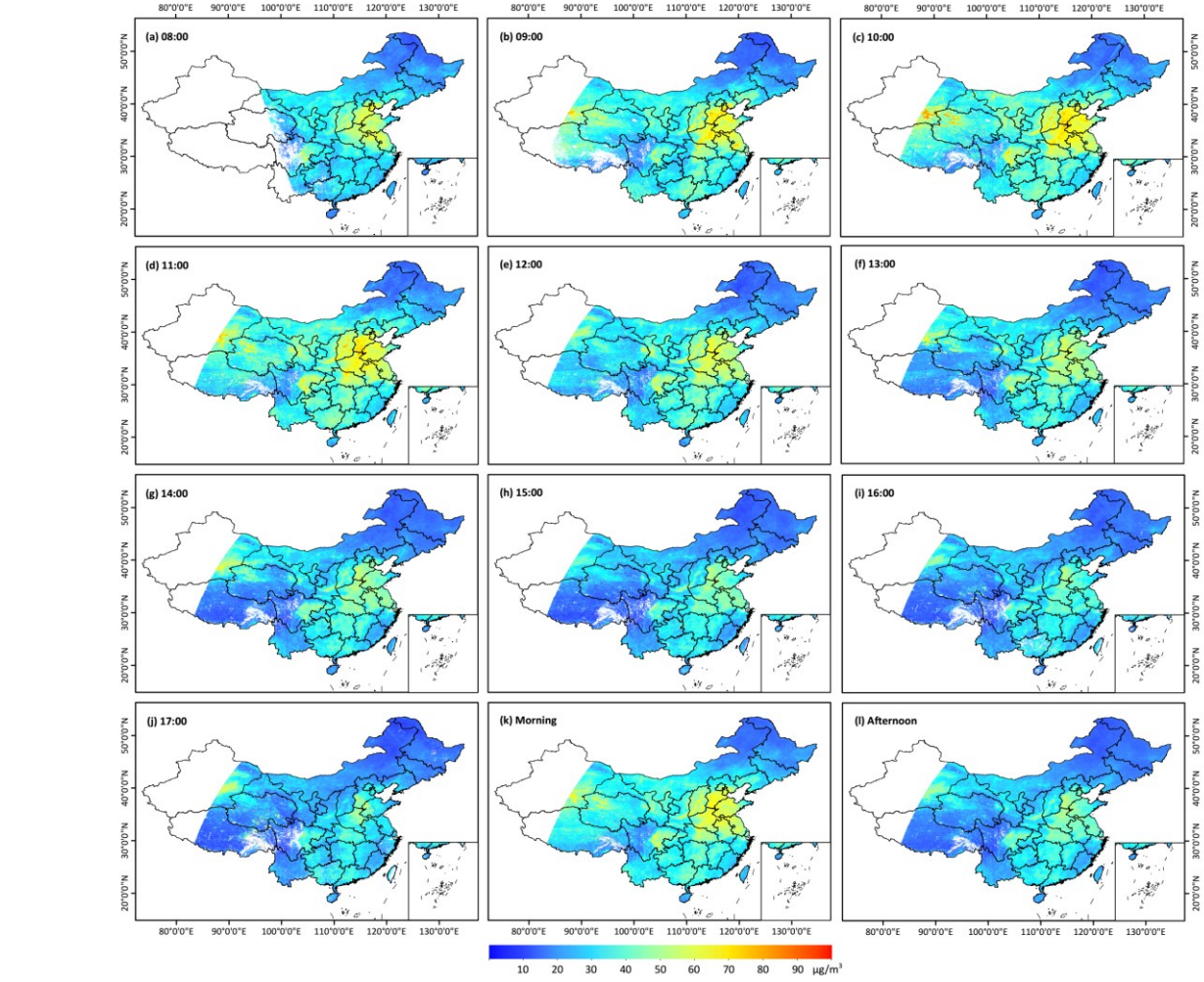


Figure 7. Himawari-8-derived hourly mean $PM_{2.5}$ maps (5 km) from (a-j) 08:00 to 17:00 local time, and in (k) morning (08:00–12:00), and (l) afternoon (13:00–17:00) in 2018 in China.

https://doi.org/10.5194/acp-2020-1277
Preprint. Discussion started: 6 January 2021

Atmospheric Chemistry and Physics Discussions

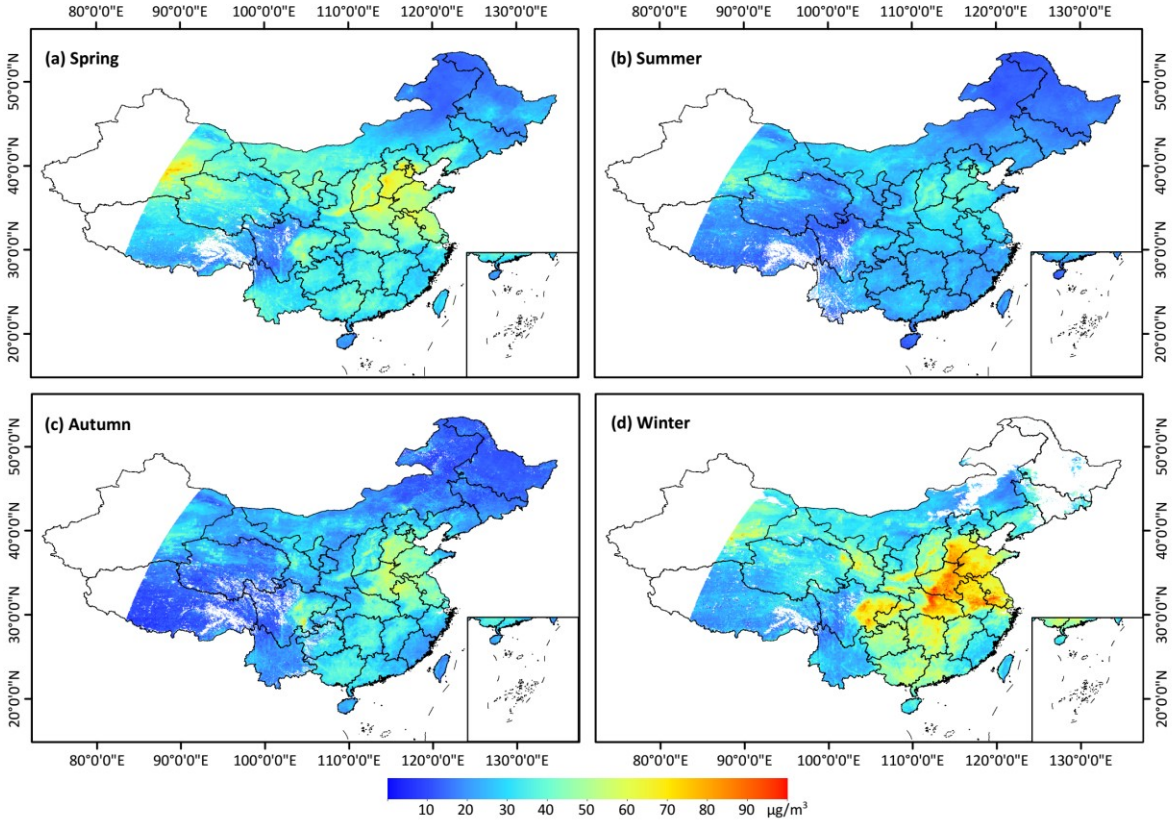


Figure 8. Himawari-8-derived seasonal mean $PM_{2.5}$ maps (5 km) in 2018 across China.