Article

High-Resolution Mapping of Aerosol Optical Depth and Ground Aerosol Coefficients for Mainland China

Lianfa Li 1,2

1 State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; lilf@igsnrr.ac.cn; Tel.: +86-10-64888362
2 College of Resources and Environment, University of the Chinese Academy of Sciences, Beijing 100049, China

Abstract: Aerosols play an important role in climate change, and ground aerosols (e.g., fine particulate matter, abbreviated as PM 2.5) are associated with a variety of health problems. Due to clouds and high reflectance conditions, satellite-derived aerosol optical depth (AOD) products usually have large percentages of missing values (e.g., on average greater than 60% for mainland China), which limits their applicability. In this study, we generated grid maps of high-resolution, daily complete AOD and ground aerosol coefficients for the large study area of mainland China from 2015 to 2018. Based on the AOD retrieved using the recent Multi-Angle Implementation of Atmospheric Correction advanced algorithm, we added a geographic zoning factor to account for variability in meteorology, and developed an adaptive method based on the improved full residual deep network (with attention layers) to impute extensively missing AOD in the whole study area consistently and reliably. Furthermore, we generated high-resolution grid maps of complete AOD and ground aerosol coefficients. Overall, compared with the original residual model, in the independent test of 20% samples, our daily models achieved an average test R 2 of 0.90 (an improvement of approximately 5%) with a range of 0.75–0.97 (average test root mean square error: 0.075). This high test performance shows the validity of AOD imputation. In the evaluation using the ground AOD data from six Aerosol Robotic Network monitoring stations, our method obtained an R 2 of 0.78, which further illustrated the reliability of the dataset. In addition, ground aerosol coefficients were generated to provide an improved correlation with PM 2.5. With the complete AOD data and ground coefficients, we presented and interpreted their spatiotemporal variations in mainland China. This study has important implications for using satellite-derived AOD to estimate aerosol air pollutants.

Keywords: aerosol optical depth; ground aerosol coefficient; simulation; adaptive full residual deep network; mainland China; reliability

1. Introduction

Aerosols are tiny particles in the air that come from many sources, including anthropogenic sources, such as coal, petroleum, burning wood and biofuels; and natural sources, such as the desert, wild fires and dust [1]. They can control the energy from the Sun to the Earth, and scatter the energy back into space, thereby considerably influencing the climate [2,3]. Ground aerosols, such as coarse particulate matter (PM 10) or fine particular matter (PM 2.5), also have adverse health effects, ranging from asthma [4], stroke [5] and heart disease [6], to pregnancy complications [7] and mortality [8]. Recent studies [9–11] also show that aerosol droplets may be an important route of human-to-human transmission of COVID-19. The spatiotemporal distributions and variations of aerosols are critical for studying climate change, and for estimations of exposure for health evaluations. Although a worldwide or countrywide ground network for monitoring aerosols, such as sampling stations of PM 10, PM 2.5 or Aerosol Robotic Network (AERONET) aerosol optical depth (AOD), has been established, due to limited spatial coverage, such a network is insufficient for monitoring and measuring spatiotemporal variation of aerosols. For...
example, there are approximately 1500 monitoring stations of PM$_{2.5}$ in mainland China. Many of these stations are located in urban rather than rural areas and there are large gap areas between them. They are limited and insufficient to capture spatiotemporal variability of PM$_{2.5}$ in mainland China where the altitude, topography, meteorology and geographic environment vary greatly.

With the global geospatial coverage and a regular temporal resolution, remote sensing provides AOD products to measure and monitor aerosols worldwide. The remote sensors of AOD products [12] range from the early Advanced Very High Resolution Radiometer (AVHRR), the Total Ozone Mapping Spectrometer (TOMS) and the Sea-viewing Wide Field of View Sensor (SeaWiFS); to the Global Ozone Monitoring Experiment (GOME), the Moderate Resolution Imaging Spectroradiometer (MODIS), the Multispectral Imaging Spectroradiometer (MISR), the Polarization and Directionality of the Earth’s Reflectances (POLDER), Ozone Mapping Imaging (OMI) and GOME-2; to the recent Visible Infrared Imaging Radiometer (VIIRS), Advanced Baseline Imager (ABI) and Advanced Himawari Imager (AHI). Depending on the types and characteristics of the sensors, different retrieval algorithms have been developed, including contrast reduction [13], Dark Target (DT) [14], Deep Blue (DB) [15,16], Multi-Angle Implementation of Atmospheric Correction (MAIAC) [17,18], the Simplified Aerosol Retrieval Algorithm (SARA) [19], polarization [20] and machine learning technology [21]. For the retrieval algorithms of AOD, the complexity of the Earth’s surface, along with clouds and high reflectance conditions, has important influences on their quality and uncertainty, especially when there are numerous gaps in the data [22].

Given its wide field of view, MODIS provides daily global observations of the Earth and has been widely used as a mainstream AOD sensor. Fan et al. [12] showed a higher correlation between DT and DB AOD retrieved from MODIS than AOD retrieved by many other sensors. As the improved version of the retrieval algorithm of MODIS AOD, MAIAC is designed to perform simultaneous retrieval of aerosol properties and surface bidirectional reflectance. Compared with DT and DB, MAIAC can work over both dark and bright surfaces, and retrieve AOD at a higher spatial resolution (1 km). However, similar to other AOD products, the retrieval of satellite-borne AOD depends on cloudless skies and suitable ground surfaces. Due to cloud cover [23], snow [24] and other high reflectance surfaces [25], AOD retrieved by MAIAC also has a large proportion of missing values, which are actually non-random [26]. For example, on average, the proportion of missing AOD in mainland China typically exceeds 60% each day [27]. Such a high missing proportion severely affects the applicability of the AOD data products.

AOD is defined as the vertical integral of the aerosol extinction coefficient at all altitudes in the orientation between the ground and the top of the atmosphere, and the aerosol’s vertical profile presents an uneven and probably exponential distribution; PM$_{2.5}$ is particulate matter near the ground, with diameters not exceeding more or less 2.5 $\mu$m. Derived from AOD, the aerosol extinction coefficient on the ground can be used as an indirect yet better proxy variable to represent the distribution of ground aerosols such as PM$_{10}$ and PM$_{2.5}$. Therefore, AOD cannot be directly used as a good predictive indicator for PM$_{2.5}$. In order to predict the concentration of PM$_{10}$ or PM$_{2.5}$, AOD can be converted into the ground aerosol extinction coefficient (simplified to ground aerosol coefficient, abbreviated as GAC).

Imputation is “assignment by inference” for the spatial gaps in the satellite data. For imputation of satellite-borne AOD, spatial, regression, or in combination, models have been developed, including early nearest neighbors [30], changes to the original retrieval algorithm [31,32], linear regression plus Kriging [33], generalized additive model (GAM) [27], and feed-forward neural network [34]. However, these methods have limited test performance and generalization, as shown by their low validation $R^2$ (0.18–0.44) with AERONET AOD. Mainland China has high variability in topography, meteorology and aerosol emission sources, and these methods are limited for imputation of a large study region, such as
mainland China. Compared with the traditional regression spatial statistics or machine learning methods, modern deep learning methods present powerful performance and generalization. The high missing proportion of AOD makes convolutional neural networks (CNN) inappropriate. Recently, a full residual deep network [35] was developed as a novel deep learning method to impute missing AOD with high generalization and natural spatial smoothing in surface modelling. However, for a large area such as mainland China, high variability in meteorology, geography and altitude needs to be considered in imputation of missing AOD.

In this study, we generated spatial maps of high-resolution (1 × 1 km$^2$), daily complete AOD and ground aerosol extinction coefficients of mainland China from 2015 to 2018. Using MAIAC AOD as the base data, we developed an improved full residual attention deep learning method to reliably impute extensive missing AOD and produce the high-resolution complete AOD dataset. Taking into account the important impact of meteorology on AOD and its missing values, but due to the lack of publicly available high-resolution parameters, we combined ground monitoring data with meteorological reanalysis data, topography, coordinates and time index to generate reliable high spatiotemporal resolution surfaces of meteorological factors.

In order to account for high variability in meteorology and land-use for a large area such as mainland China, the geographic zones were also included as a one-hot indicator through seven binary variables in the model; the attention layers were added to the full residual network to enhance the influence of important predictive covariates and reduce the bias from uncorrelated confounding covariates [36] for imputation in a large area. In deep learning, the weight parameters in the attention layer are used to allow the model to focus and place more “attention” (high weights) on the important parts rather than the minor or noisy parts (low weights) in the input sequence or nodes as needed [37], and it has been widely applied to solve complex problems of natural language processing [38] or computer vision [39].

In order to improve the effectiveness of model training, we developed the adaptive method of heuristic searching to find optimal hyperparameters (minibatch size, learning rate, number of hidden layers and number of nodes) to obtain a better model: when the test performance was lower than the threshold (for example, $R^2 < 0.75$), the system can search for optimal hyperparameters to retrain the (changed) model, thereby improving test performance and generalization ability. In deep learning, these hyperparameters have important influence upon learning and testing [40]. As shown in this study, the adaptive method with the incorporation of the regional factor and attention layers obtained high generalization in practice and modelling flexibility, and could cope well with the high variability and complexity under the meteorological and geographic conditions of a large area such as mainland China. Further, based on the imputed AOD, planetary boundary layer height (PBLH) and relative humidity, the grid maps of national ground aerosol coefficients were generated and shared.

The rest of this paper is organized as follows: Section 2 describes the study area, data and methods, including preprocessing, adaptive imputation modelling, evaluation and conversion to ground aerosol coefficient; Section 3 presents the results, the discussion of the results and data availability; and Section 4 draws the conclusions.

2. Materials and Methods

2.1. Study Area

The study area is mainland China, approximately between 18 and 54° north latitude and 73 and 135° east longitude. It covers an area of approximately 9.6 million square kilometers and had a population of approximately 1.4 billion in 2019 [41]. From the deserts in the arid north to the humid subtropical forests in the south, China has a vast and diverse landscape. The climate of the study area is mainly dominated by dry seasons and humid monsoons. Due to the highly complex topography, China’s climate varies from region to region. Mainland China has multiple aerosol sources including coal combustion, motor
vehicle emissions, and industrial sources, etc. In this study area, sulfate, ammonium, and nitrate from coal burning and vehicle emissions are the main causes of regional haze \[42,43\]; the continuous expansion of the deserts (e.g., Gobi Desert) has led to sandstorms, which is one of the main sources of aerosols \[44\]. Supplementary Figure S1 shows the study area of mainland China including its altitude and seven geographic zones, as well as six AERONET monitoring sites. The geographic zones are used to represent the differences in climate and land-use among different regions and are encoded using one-hot binary variables in the models to account for regional differences in mainland China. For imputation modeling, however, we trained a national model for each day, not separate regional models.

2.2. Original Satellite-Derived AOD Data and AERONET Data

The advanced MAIAC algorithm retrieves the daily AOD at a $1 \times 1$ km$^2$ spatial resolution from the MODIS Aqua (crossover approximately at 1:30 p.m. local time) and Terra (crossover approximately at 10:30 a.m. local time) sensors \[17,25\]. The DT algorithm associates the reflectance in visual bands with the 2.1 μm region to retrieve the surface reflectance and the DB algorithm uses an angular distribution model of surface reflectance derived from the time series of cloudless observations. As an improvement over the DT and DB algorithms, MAIAC uses time series and image processing to perform simultaneous retrievals of aerosol properties and surface bidirectional reflectance. Thus, it works on both dark vegetations and bright deserts at a higher spatial resolution \[45\].

We obtained 18 MODIS Sinusoidal tiles \[46\] of MAIAC AOD from the NASA data website (https://lpdaac.usgs.gov/products/mcd19a2v006/ (accessed on 6 January 2019)). These tiles together cover the entire mainland China with the sinusoidal projection (https://en.wikipedia.org/wiki/Sinusoidal_projection (accessed on 6 January 2019)). We mosaicked and cropped these tiles as a large grid to cover mainland China, and reprojected it onto the Beijing 1954 coordinate system (https://epsg.io/2412 (accessed on 6 January 2019)). The original MAIAC AOD data also provides the quality assurance flags that indicate cloud, land, water or snow pollution, which can be used to filter out invalid AOD.

In order to evaluate our imputation, the ground measurements of the AERONET data AOD from 6 AERONET monitoring sites (Supplementary Figure S1 for their spatial distribution) were gathered, and these data samples were measured when the sky was cloudless. We first obtained the averages of the 5-min AERONET data in the 60 min intervals to match the overpass times of the Aqua and Terra sensors; then the spectral linear interpolation was used to estimate AOD at 550 nm in the log-log space based on the AOD at the two nearest wavelengths \[47\]. The interpolated AERONET AOD at 550 nm was averaged and used as the “ground truth” to evaluate our imputed AOD.

2.3. Covariates

2.3.1. Coarse-Resolution Reanalysis Data

The Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) is a global reanalysis data product, which enables assimilation of modern hyperspectral radiance and microwave observations, climate parameters, the GEOS model and the GSI along with GPS-Radio Occultation datasets. Spatial resolution of MERRA2 is about 0.625° × 0.5°. MERRA2 AOD is estimated by assimilation of GEOS-5, the Goddard Chemistry, Aerosol, Radiation and Transport (GOCART) aerosol module, MODIS AOD data and AERONET ground-based observations. MERRA2 coarse-level ADO can provide background information of AOD in imputation modelling. Planetary boundary layer (PBL) is a critical factor in the turbulent mixing and vertical distribution of air pollutants, and its height determines the volume available for pollution dispersion and transport, and significantly affects vertical structure and turbulent mixing aerosols. We obtained reanalysis PBL height from the Goddard Earth Sciences Data and Information Services Centre “GIOVANNI” platform (https://giovanni.gsfc.nasa.gov/giovanni (accessed on 26 January 2019)). We also extracted the daily cloud fraction from the YD08_D3 MODIS Aqua and Terra Daily L3 Global 1Deg CMA datasets \[48,49\].
2.3.2. High-Resolution Meteorological Data by Spatiotemporal Interpolation

Since there are no publicly available meteorological surface grids to match our target resolution (1 × 1 km²), we used the full residual deep network [35] as the interpolation method to generate the daily 1 × 1 km² grids of meteorological variables. Compared with tree-based or other existing machine learning methods, the full residual deep network shows good performance and generalization in spatiotemporal interpolation of continuous surfaces [35,50]. Due to the large difference in spatial resolution (25 × 25 km² to 62.5 × 50 km² vs. 1 × 1 km²), the cubic or bilinear interpolation, if used directly to convert the meteorological reanalysis data to target resolution, may cause the resulting grid to lose the locality of spatial variations [50]. For our interpolation of meteorology, the input covariates included latitude, longitude, day of year, elevation and meteorological reanalysis data (Supplementary Table S1). We collected the hourly meteorological measurement data from approximately 900 monitoring stations of the China Meteorological Data Service (http://data.cma.cn (accessed on 6 January 2019)). These data were converted into daily data and used as the dependent variable to train the interpolation models. Finally, we generated the high-resolution grids of the following meteorological variables, including daily air pressure (hPa), air temperature (°C), relative humidity (%) and wind speed (m/s). For grid interpolation of these variables, the reanalysis data provided background covariates including air pressure, air temperature, wind speed and ozone. For details, please refer to Supplementary Section S1 and our published papers [51,52].

2.3.3. Geographic Zone

The geographic zone data can be used to distinguish the differences in climate, elevation and land-use among different locations. We obtained the geographic zone data of China from the Resource Environmental Science and Data Centre, Chinese Academy of Sciences (http://www.resdc.cn (accessed on 1 June 2019)). In total, we obtained seven zones (Supplementary Figure S1): Northeast China, Northwest China, North China, Southwest China, East China, Central China and South China. For training a daily-level national model, the one-hot coding method [53] was used to encode the geographic zoning factor through seven binary (0 or 1) variables to include it in the model to account for the zonal variability in AOD.

2.3.4. Other Spatial and/or Temporal Variables

The other spatial and temporal variables included either the coordinates (x and y) and their derivatives (squares of x and y, and product of x and y), elevation, time index, or combination. The elevation, and the coordinates and their derivatives were used to capture spatial variability of AOD.

2.4. Methods

2.4.1. The Preprocessing of Satellite-Derived AOD and Covariates

The preprocessing (Supplementary Figure S2) included data cleaning, projection conversion, filtering, temporal alignment, AOD fusing and merging and resampling. Data cleaning was to remove the invalid values in the original AOD and covariate data according to the normal value range. The outer fences were defined as (mean − 5 × IQR, mean + 5 × IQR) (IQR: interquartile range) to overlay with the normal range to filter out the invalid values. With the Beijing 1954 coordinate system as the target projection, projection conversion was performed for the AOD and covariate data. For the MAIAC AOD data, the quality assurance flags in the original dataset were used to filter out invalid AOD measurements due to the influences of cloud and high reflectance. Temporal alignment was performed, including averaging the fine timescale data (e.g., hourly meteorological measurements), on the target timescale (daily), and bilinear interpolation of coarse or misplaced timescale data to the target timescale. For fusing and merging of AOD data, see Supplementary Section S2. In addition, the meteorological reanalysis AOD and PBLH play an important role in the formation and spread of aerosols, but they have a coarse resolution
(approximately 25 × 25 km$^2$ or 62.5 × 50 km$^2$). We suggest bilinear interpolation \cite{54} as the resampling method to convert the coarse-resolution reanalysis data to the fine-resolution data.

2.4.2. Adaptive Imputation Modelling

Compared with other methods, the encoder–decoder full residual network has better generalization ability and better spatial variation simulation \cite{35}, so for imputation of missing AOD, it is used as the base model with the addition of attention layers (Supplementary Figure S3). The model was constructed based on the encoder–decoder architecture, which has a symmetrical structure from the encoding layers to the decoding layers. For the encoding layers, the number of nodes of a hidden layer has a downward trend; the middle hidden layer has the least number of nodes, as the coding layer of latent representation; for the decoding layers, the number of nodes of the hidden layers has an upward trend, corresponding to each encoding layer. The architecture can efficiently learn latent representation from raw features. Full residual connections are introduced to the encoder–decoder architecture and they can considerably improve learning efficiency and generalization ability.

Based on the full residual architecture, attention layers were added to enhance the contribution of important predictive covariates and reduce overfitting in the adaptive base model \cite{55} (Supplementary Figure S3).

In the architecture, each attention layer was a dense layer whose input was from the input layer or the output of the last attention layer.

Each attention layer generated a weight vector, where large weights emphasized the impacts of important covariates, and small weights reduced the impacts of secondary covariates:

$$h^{(l)} = a^{(l-1)}_h h^{(l-1)}_i + h^{(l-1)}_i (i = 1, 2, \ldots, m),$$

where $a^{(l-1)}_h$ is the weight of the $i$th node of the $l$-1 attention layer, $h^{(l-1)}_i$ is the $i$th input node of the $l$-1 attention layer and $i$ is the node index; and the residual connection was used to alleviate vanishing gradient in the attention layer. The softmax activation function was used to generate the weights of each attention layer (Supplementary Section S3).

For AOD imputation of a large area, the attention layers were used to highlight critical contributions of the important predictive covariates, thus reducing over-fitting in the full residual deep network.

For daily imputation of missing AOD in mainland China, the following loss function was defined:

$$L(\theta_{\mathbf{W}, \mathbf{b}}) = \frac{1}{N} \left[ \ell_y \left( y, \hat{y}_{\theta_{\mathbf{W}, \mathbf{b}}} (x) \right) + \ell_x \left( x, f_{\theta_{\mathbf{W}, \mathbf{b}}} (x) \right) \right] + \Omega(\theta_{\mathbf{W}, \mathbf{b}}),$$

where $x$ is the matrix of covariates used for imputation, $\theta_{\mathbf{W}, \mathbf{b}}$ denotes the set of parameters of weights ($\mathbf{W}$) and bias ($\mathbf{b}$), $y$ is the ground truth AOD, $\hat{y}_{\theta_{\mathbf{W}, \mathbf{b}}} (x)$ is the estimates of AOD, and $f_{\theta_{\mathbf{W}, \mathbf{b}}} (x)$ is the estimates of the input, $x$ by the network. $\ell_x$ is the mean square error (MSE) loss function for the input $x$, as a regularizer for the imputed $y$. $\Omega(\theta_{\mathbf{W}, \mathbf{b}})$ is the elastic regularizer ($w_i = 0.5, i = 1, 2$) \cite{56}.

In order to improve the learning efficiency, we developed an adaptive imputation method that could heuristically search optimal hyperparameters including a learning rate, a minibatch size, the number of attention layers, the number of hidden layers and the number of nodes in each hidden layer. Figure 1 shows the heuristic search process, which is described in detail in Supplementary Section S4.
we used Python to implement the adaptive imputation modelling to improve the efficiency of model learning.

2.4.3. Training and Evaluation

Among the daily samples with AOD available from the MAIAC data without cloud coverage, 20% were randomly selected for validation and 20% were randomly used for the independent test, and the remaining 60% were used to train the daily models. R² and root mean square error (RMSE) were used as the performance metrics. In the evaluation, training and testing metrics were summarized and compared. Furthermore, the spectrally interpolated AOD from AERONET was used as the third-party ground-truth to evaluate our imputation results, and R² and RMSE are also reported in the validation.

If the missing proportion of one day was too high (>75%), the available samples of the previous day and the next day were combined with the data of the day for training and testing. Correspondingly, the day index (−1: the previous day; 0: the day; 1: the next day) was included as a time covariate in the predictors to account for the day variation. If the size of the combined samples was still small, we used a similar method to include the data of more days (two days each time). For training and testing, similar proportions were used to split the merged samples (60%: training; 20%: validation; 20%: testing).

2.4.4. Conversion of AOD into Ground Aerosol Coefficient

With the complete AOD (original values plus imputed values), downscaled PBLH and relative humidity (RH), the following optimal formula [29] was used to convert the complete AOD into the ground aerosol coefficient:

\[
GAC = \frac{AOD \times (1 - RH/100)^g}{PBLH},
\]  

where \(AOD\) is the original and imputed AOD, \(GAC\) denotes the ground aerosol coefficient and \(g\) is the parameter to be solved. Automatic differentiation [57] can be used to find an optimal \(g\). For different geographic regions, we also fitted different \(g\) to show the influence of different zones on the GAC conversion in mainland China.
3. Results and Discussion

3.1. Descriptive Statistics of AOD and Fusion of High-Resolution Meteorology

For mainland China, the mean AOD of 2015–2018 was 0.22 with a standard deviation of 0.16. Figure 2 shows the spatial distribution of mean AOD (a) and seasonal means (b for spring, c for summer, d for autumn and d for winter) of AOD for 2015–2018. The average AOD in summer and spring was higher than that in autumn and winter (0.25–0.26 vs. 0.19–0.22; Table 1). In addition, the average AOD in central and eastern China is higher than in other regions (0.45–0.47 vs. 0.13–0.39); the average AOD in southern China is higher than that in northern China (0.36 vs. 0.18); the northeast, northwest and southwest regions have lower average AOD values than other regions (0.13–0.19 vs. 0.30–0.47; Figure 2 and Table 2). As economic development drove high emission sources of aerosols, the AOD values of the central and eastern developed regions were higher than other regions.

Overall, the average proportion of missing values in the study area from 2015 to 2018 was 63%, and the standard deviation was 18% (Table 1). Figure 3 shows the spatial distribution of the mean (a) and seasonal means (b for spring, c for summer, d for autumn and d for winter) of the missing proportion of AOD in mainland China. In summer and autumn, the missing proportion of AOD is higher than that in spring and winter (67% vs. 53–64%). For mainland China, due to the influence of precipitation and cloud cover [27], the average missing proportion of AOD in the southern region is higher than that in the northern region (81% vs. 57%); the average missing proportions of AOD in eastern, central, southern and southwest China are higher than that in the other regions (70–87% vs. 50–66%; Table 2).

Table 1. Mean and seasonal means of aerosol optical depth (AOD) and its missing proportion for 2015–2018 in mainland China.

|                      | 2015–2018 | Spring | Summer | Autumn | Winter |
|----------------------|-----------|--------|--------|--------|--------|
| Mean AOD             | 0.22      | 0.25   | 0.26   | 0.22   | 0.19   |
| Standard deviation of AOD | 0.16    | 0.19   | 0.18   | 0.16   | 0.15   |
| Mean missing proportion | 63%     | 64%    | 67%    | 67%    | 53%    |
| Standard deviation of missing proportion | 18% | 23%    | 18%    | 21%    | 22%    |

Table 2. Mean and regional means of AOD and its missing proportion for 2015–2018 in mainland China.

|                    | Northeast China | North China | East China | Central China | South China | Northwest China | Southwest China |
|--------------------|-----------------|-------------|------------|---------------|-------------|-----------------|-----------------|
| Mean AOD           | 0.19            | 0.30        | 0.45       | 0.47          | 0.39        | 0.18            | 0.13            |
| Standard deviation of AOD | 0.06 | 0.14       | 0.12       | 0.11          | 0.06        | 0.11            | 0.16            |
| Mean missing proportion | 66% | 50%        | 74%        | 76%           | 87%         | 52%             | 70%             |
| Standard deviation of missing proportion | 13% | 7%         | 11%        | 9%            | 4%          | 17%             | 17%             |
Figure 2. Overall mean distribution of 2015–2018 AOD (a) and seasonal mean distributions (b) for spring, (c) for summer, (d) for autumn and (e) for winter of AOD from 2015 to 2018.
Figure 3. Spatial distributions of the proportions of missing AOD for 2015–2018 (a), the springs of 2015–2018 (b), the summers of 2015–2018 (c), the autumns of 2015–2018 (d) and the winters of 2015–2018 (e).
For the reanalysis of PBLH and AOD, compared with the original coarse-resolution data, the high-resolution (1 × 1 km²) grid resampled using bilinear interpolation showed a spatially smoothing surface (Supplementary Figure S4 for the PBLH example). Sensitivity analysis showed that bilinear interpolation and the downscaling method [52,58] had similar results for the imputation of missing AOD.

Table 3 shows the test performance of the fusion of high-resolution meteorological variables for mainland China. The results show high test performance and generalization for the trained models of the meteorological factors (air pressure, air temperature, relative humidity and wind speed). Due to the volatility of wind, estimating wind speed is challenging. Thus, we performed bootstrap aggregating (bagging) of the full residual deep base networks for the estimation of wind speed and obtained an R² of 0.79.

| Type            | Model                          | Test R² | Test RMSE | Modelling Way |
|-----------------|--------------------------------|---------|-----------|---------------|
| Air pressure    | Generalized additive model     | 0.99    | 5.14 hPa  | One model     |
| Air temperature | Full residual deep network     | 0.92    | 3.42 °C   | One model     |
| Relative humidity | Full residual deep network     | 0.88    | 0.11%     | One model     |
| Wind speed      | Full residual deep network     | 0.79    | 0.56 m/s  | Bagging of 100 base models |

3.2. Reliability of Imputed AOD and Simulated GAC

For imputation of AOD, we trained 1461 daily-level models with attention layers from 1 January 2015 to 31 December 2018. For each daily-level model, the learning procedure was stable, and the model parameters also converged to an optimal solution (Supplementary Figure S5 for the learning curves of training and testing R² for four seasonal days). Figure 4 shows the time series of test R² and test RMSE for the 1461 models, and Table 4 shows the summary of test performance metrics (Supplementary Figure S6 for the scatter plots between observed AOD and predicted AOD in independent tests on four seasonal days). In total, the R² had a mean of 0.90, ranging from 0.75 to 0.97, with a standard deviation of 0.04 and an IQR of 0.05; the RMSE had a mean of 0.08, ranging from 0.03 to 0.32, with a standard deviation of 0.04 and an IQR of 0.05. The model’s daily-level test performance seemed to remain stable with small changes, as shown in Figure 4 (Supplementary Figure S7 for the histograms of test R² and RMSE). The analysis showed that, without adding data from other days, a higher AOD missing proportion was one of the main reasons for the lower test R². Due to the high missing proportion, the sample size in training and testing was small, which might have a negative impact on the testing performance of the trained model [59]. In addition, the high cloud proportion was the main reason for the high missing proportion; accordingly, this was also one of the main reasons for the low test R².

Figure 4. Time series of test R² (a) and test RMSE (b) for the trained daily-level imputation models.
Table 4. Summary of the test performances for daily imputation models.

| Type   | Metrics | Mean | Median | Minimum | Maximum | Standard Deviation | IQR  |
|--------|---------|------|--------|---------|---------|--------------------|------|
| Training | $R^2$   | 0.90 | 0.91   | 0.75    | 0.97    | 0.04               | 0.05 |
|        | RMSE    | 0.08 | 0.07   | 0.03    | 0.33    | 0.03               | 0.03 |
| Testing | $R^2$   | 0.90 | 0.91   | 0.75    | 0.97    | 0.04               | 0.05 |
|        | RMSE    | 0.08 | 0.07   | 0.03    | 0.32    | 0.03               | 0.03 |

As shown in the summary and the boxplots (Figure 5), there are almost no differences between training performance metrics and testing performance metrics, indicating little over-fitting and reliability for our imputation models. Due to the low initial test performance (<0.75), our adaptive method automatically retrained the imputation models of 30 days after searching for their optimal hyperparameters. The default configuration of hyperparameters (minibatch size: 1024; number of attention layers: 2; learning rate: 0.1; number of layers: 12; numbers of nodes: (15, 15, 128, 64, 32, 16, 8, 16, 32, 64, 128, 15 and 16)) worked fine for most (approximately 98%) daily imputation models.

![Figure 5](image.png)

Figure 5. Boxplots of training and testing $R^2$ (a) and RMSE (b) for the imputation of Multi-Angle Implementation of Atmospheric Correction AOD.

Sensitivity analysis was conducted to evaluate the contributions of geographic zoning and attention layers. Compared with the full residual deep network, the analysis shows that the climate zone and attention layers made average improvements in test $R^2$ of 2% and 3.5%, respectively. Both contributed an average improvement of 5% in test $R^2$. The analysis shows that for such a large area as mainland China, it is important to include geographic zones to account for variability in meteorology and land-use, and adding the attention levels to highlight important predictors in imputing missing AOD. For the large study area with high complexity in regional heterogeneity, the attention layers were learned to improve the model sensitivity to the influence of the covariates on the predictions through different attention weights in different regions. Therefore, by focusing more on the important parts rather than the minor or noisy parts in the input covariates in different regions, the testing performance was improved. The importance scores adaptively varied with the input for the model to obtain the high test performance. Statistics on the results showed that the covariates with a higher importance score included MERRA2 AOD and PBLH, relative humidity, air pressure, altitude, latitude and longitude.

In total, 3608 samples collected from six AERONET sites were used to validate the imputed values. In short, we obtained an $R^2$ of 0.78 (correlation: 0.88) and an RMSE of 0.27 in the validation using AERONET data (Figure 6a). The residuals from the AERONET AOD presented a normal distribution with a mean (0.02) close to 0 (Figure 6b), indicating that the majority of the spatiotemporal variance was reliably captured by our imputations. One thousand, three hundred and thirty-eight samples of the AERONET AOD were collected under a cloudy sky where MAIAC AOD was missing. The imputed AOD under the cloudy sky had an $R^2$ of 0.68 (correlation: 0.83; RMSE: 0.31). In a cloudy sky, aerosols
may be activated and turned into cloud droplets, which may reduce AOD and increase cloud optical depth [60,61]. It is possible for the imputed AOD to be overestimated under a cloudy sky. In the proposed method, the cloud fraction was used to account for this difference in AOD between the clear sky and cloudy sky. As shown in the result, the proposed method achieved a high correlation (0.83) of imputed AOD with AERONET AOD, indicating its reliability under the cloudy sky when extensive AOD was missing.

![Figure 6. Scatter plot of training and testing R² (a) and a histogram of the residuals (b) in the imputation of MAIAC AOD of 2015–2018.](image)

Using the complete AOD, and downscaled PBLH and RH, the grid map dataset of ground aerosol coefficient was simulated. The national optimal $g$ in Equation (3) is 0.21. The sensitivity analysis showed that the optimal $g$ varied for each region. However, for the estimation of ground PM$_{2.5}$, the difference between the national optimal $g$ and the regional optimal $g$ was very small, as shown by their small difference in the RMSE of Supplementary Table S2.

AOD is defined as the integral of aerosol extinction coefficient at all altitudes and presents a probably exponential distribution along the vertical profile. The temporally varying mixing vertical height has a great influence on the relationship between AOD and ground aerosol concentrations [62]. Typically, due to the lower PBL height (Supplementary Figure S8) in winter, column-integrated AOD is higher in summer than in winter, which is quite different from the seasonal trend of ground aerosols such as PM$_{10}$ and PM$_{2.5}$ (high in winter and low in summer) [62]. Thus, PBLH and RH played a critical role in the conversion from AOD to GAC. As shown in Figures 9 and 12, the early spring day seemed to have higher AOD and the winter seemed to have higher GAC. The increase in AOD may cause the ground surface to cool, which may impede the development of PBL so that more aerosols such as PM$_{2.5}$ may be concentrated on the ground [63]. The complex spatiotemporal interactions between meteorological factors (including relative humidity, wind speed and PBLH, etc.) and aerosols may lead to uncertainty in the estimation of these variables [64,65]. Studies have shown that satellite-derived [66] and MERRA2 [67,68] PBLH have high uncertainty. Since the satellite did not capture the shallow surface aerosol layers in winter, the uncertainty of PBLH estimates might have a negative impact on ground aerosol estimates.

Despite the uncertainty in the MERRA2 PBLH and RH estimates, GAC converted from AOD can better represent the temporal and seasonal trend of ground aerosol concentrations (Figures 7 and 8) and can be used as a better predictive covariate for estimating PM$_{2.5}$. The results show that the correlation (0.51) between GAC and PM$_{2.5}$ was improved by 0.21, compared with that (0.30) of the AOD. Therefore, compared with AOD, GAC is a
better proxy variable used to represent temporal trend and spatial distribution of ground aerosol pollutants.

Figure 7. Time series of the mean (a) and standard deviation (b) of AOD in mainland China for 2015–2018.

Figure 8. Time series of the mean (a) and standard deviation (b) of ground aerosol coefficient (GAC) in mainland China for 2015–2018.

In order to evaluate the reliability of the imputation for mainland China in practice, we compared the original and complete AODs, and the MAIAC land surface reflectance of true color (bands 1, 4 and 3) and GAC. Figures 9–12 show the original (a) and complete (b) AOD, the land surface reflectance (c) and the ground aerosol coefficient (d) for four typical seasonal days (Figure 9: 1 April 2015; Figure 10: 1 July 2016; Figure 11: 1 October 2017; Figure 12: 30 December 2018). The comparison shows the natural and reasonable spatial smoothness of the transitions between the observed (original) AOD and imputed AOD on each of the four days. The land surface reflectance of true color shows that highly reflective surface coverage was a dominant factor for the extensive missingness of MAIAC AOD.
Figure 9. Imputation of MAIAC AOD for a spring day (1 April) of 2015: original AOD (a), complete AOD (b), MAIAC ground reflectance of true color (c) and GAC (d).
Figure 10. Imputation of MAIAC AOD for a summer day (1 July) of 2016: original AOD (a), complete AOD (b), MAIAC ground reflectance of true color (c) and GAC (d).
Figure 11. Imputation of MAIAC AOD for an autumn day (1 October) of 2017: original AOD (a), complete AOD (b), MAIAC ground reflectance of true color (c) and GAC (d).
Figure 12. Imputation of MAIAC AOD for a winter day (30 December) of 2018: original AOD (a), complete AOD (b), MAIAC ground reflectance of true color (c) and GAC (d).

3.3. Spatiotemporal Distributions of AOD and GAC for Mainland China

As shown in Figure 7a, the national average AOD (including imputed AOD) presented a seasonal pattern (high in spring and low in winter). Similarly, due to the influence of average AOD, the standard deviation of AOD had a similar pattern (Figure 7b). In most areas of mainland China, the low-altitude AOD in winter is the highest because more aerosols are concentrated near the ground surface, thereby resulting in high ground aerosol loading, due to a lower atmospheric boundary layer in winter than in summer [69]. In addition, the northwest desert dust, as one of the main sources of aerosols, covers most of Xinjiang and had an important impact on the high AOD in this area; anthropogenic aerosols were one of the main reasons for the high AOD in North China and Central China.
Figure 8 showed the seasonal time series of GAC mean (a) and its standard deviation (b) from 2015 to 2018 in mainland China. In addition, from 2015 to 2018, the national annual average AOD (Figure 13a) and GAC (Figure 13b) showed a slight downward trend. The cities also showed similar GAC seasonal patterns, but with different magnitudes (Supplementary Figures S9–S12 for four megacities, i.e., Beijing in the metropolitan Jing-Jin-Tang region, Shanghai in the Yangtze River Delta, Guangzhou in the Pearl River Delta and Chengdu in the Chengdu Plain). The spatial distributions of daily GAC in four seasonal days are shown for the four megacities (Figures 14 and 15 and Supplementary Figures S13 and S14). Although Beijing, Shanghai and Guangzhou had similar seasonal patterns (high in spring or winter and low in summer), Chengdu had a different pattern (generally high in four seasons). Compared with the other three cities, Chengdu was more polluted by PM$_{2.5}$ [70]. In addition, with the exception of Chengdu and Guangzhou, the average GAC of Beijing and Shanghai also showed a moderate downward trend from 2015 to 2018 (Figure 13).

![Figure 13](image_url)

**Figure 13.** Average annual AOD (a) and ground aerosol coefficient (b) in mainland China and four megacities from 2015 to 2018.
Figure 14. Spatial distributions for four seasonal days in Beijing: a spring day (1 April 2015) (a), a summer day (1 July 2016) (b), an autumn day (1 October 2017) (c) and a winter day (30 December 2018) (d).

As an indicator of aerosol pollution (e.g., PM\(_{2.5}\) and PM\(_{10}\), and total suspended particles), the declining trend of satellite GAC might show the effectiveness of the Chinese government in controlling air pollution. For example, the China Clean Air Policy has been implemented since 2013, resulting in a gradual decline in PM\(_{2.5}\) [71,72]. On the other hand, the seasonal pattern (Figure 8) of average GAC of 2015–2018 in mainland China was consistent with the trends of PM\(_{10}\) and PM\(_{2.5}\), which are usually low in summer and high in winter.
Due to the influence of the spatiotemporally varying PBLH, AOD and GAC presented different patterns for mainland China. A typical example is the spatial distributions of AOD and GAC on 30 December of 2018 (b vs. d of Figure 12). The AOD map showed low and moderate values for most of mainland China, but the GAC map showed moderate to high pollution in parts of southern, central, eastern, northwest and northeastern China, which was consistent with actual situations [73]. Table 5 lists the simulated GAC summary of mainland China and its sub-regions. It shows that the central region had the highest average GAC from 2015 to 2018 (0.00113), indicating heavy air pollution for this region. On the contrary, the average GAC in the southwest region was the lowest (0.00032), indicating less pollution for this region. In mainland China and its sub-regions, summer air pollution is the lowest, while winter air pollution is the highest.

The complete high-resolution (1 × 1 km²) AOD and GAC across mainland China during 2015–2018 are available at https://doi.org/10.7910/DVN/RNSWRH (accessed on 30 December 2020) [74] and https://doi.org/10.7910/DVN/YDJT3L (accessed on 20 February 2021) [75], respectively, which can be downloaded in GeoTiff format. We will continue to update the datasets to cover 2019, 2020 and earlier years.

Figure 15. Spatial distributions of GAC (m⁻¹) for four seasonal days in Shanghai: a spring day (1 April 2015) (a), a summer day (1 July 2016) (b), an autumn day (1 October 2017) (c) and a winter day (30 December 2018) (d).
Table 5. Summary of the simulated GAC (m$^{-1}$) for mainland China and its subregions.

| Region          | Mainland China | Northeast China | North China | East China | Central China | South China | Northwest China | Southwest China |
|-----------------|----------------|-----------------|------------|-----------|--------------|------------|----------------|-----------------|
| **Mean GAC**    | 0.00052        | 0.00052         | 0.00075    | 0.00097   | 0.00113      | 0.00072    | 0.00038         | 0.00032         |
| **Standard deviation of GAC** | 0.00021        | 0.00042         | 0.00061    | 0.0005    | 0.00064      | 0.00038    | 0.00021         | 0.00012         |
| **Spring**      | 0.00051        | 0.00045         | 0.00059    | 0.0010    | 0.0011       | 0.00091    | 0.00038         | 0.00033         |
| **Summer**      | 0.00037        | 0.00039         | 0.00065    | 0.00076   | 0.00077      | 0.00053    | 0.00023         | 0.00027         |
| **Autumn**      | 0.00046        | 0.00045         | 0.00075    | 0.00082   | 0.0010       | 0.00062    | 0.00034         | 0.00030         |
| **Winter**      | 0.00071        | 0.00080         | 0.0010     | 0.0013    | 0.0016       | 0.00080    | 0.00058         | 0.00036         |

This study has several limitations. First, although we reported the average high test R$^2$ in the sample test (0.90) and AERONET data (0.78), the data from the six monitoring AERONET sites was limited for national validation of imputed MAIAC AOD. In the future when more AERONET sites become available, more test data can be used in imputation validation. Second, when satellite-derived AOD cannot capture shallow surface aerosol pollutants as in winter, the uncertainty of PBLH estimates will have a negative impact on GAC estimates, inevitably having ill-posed features. In the future, PBLH measured using ground-based lidar, if available, may be used to adjust satellite-derived PBLH to reduce bias in the conversion from AOD to GAC. Third, we used national instead of regional g in the conversion from AOD to GAC, although the difference in the estimation of ground aerosols such as PM$_{2.5}$ by different GAC is very small.

4. Conclusions

This article reported an adaptive method of a full residual deep network with attention layers to improve the imputation of extensive missing AOD and the mapping of ground aerosol coefficients for mainland China. To account for the regional variability of emissions and meteorology in a large study area, the geographical zoning factor was encoded as one-hot quantitative variables in the daily imputation model, and attention layers were added to highlight the contribution of important predictive covariates to improving generalization. In the adaptive method, we designed a heuristic search method to improve the optimization of hyperparameters, thereby improving the efficiency of imputation training on large satellite datasets. The daily imputation models achieved an average test R$^2$ of 0.90, ranging from 0.75 to 0.97 (average RMSE: 0.075), with an improvement of approximately 5% over the original full residual network. There were smooth transitions (no unrealistic abrupt changes) between the imputed values and the original AOD. The validation with the AERONET AOD (R$^2$: 0.78) further illustrated the reliability of the proposed method. We shared the datasets of high-resolution complete AOD and ground aerosol coefficient for mainland China from 2015 to 2018. This study has important implications for the imputation of AOD and the mapping of ground aerosol coefficients for a large area, and extensive applications of high-resolution MAIAC AOD data.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/rs13122324/s1, Section S1: Spatiotemporal interpolation of meteorological parameters, Section S2: Fusing and merging of AOD Data, Section S3: Softmax function, Section S4: Procedure of heuristic search for adaptive imputing, Table S1: Input covariates for the high-resolution interpolation of meteorology, Table S2: Sensitivity analysis of regionally-varying g in the conversion from AOD to GAC; Figure S1: Study region of mainland China, with AERONET monitoring stations, elevation and climate zones, Figure S2: Procedure of data preprocessing for imputation, Figure S3: Full residual deep network with attention layers, Figure S4: Original PBLH (a) and interpolated 1 × 1 km$^2$ PBLH by bilinear method (b), Figure S5: Learning curves of training and testing R$^2$ of the daily-level models for four seasonal days (1 April 2015 for a; 1 July 2016 for b; 1 October 2017 for c; 30 December 2018 for d), Figure S6: Scatter plots of the predicted and observed AOD for four seasonal days (1 April 2015 for a; 1 July 2016 for b; 1 October 2017 for c; 30 December 2018 for d), Figure S7: Histograms of testing R$^2$ (a) and RMSE (b) for MAIAC AOD of mainland China, Figure S8: Seasonal variation
of daily PBLH for MAIAC AOD of mainland China, Figure S9: Time series of AOD mean (a) and standard deviation (b) in Beijing, Figure S10: Time series of AOD mean (a) and standard deviation (b) in Shanghai, Figure S11: Time series of AOD mean (a) and standard deviation (b) in Guangzhou, Figure S12: Time series of AOD mean (a) and standard deviation (b) in Chengdu, Figure S13: Spatial distributions of GAC for four seasonal days in Chengdu, Figure S14: Spatial distributions of GAC for four seasonal days in Guangzhou.

**Author Contributions:** L.L. designed the research and developed the methodology; L.L. wrote the manuscript and revised the manuscript.

**Funding:** This work was supported in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDA19040501 and in part by the National Natural Science Foundation of China under Grants 41471736 and 42071369.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The daily complete high-resolution (1 × 1 km²) AOD and GAC across mainland China during 2015–2018 are available at https://doi.org/10.7910/DVN/RNSWRH and https://doi.org/10.7910/DVN/YDJT3L, respectively, which can be downloaded in GeoTiff format.

**Acknowledgments:** The support of Nvidia Corporation through the donation of the Titan Xp GPUs used in this research.

**Conflicts of Interest:** The author declares that he has no conflicts of interest.

**References**

1. NASA. Aerosol Optimal Depth. 2017. Available online: https://aeronet.gsfc.nasa.gov/new_web/Documents/Aerosol_Optical_Depth.pdf (accessed on 21 May 2019).

2. Levy, H.; Horowitz, L.W.; Schwarzkopf, M.D.; Ming, Y.; Golaz, J.C.; Naik, V.; Ramaswamy, V. The roles of aerosol direct and indirect effects in past and future climate change. *J. Geophys. Res. Atmos.* 2013, 118, 4521–4532. [CrossRef]

3. Ocko, I.B.; Ramaswamy, V.; Ginoux, P.; Ming, Y.; Horowitz, L.W. Sensitivity of scattering and absorbing aerosol direct radiative forcing to physical climate factors. *J. Geophys. Res. Atmos.* 2012, 117. [CrossRef]

4. Xing, Y.F.; Xu, Y.H.; Shi, M.H.; Lian, Y.X. The impact of PM2.5 on the human respiratory system. *J. Thorac. Dis.* 2016, 8, E69–E74. [CrossRef]

5. Kelly, F.J.; Fussell, J.C. Air pollution and public health: Emerging risks and improved understanding of risk. *Environ. Geochem. Health* 2015, 37, 631–649. [CrossRef] [PubMed]

6. Du, Y.X.; Xu, X.H.; Chu, M.; Guo, Y.; Wang, J.H. Air particulate matter and cardiovascular disease: The epidemiological, biomedical, and clinical evidence. *J. Thorac. Dis.* 2016, 8, E8–E19. [CrossRef]

7. Zhang, B.; Liang, S.W.; Zhao, J.Z.; Qian, Z.M.; Bassig, B.A.; Yang, R.; Zhang, Y.M.; Hu, K.; Xu, S.Q.; Zheng, T.Z.; et al. Maternal exposure to air pollutant PM2.5 and PM10 during pregnancy and risk of congenital heart defects. *J. Expo. Sci. Environ. Epid.* 2016, 26, 422–427. [CrossRef] [PubMed]

8. Shiraiwa, M.; Ueda, K.; Pozzer, A.; Lammel, G.; Kampf, C.J.; Fushimi, A.; Enami, S.; Arangio, A.M.; Fröhlich-Nowoisky, J.; Fujitani, Y.; et al. Aerosol Health Effects from Molecular to Global Scales. *Environ. Sci. Technol.* 2017, 51, 13545–13567. [CrossRef]

9. Leung, N.H.L.; Chu, D.K.W.; Shi, E.Y.C.; Chan, K.H.; McDevitt, J.J.; Hau, B.J.P.; Yen, H.L.; Li, Y.G.; Ip, D.K.M.; Peiris, J.S.M.; et al. Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. *Nature Med.* 2020, 26, 557. [CrossRef]

10. Liu, Y.; Ning, Z.; Chen, Y.; Guo, M.; Liu, Y.L.; Gali, N.K.; Sun, L.; Duan, Y.S.; Cai, J.; Westerdahl, D.; et al. Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. *Nature Med.* 2020, 26, 981. [CrossRef] [PubMed]

11. Liu, Y.; Ning, Z.; Chen, Y.; Guo, M.; L: Liu, Y.L.; Gali, N.K.; Sun, L.; Duan, Y.S.; Cai, J.; Westerdahl, D.; et al. Aerosol Health Effects from Molecular to Global Scales. *Environ. Sci. Technol.* 2017, 51, 13545–13567. [CrossRef] [PubMed]

12. Fan, X.H.; Xia, X.G.; Chen, H.B. Intercomparison of Multiple Satellite Aerosol Products against AERONET over the North China Plain. *Atmosphere* 2019, 10, 480. [CrossRef]

13. Kaufman, Y.J.; Joseph, J.H. Determination of surface albedos and aerosol extinction characteristics from satellite imagery. *J. Geophys. Res.* 1982, 87, 1287–1299. [CrossRef]

14. Tanre, D.; Kaufman, Y.J.; Herman, M.; Mattoo, S. Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances. *J. Geophys. Res. Atmos.* 1997, 102, 16971–16988. [CrossRef]

15. Hsu, N.C.; Jeong, M.J.; Bettenhausen, C.; Sayer, A.M.; Hansell, R.; Setfort, C.S.; Huang, J.; Tsay, S.C. Enhanced Deep Blue aerosol retrieval algorithm: The second generation. *J. Geophys. Res. Atmos.* 2013, 118, 9296–9315. [CrossRef]

16. Hsu, N.C.; Tsay, S.C.; King, M.D.; Herman, J.R. Aerosol properties over bright-reflecting source regions. *IEEE Trans. Geosci. Remote* 2004, 42, 557–569. [CrossRef]
17. Lyapustin, A.; Wang, Y.; Laszlo, I.; Kahn, R.; Korkin, S.; Remer, L.; Levy, R.; Reid, J.S. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *J. Geophys. Res. Atmos.* 2011, 116. [CrossRef]

18. Lyapustin, A.; Wang, Y.J.; Korkin, S.; Huang, D. MODIS Collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* 2018, 11, 5741–5765. [CrossRef]

19. Bilal, M.; Nichol, J.E.; Bleiweiss, M.P.; Dubois, D. A Simplified high resolution MODIS Aerosol Retrieval Algorithm (SARA) for use over mixed surfaces. *Remote Sens. Environ.* 2015, 136, 135–145. [CrossRef]

20. Tanre, D.; Breon, F.M.; Deuze, J.L.; Dubovik, O.; Ducos, F.; Francois, P.; Goloub, P.; Herman, M.; Lifermann, A.; Waquet, F. Remote sensing of aerosols by using polarized, directional and spectral measurements within the A-Train: The PARASOL mission. *Atmos. Meas. Tech.* 2011, 4, 1383–1395. [CrossRef]

21. Lanzaco, B.L.; Olcese, L.E.; Palencar, G.G.; Toselli, B.M. An Improved Aerosol Optical Depth Map Based on Machine-Learning and MODIS AOD: Development and Application in South America. *Aerosol Air Qual. Res.* 2017, 17, 1623–1636. [CrossRef]

22. Wei, X.L.; Chang, N.B.; Bai, K.X.; Gao, W. Satellite remote sensing of aerosol optical depth: Advances, challenges, and perspectives. *Crit. Rev. Environ. Sci. Technol.* 2020, 50, 1640–1725. [CrossRef]

23. Singh, M.K.; Venkatachalam, P.; Gautam, R. Geostatistical Methods for Filling Gaps in Level-3 Monthly-Mean Aerosol Optical Depth Data from Multi-Angle Imaging SpectroRadiometer. *Aerosol Air Qual. Res.* 2017, 17, 1963–1974. [CrossRef]

24. Bai, Y.; Wu, L.; Qin, Z.; Zhang, Y.; Shen, Y.; Zhou, Y. A Geographically and Temporally Weighted Regression Model for Ground Level PM2.5 Estimation from Satellite-Derived 500 m Resolution AOD. *Remote Sens.* 2016, 8, 262. [CrossRef]

25. Lyapustin, A.; Wang, Y. MODIS Multi-Angle Implementation of Atmospheric Correction (MAIAC) Data User’s Guide. 2018. Available online: https://modis-land.gsfc.nasa.gov/pdf/MCD19_UserGuide_final_Feb-6-2018.pdf (accessed on 20 February 2019).

26. Zhang, J.L.; Reid, J.S. An analysis of clear sky and contextual biases using an operational over ocean MODIS aerosol product. *Geophys. Res. Lett.* 2009, 36. [CrossRef]

27. Xiao, Q.; Wang, Y.; Chang, H.H.; Meng, X.; Geng, G.; Lyapustin, A.; Liu, Y. Full-coverage high-resolution daily PM2.5 estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote Sens. Environ.* 2017, 199, 437–446. [CrossRef]

28. Li, L.F. Optimal Inversion of Conversion Parameters from Satellite AOD to Ground Aerosol Extinction Coefficient Using Automatic Differentiation. *Remote Sens.* 2020, 12, 492. [CrossRef]

29. Wang, Z.F.; Chen, L.F.; Tao, J.H.; Zhang, Y.; Su, L. Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method. *Remote Sens. Environ.* 2010, 114, 50–63. [CrossRef]

30. Kloog, I.; Coull, B.A.; Zanobetti, A.; Koutrakis, P.; Schwartz, J.D. Acute and chronic effects of particles on hospital admissions in New-England. *PLoS ONE* 2012, 7, e34664. [CrossRef] [PubMed]

31. Li, S.S.; Chen, L.F.; Tao, J.H.; Han, D.; Wang, Z.T.; Su, L.; Fan, M.; Yu, C. Retrieval of aerosol optical depth over bright targets in the urban areas of North China during winter. *Sci. China Earth Sci.* 2012, 55, 1545–1553. [CrossRef]

32. Van Donkelaar, A.; Martin, R.V.; Levy, R.C.; da Silva, A.M.; Krzyzanowski, M.; Chubarova, N.E.; Semutnikova, E.; Cohen, A.J. Satellite-based estimates of ground-level fine particulate matter during extreme events: A case study of the Moscow fires in 2010. *Atmos. Environ.* 2011, 45, 6225–6232. [CrossRef]

33. Lv, B.; Hu, Y.; Chang, H.H.; Russell, A.G.; Bai, Y. Improving the accuracy of daily PM2.5 distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in north China. *Environ. Sci. Technol.* 2016, 50, 4752–4759. [CrossRef]

34. Di, Q.; Kloog, I.; Koutrakis, P.; Lyapustin, A.; Wang, Y.; Schwartz, J. Assessing PM2.5 Exposures with High Spatiotemporal Resolution across the Continental United States. *Remote Sens. Environ.* 2016, 170, 4712–4721. [CrossRef]

35. Li, L.; Fang, Y.; Wu, J.; Wang, J.; Y. G. Encoder-Decoder Full Residual Deep Networks for Robust Regression Prediction and Spatiotemporal Estimation. *IEEE Trans. Neural Netw. Learn. Syst.* 2020, in press. [CrossRef]

36. Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E. Hierarchical Attention Networks for Document Classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, CA, USA, 12–17 June 2016; Association for Computational Linguistics: Stroudsburg, PA, USA, 2016; pp. 1480–1489.

37. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. *arXiv preprint 2017*, arXiv:1706.03762.

38. Zhou, P.; Shi, W.; Tian, J.; Qi, Z.; Li, B.; Hao, H.; Xu, B. Attention-based bidirectional long short-term memory networks for relation classification. In *Proceedings of the 54th annual meeting of the association for computational linguistics, Berlin, Germany, 7–12 August 2016*; Association for Computational Linguistics: Stroudsburg, PA, USA; pp. 207–212.

39. Li, Y.; Zeng, J.B.; Shan, S.G.; Chen, X.L. Occlusion Aware Facial Expression Recognition Using CNN with Attention Mechanism. *IEEE Trans. Image* 2019, 28, 2439–2450. [CrossRef]

40. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

41. National Bureau of Statistics of China. Total Population. Available online: https://en.wikipedia.org/wiki/National_Bureau_of_Statistics_of_China (accessed on 21 March 2020).

42. Bai, Z.P.; Han, J.B.; Azmi, M. Insights into measurements of ambient air PM2.5 in China. *Trends Environ. Anal.* 2017, 13, 1–9. [CrossRef]

43. Pui, D.Y.H.; Chen, S.C.; Zuo, Z.L. PM2.5 in China: Measurements, sources, visibility and health effects, and mitigation. *Particuology* 2014, 13, 1–26. [CrossRef]
44. Waghorn, T. Fighting Desertification. Available online: https://www.forbes.com/sites/terrywaghorn/2011/03/07/fighting-desertification/?sh=4dec8b795fd7 (accessed on 21 January 2020).

45. Liu, N.; Zou, B.; Feng, H.H.; Wang, W.; Tang, Y.Q.; Liang, Y. Evaluation and comparison of multivariate implementation of the atmospheric correction algorithm, Dark Target, and Deep Blue aerosol products over China. Atmos. Chem. Phys. 2019, 19, 8243–8268. [CrossRef]

46. NASA MODIS Grids. Available online: https://modis-land.gsfc.nasa.gov/MODLAND_grid.html (accessed on 1 June 2021).

47. Eck, T.F.; Holben, B.N.; Reid, J.S.; Dubovik, O.; Smirnov, A.; O’Neill, N.T.; Slutsky, I.; Kinne, S. Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. J. Geophys. Res. Atmos. 1999, 104, 31333–31349. [CrossRef]

48. NASA MODIS Atmosphere. Available online: https://modis-images.gsfc.nasa.gov/products.html (accessed on 3 September 2020).

49. Oreopoulos, L.; Cho, N.; Lee, D.; Kato, S.; Huffman, G. An examination of the nature of global MODIS cloud regimes. JGR Atmos. 2014, 119, 8362–8383. [CrossRef]

50. Li, L.F.; Franklin, M.; Girguis, M.; Lurmann, F.; Wu, J.; Pavlovic, N.; Breton, C.; Gilliland, F.; Habre, R. Spatiotemporal imputation of MAIAC AOD using deep learning with downscaling. Remote Sens. Environ. 2020, 237. [CrossRef] [PubMed]

51. Fang, Y.; Li, L. Estimation of high-precision high-resolution meteorological factors based on machine learning. J. Geo-Inf. Sci. 2019, 21, 799–813.

52. Li, L. Geographically Weighted Machine Learning and Downscaling for High-Resolution Spatiotemporal Estimations of Wind Speed. Remote Sens. 2019, 11, 1378. [CrossRef]

53. Wikipedia. One-Hot. Available online: https://en.wikipedia.org/wiki/One-hot (accessed on 20 June 2019).

54. Press, H.W.; Teukolsky, A.S.; Vetterling, T.W.; Flannery, P.B. Numerical Recipes in C: The Art of Scientific Computing, 2nd ed.; Cambridge University Press: New York, NY, USA, 1992; pp. 123–128.

55. Devlin, J.; Chang, M.W.; Lee, K. Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv 2018, arXiv:1810.04805. [CrossRef]

56. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. J. R. Stat. Soc. B 2005, 67, 768. [CrossRef]

57. Baydin, G.A.; Pearlmutter, B.; Radul, A.A.; Siskind, J. Automatic differentiation in machine learning: A survey. J. Mach. Learn. Res. 2018, 18, 1–43.

58. Malone, P.B.; McBratney, B.A.; Minasny, B.; Wheeler, I. A general method for downscaling earth resource information. Comput. Geosci. 2012, 41, 119–125. [CrossRef]

59. Keshari, R.; Ghosh, S.; Chhabra, S.; Vatsa, M.; Singh, R. Unravelling Small Sample Size Problems in the Deep Learning World. In Proceedings of the 2020 IEEE Sixth International Conference on Multimedia Big Data (Bigmm 2020), New Delhi, India, 24–26 September 2020; pp. 134–143. [CrossRef]

60. Gonzalez, J.A.; Calbo, J. Aerosol Effects on the Cloud Optical Depth Retrieval from Atmospheric Transmittance. AIP Conf. Proc. 2013, 1531, 556–559. [CrossRef]

61. Yuan, T.L.; Li, Z.Q.; Zhang, R.Y.; Fan, J.W. Increase of cloud droplet size with aerosol optical depth: An observation and modeling study. J. Geophys. Res. Atmos. 2008, 113. [CrossRef]

62. Li, J.; Carlson, E.B.; Lacis, A.A. How well do satellite AOD observations represent the spatial and temporal variability of PM2.5 concentration for the United States? Atmos. Environ. 2015, 102, 260–273. [CrossRef]

63. Wang, H.F.; Li, Z.Q.; Lv, Y.; Xu, H.; Li, K.T.; Li, D.H.; Hou, W.Z.; Zheng, F.X.; Wei, Y.Y.; Ge, B.Y. Observational study of aerosol-induced impact on planetary boundary layer based on lidar and sunphotometer in Beijing. Environ. Pollut. 2019, 252, 897–906. [CrossRef] [PubMed]

64. Seidel, D.J.; Ao, C.O.; Li, K. Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis. J. Geophys Res. Atmos. 2010, 115. [CrossRef]

65. Su, T.N.; Li, Z.Q.; Kahn, R. Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: Regional pattern and influencing factors. Atmos. Chem. Phys. 2018, 18, 15921–15935. [CrossRef]

66. Liu, Z.; Mortier, A.; Li, Z.Q.; Hou, W.Z.; Goloub, P.; Lv, Y.; Chen, X.F.; Li, D.H.; Li, K.T.; Xie, Y.S. Improving Daytime Planetary Boundary Layer Height Determination from CALIOP: Validation Based on Ground-Based Lidar Station. Adv. Meteorol. 2017, 2017. [CrossRef]

67. Jordan, N.S.; Hoff, R.M.; Bacmeister, J.T. Validation of Goddard Earth Observing System-version 5 MERRA planetary boundary layer heights using CALIPSO. J. Geophys Res. Atmos. 2010, 115. [CrossRef]

68. Molod, A.; Salmins, H.; Collof, A. The Use of MERRA-2 Near Surface Meteorology to Understand the Behavior of Planetary Boundary Layer heights Derived from Wind Profiler Data Over the US Great Plains. In Proceedings of the American Geophysical Union, Fall Meeting 2017, New Orleans, LA, USA, 15–17 December 2017.

69. Yang, Q.Q.; Yuan, Q.Q.; Yue, L.W.; Li, T.W.; Shen, H.F.; Zhang, L.P. The relationships between PM2.5 and aerosol optical depth (AOD) in mainland China: About and behind the spatio-temporal variations. Environ. Pollut. 2019, 248, 526–535. [CrossRef]

70. Chugh, N. Three Reasons Why You Should Bet On Guangzhou as the Next Big Global City. Available online: https://www.forbes.com/sites/nishthachugh/2017/06/15/three-reasons-why-you-should-bet-on-guangzhou-as-next-big-global-city/?sh=6430498c2dc9 (accessed on 7 January 2021).

71. Baidu. Beijing Clean Air Action Plan for 2013–2017 (in Chinese). Available online: https://baike.baidu.com/item/%E5%8C%97%EF%BC%8A%E6%88%91%E7%BD%9A%E6%80%94%E8%A1%8C%E5%8A%A8%E8%AE%A1%E5%88%92/10348277?fr=aladdin (accessed on 1 February 2021).
72. Huo, K.; Zhang, Q.; Tong, Z.; Cheng, J.; Liu, Y. *Ways and Health Benefits of Improving Air Quality in China in the Medium and Long Term*; Qinghua University: Beijing, China, 2020. (In Chinese)

73. Ministry of Ecology and Environment of People’s Republic of China. The National Air Quality Forecast Consultation in Mid-to-Early January 2019 (in Chinese). 2018. Available online: https://www.mee.gov.cn/xxgk2018/xxgk/xxgk15/201812/t20181230_688171.html (accessed on 10 March 2019).

74. Li, L.; Wu, J. *A Complete Aerosol Optical Depth Dataset with High Spatiotemporal Resolution for Mainland China*, 1st ed.; IGSNRR, Ed.; Harvard Dataverse: Cambridge, MA, USA, 2020. [CrossRef]

75. Li, L.; Wu, J. *Mapping of Ground Aerosol Coefficient with High Spatiotemporal Resolution for Mainland China*; Harvard Dataverse: Cambridge, MA, USA, 2021. [CrossRef]