

# MACHINE LEARNING BASED BIAS CORRECTION FOR MODIS AEROSOL OPTICAL DEPTH IN BEIJING

Mengjuan Wang <sup>1</sup>, Meng Fan <sup>2</sup>, Zhibao Wang <sup>1,3\*</sup>, Liangfu Chen <sup>2</sup>, Lu Bai <sup>4</sup>, Yuanlin Chen <sup>1</sup>, Mei Wang <sup>1</sup>

<sup>1</sup> School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China

<sup>2</sup> State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute of Chinese Academy of Sciences, Beijing 100101, China

<sup>3</sup> Bohai-Rim Energy Research Institute, Northeast Petroleum University, Qinhuangdao 066004, China

<sup>4</sup> School of Computing, Ulster University, Belfast BT15 1ED, UK

**KEY WORDS:** Aerosol optical depth, bias correction, machine learning, artificial neural network, support vector regression.

## ABSTRACT:

Aerosol refers to suspensions of small solid and liquid particles in the atmosphere. Although the content of aerosol in the atmosphere is small, it plays a crucial role in atmospheric and the climatic processes, making it essential to monitor. In areas with poor aerosol characteristics, satellite-based aerosol optical depth (AOD) values often differ from ground-based AOD values measured by instruments like AERONET. The use of 3km DT, 10km DT and 10km DTB algorithms in Beijing area has led to significant overestimation of AOD values, highlighting the need for improvement. This paper proposes the use of machine learning techniques, specifically support vector regression (SVR) and artificial neural network (ANN), to correct the deviation of AOD data. Our approach leverages ground-based monitoring data, meteorological reanalysis data and satellite products to train the models. Our results show that the ANN model outperforms the SVR model achieving R<sup>2</sup>, RMSE and Slope values of 0.88, 0.12 and 0.97, respectively, when applied to nearly two decades of data from 2001 to 2019. This study significantly improves the accuracy of MODIS AOD values, reducing overestimation and bringing them closer to ground-based AOD values measured by AERONET. Our findings have important applications in climate research and environmental monitoring.

## 1. INTRODUCTION

### 1.1 Introduction

Aerosols are tiny particles that can be solid, liquid, or a mixture of both, with an aerodynamic diameter of fewer than 100 µm. When carried by the atmosphere, they are referred to as atmospheric aerosols. Despite their small quantity, aerosols have a significant impact on the physical and chemical processes in the atmosphere, as well as a crucial role in the climate system (Luo et al., 2014). Aerosols particles change the solar radiation reaching the Earth's surface by scattering and absorbing solar short-wave and Earth long-wave radiation, thereby affecting the climate. Additionally, aerosol particles can indirectly influence the climate by acting as cloud particles, which can that and alter the concentration of various chemical components in the atmosphere and contribute to the cloud formation. Aerosols have substantial impacts on the Earth's radiation budget balance, cloud and precipitation formation, regional air pollution visibility, human health (Colbeck and Lazaridis, 2010). The occurrence of winter haze weather significantly reduces the visibility of cities, causing inconvenience in people's daily lives. Recent studies have indicated that Particulate Matter (PM) in polluted air, particularly inhalable fine particles like PM<sub>2.5</sub>, is linked to heart and respiratory diseases (T. Li et al., 2018; Ni et al., 2015; Tecer et al., 2008). These particles can pose a threat to human health. Research shows that aerosol and PM<sub>2.5</sub> have strong correlation (Guo et al., 2017; Li et al., 2015). Therefore, monitoring and controlling the atmospheric environment are critical.

Aerosol optical depth (AOD) is an important parameter for characterizing the optical properties of atmospheric aerosols. It represents the integrated extinction coefficient over the entire atmosphere in the vertical direction and is the primary optical parameter that can be retrieved using ground-based and satellite remote sensing methods. Satellite detection technology is increasingly being used to monitor and assess environmental changes (Li et al., 2020; Shi et al., 2022; Yang et al., 2013; J. Zhang et al., 2021). Satellites equipped with various sensors and cameras can capture detailed images and data of the Earth's surface, atmosphere, and oceans (Bovensmann et al., 2010; Yu et al., 2021; X. Zhang et al., 2021a, 2021b; Zhu et al., 2021). While ground-based observation provides high temporal resolution data on aerosol characteristics at a specific location, it has limited spatial coverage and cannot monitor large areas (Z. Li et al., 2018). In contrast, satellite-based monitoring of aerosols is advantageous due to its wide range, low cost and lack of geographical constraints (Dubovik et al., 2019). As satellite detection technology and capabilities continue to improve, the role of satellite remote sensing in atmospheric aerosol research is also increasing. Thus, maximising the potential of satellite remote sensing for AOD-related research real-life applications is of paramount importance.

Satellite remote sensing has become increasing valuable for filling gaps in AOD data where the characteristics are not readily apparent. However, the accuracy of satellite products is limited by assumptions regarding aerosol and surface properties, and AOD adjustment through topography and canopy (Weber et al., 2010). However, ground measurements are not subject to these limitations, which is why satellite

\* Corresponding author

derived AOD values are often verified by AERONET AOD values (Estellés et al., 2012). Currently, Moderate Resolution Imaging Spectroradiometer (MODIS) is among the most used sensors for aerosol detection. MODIS is one of the primary sensors on the Terra and Aqua satellites, which collaborate to repeatedly observe the Earth’s surface every 1-2 days.

Studies indicate that C6.1 dark target algorithm (3km, 10km DT) and a combination of the dark target and deep blue algorithm (10km DTB) AOD products in Beijing are often overestimated(Wei and Sun, 2016). Nevertheless, there is still room for improvement in the accuracy of MODIS AOD (AODM). An effective approach is to use satellite observations as input and ground observations as output to build models based on both sources. These models can correct AOD values obtained from satellite observations, with sufficient training data. The models can be adapted to different retrieval scenarios and outperform deterministic algorithms. Machine learning methods satisfy these requirements and have been widely adopted in the aerosol science. For instance, (Lanzaco et al., 2016a) used machine learning techniques to correct the AOD value at 550nm, resulting in a significant improvement in the AOD value obtained from satellite data. (Palancar et al., 2016) used a similar method to obtain AOD values of 340 nm from MODIS and calculated the aerosol radiative forcing in the UV-B region, and the obtained values were consistent with those obtained using AERONET AOD (AODA) as input. These studies use AODA to correct AODM biases but are not applicable in areas with low AERONET sites density or for different underlying surfaces, which limits their reliability on a global scale.

In this paper, a method to improve AODM is proposed using AODA as reference. AOD products of Three different algorithms of AOD products, including 3km DT,10km DT, 10km DTB from Terra and Aqua satellites in Beijing area, are employed in this method. The AERONET site data in Beijing area is used as the ground truth for bias correction. The method leverages machine learning technology to improve the correlation between AODM and AODA and trains the model using three types of AOD inversion products from Beijing area to correct the overestimation phenomenon. The method preserves abnormal data in the dataset to make it applicable to original data. Moreover, the method requires only annual average trend data of meteorological variables, meteorological conditions as possible inputs to keep the requirements minimal.

## 2. DATASET

### 2.1 AERONET data

AERONET (“AERONET,” n.d.) is a federation of ground-based remote sensing aerosol networks jointly established by NASA and LOA - PHOTONS (CNRS), provides high-quality measurements of aerosol optical properties (Holben et al., 1998). With around 500 sites worldwide, it represents a key resource for aerosol observation. The aerosol products derived from AERONET observations are classified into three levels: Level 1.0, Level 1.5, and Level 2.0. In this study, the Level 2.0 data from four aerosol sites in Beijing were used. The detailed information of these sites is presented in Table 1.

**Table 1.** Statistics of AERONET ground stations in China

Site name	longitude, latitude	observation time	MODIS Terra/Aqua data volume
-----------	------------------------	---------------------	---------------------------------

Beijing	116.381, 39.977	2001.05- 2018.11	438/361
Beijing_	116.31, PKU	2016.07- 2019.5	134/107
Beijing_	116.379, RADI	2010.05- 2019.05	67/166
Beijing-	116.317, CAMS	2012.08- 2019.05	261/189

AERONET does not provide the AOD observation value at 550 nm band. However, the 550 nm Ångström index can be obtained by interpolation of the Ångström index at 500 nm and 675nm bands. The estimation of AOD at 550 nm can be obtained by using the equations (1) and (2), as proposed by (Eck et al., 1999).

$$\alpha_{500-675} = -\frac{\ln(\tau_{500}/\tau_{675})}{\ln(500/675)}, \quad (1)$$

$$\tau_{550} = \tau_{675}(500/675)^{-\alpha_{500-675}}, \quad (2)$$

In the above equation:  $\tau_{550}$  and  $\tau_{675}$  are the AOD values of 500nm and 657nm respectively,  $\alpha_{500-675}$  is the Ångström index of 500-657nm,  $\tau_{550}$  is the AOD value of 550nm obtained by interpolation.

### 2.2 MODIS data

The United States National Aeronautics and Space Administration (NASA) launched a series of comprehensive Earth observation satellites in 1991. Terra and Aqua are two solar polar orbit synchronous satellites that were launched in 1999 and 2002 respectively. They were known as “morning star” and “afternoon star” because of their transit times of 10:30 a.m. and 1:30 p.m., respectively. These satellites can acquire daily global coverage data using a spectrometer with a measurement range from 0.4µm to 14.4µm. Since aerosol absorption levels in the mid-infrared region are low, surface reflectance can be determined, from which AOD can be derived (Lanzaco et al., 2017). However, different inversion algorithms may yield different AOD results.

To correct for bias in the AOD products, this work uses the AOD products of the Terra (from March 2001 to May 2019) and Aqua satellite (from July 2002 to May 2019) MODIS 3km DT,10km DT, 10km DTB. The corrected data is verified using the AOD data from AERONET stations in Beijing to ensure accuracy.

### 2.3 Meteorological data

Meteorological conditions play an essential role in predicting AOD, and they have a strong influence on the concentration distribution, chemical composition, and optical properties of AOD. Variations in AOD have been found to be significantly linked to meteorological factors. Therefore, in this study, the fifth-generation global climate reanalysis dataset ERA5 (“Climate Data Store,” n.d.), was published by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 provides hourly atmospheric, terrestrial, and oceanic climate variables, terrestrial and oceanic climate variables at a spatial resolution of 0.25° × 0.25° (Jiang et al., 2021). In this study, meteorological parameters extracted include the U wind

component (u10) and V wind component vectors (v10), 2m dew point temperature (d2m), 2m air temperature (t2m), surface temperature (skt), and atmospheric pressure (sp). Uv10 is calculated based on u10 and v10.

To capture the annual trend of meteorological conditions, we use the day of the year as a parameter. However, since January 1 and December 31 have the same average, this parameter is modified to create a more representative parameter. Specifically, we set the variable to 1 on January 1 and increase it to a maximum of 183 on July 1 and 2, then decrease it again to a minimum on December 31 to better reflect the seasonality of meteorological conditions (Lanzaco et al., 2016b).

## 2.4 Data pre-processing

The feature sample set was constructed from data gathered between 2001 to 2019, and a range of different features were used for model training. These parameters are listed in Table 2. Time data features were specified by selecting the year, month, and day. The transit times for Terra and Aqua satellites are 10:30 AM and 1:30 PM, respectively. In order to maintain consistency in timing, the ERA5 data was selected are 10:00 and 13:00 because the time resolution of the ERA5 data is in hours.

**Table 2.** Feature information table

Type	Data	Temporal resolution	Spatial resolution	Data origin
Ground station data	AERONET _AOD	15 minutes	—	AER
				OSOL
				ROB
				OTIC
				NET
Satellite data	3km DT AOD	1 day	3km	MODIS
	10 km DT AOD	1 day	10km	MODIS
	10km DTB AOD	1 day	10km	MODIS
Meteorological data	uv10	1 hour	0.25°×0.25°	ERA5
	d2m	1 hour	0.25°×0.25°	ERA5
	t2m	1hour	0.25°×0.25°	ERA5
	skt	1 hour	0.25°×0.25°	ERA5
	sp	1hour	0.25°×0.25°	ERA5
Time	days	1day	—	—
	year	—	—	—
	month	—	—	—
Geographic data	lons	—	—	—
	lats	—	—	—

## 3. METHODOLOGY

Machine learning is a subfield of artificial intelligence that focuses on learning from data to improve predictions. In this study, two types of machine learning algorithms, Artificial Neural Network (ANN) and Support Vector Regression (SVR)

were used to correct the bias between AODM and AODA. The datasets used in this study were MODIS Terra and MODIS Aqua, which were divided into training and test sets. For each machine learning algorithm, the training set was split into “input” and “output” variables, and the algorithm searched for the relationships between them.

The goal of the study was to obtain AODM that closely matched AODA. The inputs used for the machine learning models included the three AODM products, meteorological variables, time (year, month, and day), latitude and longitude, and the annual average trend of meteorological conditions, while AODA was used as the output variable.

### 3.1 ANN

ANN is a machine learning method inspired by biological systems (Basheer and Hajmeer, 2000). It is a complex network comprising interconnected neurons that work in parallel. In a typical ANN structure, the input is connected to one or more layers of neurons, which are connected to the output. The connection weights between the neurons are parameters that are adjusted during the training process. The training process is iterative, where the Root Mean Square Error (RMSE) is calculated in each cycle and the connection weights are adjusted based on the results until the network can correctly predict the output. This type of process is called supervised learning because it provides realistic outputs.

When training an ANN model, the dataset is randomly split into a training set and a test set. The training set is used to train the weights of the neural network while the test set is used to evaluate the performance of the trained model on new, unseen data. The convergence of the training process is evaluated using the RMSE and the number of training iterations. Once the model is trained, a scatter plot is generated from the test set data, and the regression line is determined to evaluate the accuracy of the predicted outputs. This validation is rigorous and independent since the test set is randomly selected from all available temporal and spatial datasets.

### 3.2 SVR

Support Vector Machine (SVM) was first introduced by Cortes and Vapnik and developed into classification and regression problems (Sain, 1996). SVM is widely accepted in the field of learning because of its ability to generalize. One of its important features is its capability to project linearly inseparable data into a high dimensional feature space, making it linearly separable. SVMs were later extended to regression problems, named SVR (Schölkopf et al., 2000; Smola and Schölkopf, 2004). In this study we used the SVR method proposed by LIBSVM (Chen, Fan and Lin 2006, Fan et al. 2005).

The goal of the SVR is to find a function that has a minimum deviation from the output of the training dataset, while also maintaining a flatness to control the complexity of the system and the training error. The penalty factor of the model is continuously adjusted through the training set, and the coefficient of determination (R2) and RMSE of the current model are evaluated in each iteration of training. The convergence of the training is determined by using the coefficient of R2, RMSE, and the penalty factor training iterations.

### 3.3 Technical flow chart

In this study, two machine learning algorithms: ANN and SVR were used to correct the bias between AODM and AODA. This study aims to approximate the AODM as close as possible to the AODA. Two dataset - MODIS Terra and MODIS Aqua were used together with meteorological data. The sample set is created by spatial temporal matching the features from satellite data with ground station (AERONET) data. As shown in Figure 1, machine learning algorithms were used to map the features inputs to correct AOD.

### 3.4 Correlation analysis

Feature selection was performed based on the correlation analysis, as shown in Figure 2. Pearson correlation coefficients were calculated between ground stations and all characteristic

parameters. The AODA was positively correlated with 3km DT, 10km DT and 10km DTB, and the correlation exceeded 0.9. Temperature affects the Brownian motion of aerosol and the vertical distribution of aerosol, so that the change of its mass concentration satisfies the negative exponential law. All meteorological characteristic parameters were positively correlated. Among them, the strongest correlation was d2m with a correlation coefficient of 0.4, and the AODA was positively correlated with longitude (lons) and latitude (lats). Moreover, according to Figure 2, the correlation between year, month, sp and AODA was negative, and they were removed from the sample. Based on the results from correlation analysis, the final selected features train the ANN and SVR models are 3km DT, 10km DT, 10km DTB, day, lons, lats, uv10, d2m, t2m, skt, and days. The data was split into a training set of 70% and a test set of 30%.

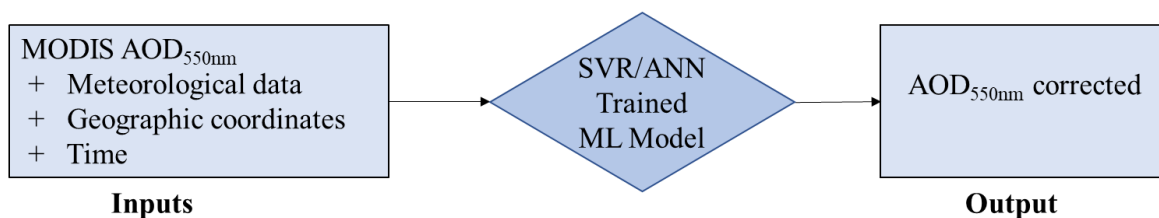


Figure 1. Flowchart of the proposed bias correction method

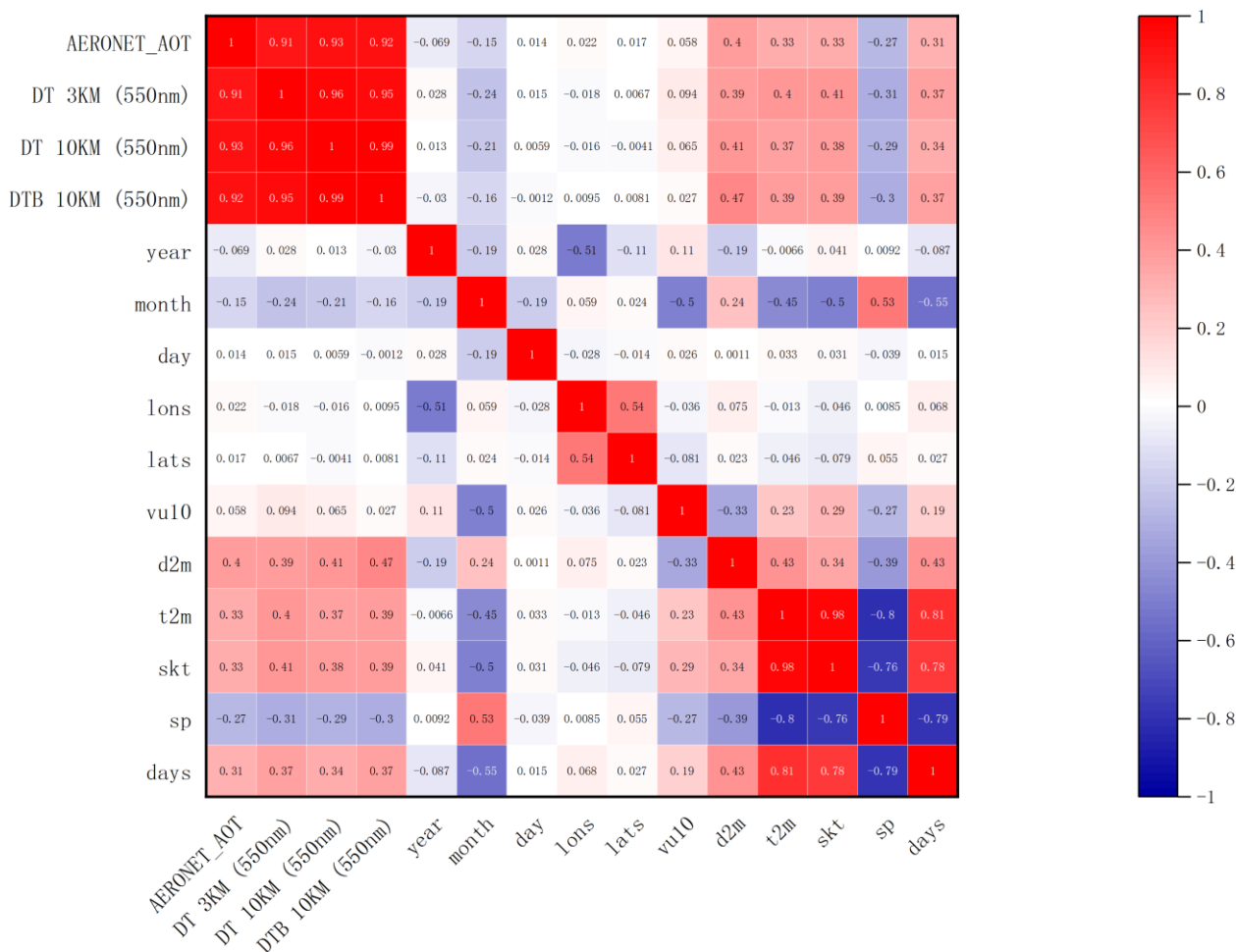


Figure 2. Correlation matrix of AOD Features

#### 4. RESULTS AND DISCUSSIONS

##### 4.1 Comparison results of AOD values from MODIS data and AERONET

We first compare the AODM with AODA. Three MODIS AOD products that were used in this study, including 3km DT, 10km DT, and 10km DTB in Beijing region. The results are shown in

Figure 3, R2 of the 3km DT Terra dataset is 0.578, and that of the Aqua dataset is 0.167. R2 of the 10km DT Terra dataset is 0.639, and that of Aqua is 0.585. R2 of the 10km DTB Terra dataset is 0.722, and that of Aqua is 0.657. Compared with the AODA, AODM is overestimated and needs to be improved. Therefore, it is necessary to propose a method to improve the accuracy of AODM.

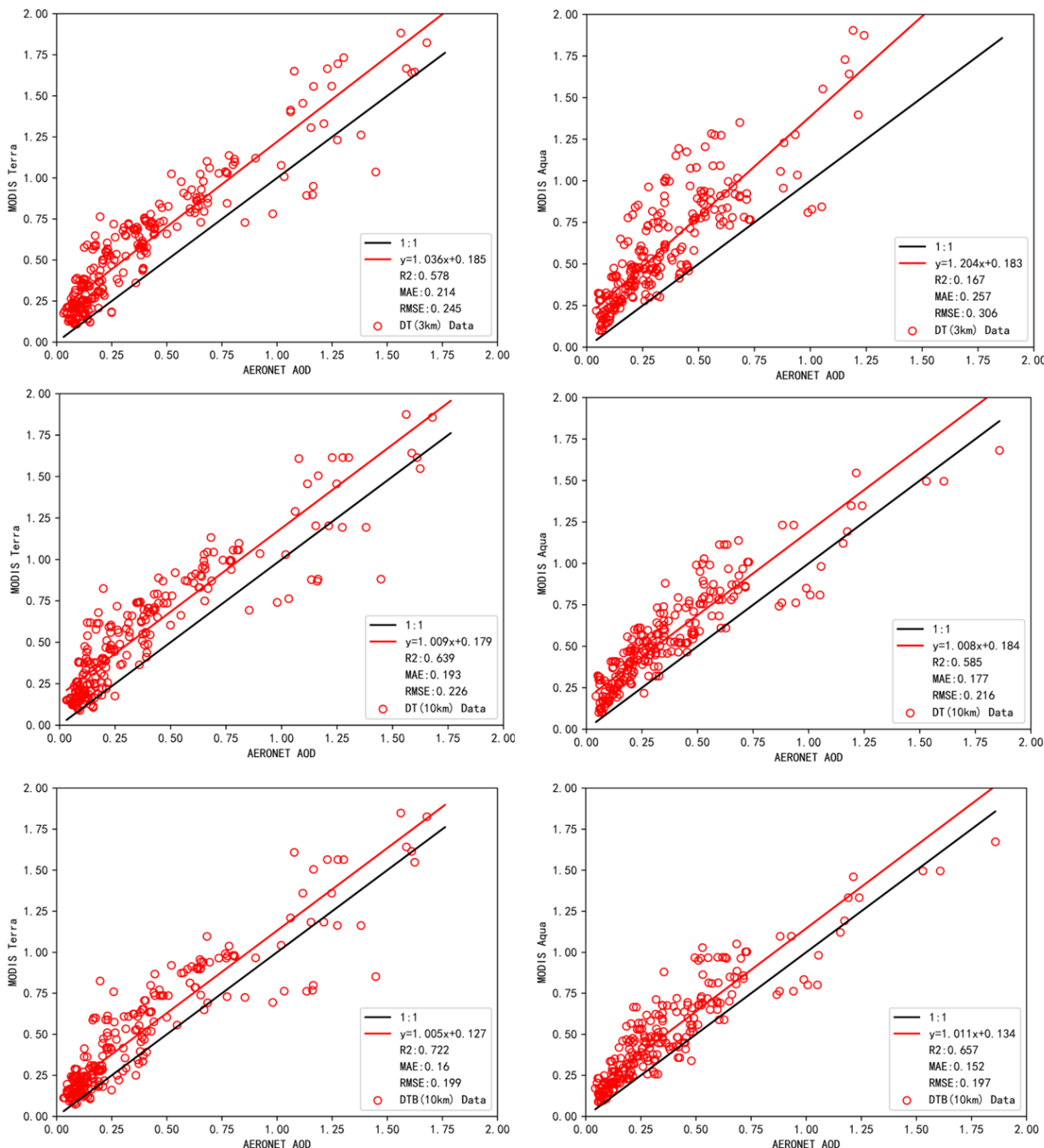


Figure 3. Comparison between 3km DT, 10km DT, and 10km DTB AOD products with AERONET AOD products

##### 4.2 Determine the optimal parameters

To evaluate the bias correction effect of ANN and SVR model in AOD model, the training set and test set are used, and the

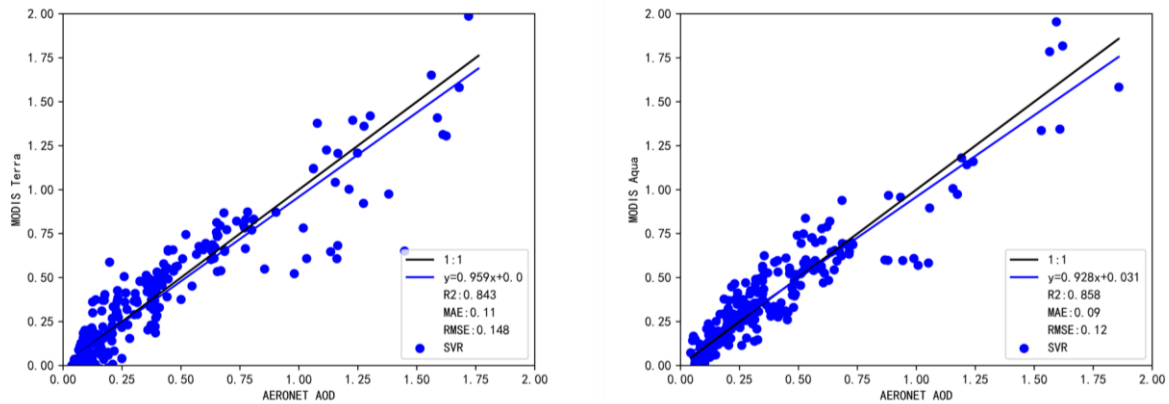
accuracy is verified by 10-fold cross validation method. In the process of model training, the parameters of the model are fine-tuned to finally determine the optimal parameters. The specific parameters are shown in Table 3.

**Table 3.** The model optimizes the relevant parameters

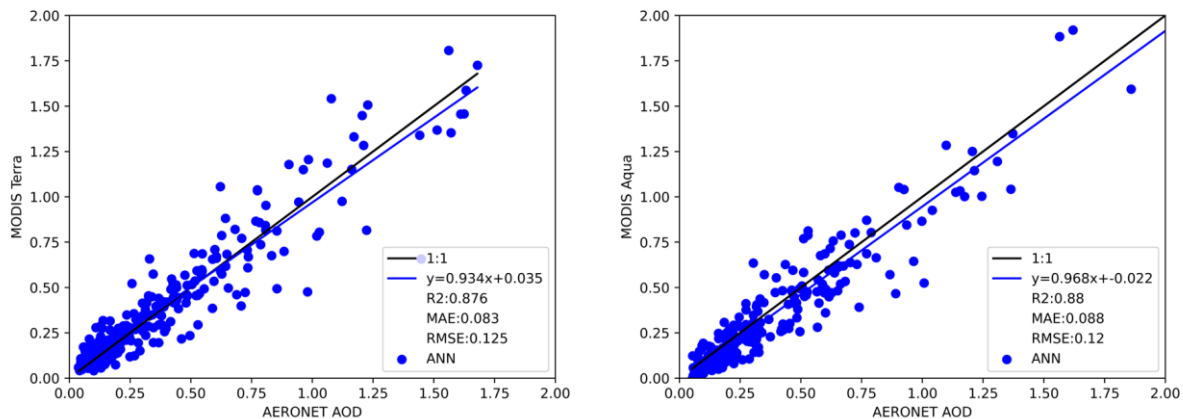
Algorithm name	Name of parameter	Value
ANN	Number of input layer neurons	40
	Number of hidden layer neurons	20
	Number of output layer neurons	1
	Number of iterations	1000
SVR	Penalty factor	5000

### 4.3 Bias correction results

After applying our proposed machine learning based bias correction method, the overestimation has been significantly improved. The results from SVR and ANN models are shown in Figures 4 and 5 respectively. After correction, R2 of the Terra dataset is 0.843 (SVR) and 0.876 (ANN), MAE is 0.11 and 0.083, and RMSE is 0.148 and 0.125. R2 of the Aqua dataset is 0.858 and 0.88, MAE is 0.09 and 0.088, and RMSE is 0.12 and 0.12. Based on the above analysis, it is noted that the deviation correction effect of ANN on AOD is better than that of the SVR method.



**Figure 4.** SVR model deviation correction results



**Figure 5.** ANN model deviation correction results

## 5. CONCLUSIONS

In this study, we propose a machine learning based method to correct MODIS DT AOD at 3km, 10km, and MODIS DTB at 10km. MODIS AOD data and meteorological data are used as input to the proposed machine learning models, including SVR and ANN. Four ground stations from AERONET located in Beijing are used as ground truth. The results show that our proposed methods significantly improve the overestimation from the MODIS AOD and ANN performance is better than SVR. In future, this machine learning based method can also be applied to other parts of the world, especially in areas with poor aerosol characteristics, such as Africa and Australia, where the reliability of MODIS algorithm is relatively poor. Extending the practicability of this method to define a region around each site and train the model for each region separately can improve this

research to a certain extent. Estimation of PM2.5 and PM10 based on satellite measurements is a topic very relevant to atmospheric science, and there is a strong correlation between aerosol and PM. By using the measurement data of particulate matter in a certain area and combining with the method used in this study, it is possible to obtain better PM2.5 and PM10 from satellite measurement data.

## REFERENCES

- AERONET [WWW Document], n.d. URL <https://aeronet.gsfc.nasa.gov/> (accessed 2.20.23).
- Basheer, I.A., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design, and application. *J. Microbiol. Methods* 43, 3–31.

- Bovensmann, H., Buchwitz, M., Burrows, J.P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretnér, A., Erzinger, J., 2010. A remote sensing technique for global monitoring of power plant CO<sub>2</sub> emissions from space and related applications. *Atmospheric Meas. Tech.* 3, 781–811. <https://doi.org/10.5194/amt-3-781-2010>
- Climate Data Store [WWW Document], n.d. URL <https://cds.climate.copernicus.eu#!/home>
- Colbeck, I., Lazaridis, M., 2010. Aerosols and environmental pollution. *Naturwissenschaften* 97, 117–131.
- Dubovik, O., Li, Z., Mishchenko, M.I., Tanré, D., Karol, Y., Bojkov, B., Cairns, B., Diner, D.J., Espinosa, W.R., Goloub, P., 2019. Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives. *J. Quant. Spectrosc. Radiat. Transf.* 224, 474–511.
- Eck, T.F., Holben, B., Reid, J., Dubovik, O., Smirnov, A., O’neill, N., Slutsker, I., Kinne, S., 1999. Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *J. Geophys. Res. Atmospheres* 104, 31333–31349.
- Estellés, V., Campanelli, M., Utrillas, M., Expósito, F., Martínez-Lozano, J., 2012. Comparison of AERONET and SKYRAD4. 2 inversion products retrieved from a Cimel CE318 sunphotometer. *Atmospheric Meas. Tech.* 5, 569–579.
- Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M., Zhai, P., 2017. Impact of diurnal variability and meteorological factors on the PM<sub>2.5</sub> - AOD relationship: Implications for PM<sub>2.5</sub> remote sensing. *Environ. Pollut.* 221, 94–104. <https://doi.org/10.1016/j.envpol.2016.11.043>
- Holben, B.N., Eck, T.F., Slutsker, I. al, Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y., Nakajima, T., 1998. AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.* 66, 1–16.
- Jiang, T., Chen, B., Nie, Z., Ren, Z., Xu, B., Tang, S., 2021. Estimation of hourly full-coverage PM<sub>2.5</sub> concentrations at 1-km resolution in China using a two-stage random forest model. *Atmospheric Res.* 248, 105146.
- Lanzaco, B.L., Olcese, L.E., Palancar, G.G., Toselli, B.M., 2017. An improved aerosol optical depth map based on machine-learning and MODIS data: Development and application in South America.
- Lanzaco, B.L., Olcese, L.E., Palancar, G.G., Toselli, B.M., 2016a. A method to improve MODIS AOD values: Application to South America.
- Lanzaco, B.L., Olcese, L.E., Palancar, G.G., Toselli, B.M., 2016b. A method to improve MODIS AOD values: Application to South America.
- Li, J., Carlson, B.E., Lacis, A.A., 2015. How well do satellite AOD observations represent the spatial and temporal variability of PM<sub>2.5</sub> concentration for the United States? *Atmos. Environ.* 102, 260–273. <https://doi.org/10.1016/j.atmosenv.2014.12.010>
- Li, J., Pei, Y., Zhao, S., Xiao, R., Sang, X., Zhang, C., 2020. A Review of Remote Sensing for Environmental Monitoring in China. *Remote Sens.* 12, 1130. <https://doi.org/10.3390/rs12071130>
- Li, T., Hu, R., Chen, Z., Li, Q., Huang, S., Zhu, Z., Zhou, L.-F., 2018. Fine particulate matter (PM<sub>2.5</sub>): The culprit for chronic lung diseases in China. *Chronic Dis. Transl. Med.* 4, 176–186.
- Li, Z., Xu, H., Li, K., Li, D., Xie, Y., Li, L., Zhang, Y., Gu, X., Zhao, W., Tian, Q., 2018. Comprehensive study of optical, physical, chemical, and radiative properties of total columnar atmospheric aerosols over China: An overview of Sun–Sky Radiometer Observation Network (SONET) measurements. *Bull. Am. Meteorol. Soc.* 99, 739–755.
- Luo, Y., Zheng, X., Zhao, T., Chen, J., 2014. A climatology of aerosol optical depth over China from recent 10 years of MODIS remote sensing data. *Int. J. Climatol.* 34, 863–870.
- Ni, L., Chuang, C.-C., Zuo, L., 2015. Fine particulate matter in acute exacerbation of COPD. *Front. Physiol.* 6, 294.
- Palancar, G.G., Olcese, L.E., Lanzaco, B.L., Achad, M., López, M.L., Toselli, B.M., 2016. Aerosol radiative forcing efficiency in the UV-B region over central Argentina. *Atmospheric Res.* 176, 1–9.
- Sain, S.R., 1996. The nature of statistical learning theory.
- Schölkopf, B., Smola, A.J., Williamson, R.C., Bartlett, P.L., 2000. New support vector algorithms. *Neural Comput.* 12, 1207–1245.
- Shi, K., Bai, L., Wang, Z., Tong, X., Mulvenna, M.D., Bond, R.R., 2022. Photovoltaic Installations Change Detection from Remote Sensing Images Using Deep Learning, in: *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*. Presented at the IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Kuala Lumpur, Malaysia, pp. 3231–3234. <https://doi.org/10.1109/IGARSS46834.2022.9883738>
- Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. *Stat. Comput.* 14, 199–222.
- Tecer, L.H., Alagha, O., Karaca, F., Tuncel, G., Eldes, N., 2008. Particulate matter (PM<sub>2.5</sub>, PM<sub>10-2.5</sub>, and PM<sub>10</sub>) and children’s hospital admissions for asthma and respiratory diseases: a bidirectional case-cross-over study. *J. Toxicol. Environ. Health A* 71, 512–520.
- Weber, S.A., Engel-Cox, J.A., Hoff, R.M., Prados, A.I., Zhang, H., 2010. An improved method for estimating surface fine particle concentrations using seasonally adjusted satellite aerosol optical depth. *J. Air Waste Manag. Assoc.* 60, 574–585.
- Wei, J., Sun, L., 2016. Comparison and evaluation of different MODIS aerosol optical depth products over the Beijing-Tianjin-Hebei region in China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 10, 835–844.
- Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., Dickinson, R., 2013. The role of satellite remote sensing in climate change studies. *Nat. Clim. Change* 3, 875–883.

Yu, Z., Wang, Z., Bai, L., Chen, L., Tao, J., 2021. Remote Sensing Inversion of PM10 Based on Spark Platform, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Presented at the IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Brussels, Belgium, pp. 1685–1688. <https://doi.org/10.1109/IGARSS47720.2021.9554323>

Zhang, J., Wang, Z., Bai, L., Song, G., Tao, J., Chen, L., 2021. Deforestation Detection Based on U-Net and LSTM in Optical Satellite Remote Sensing Images, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Presented at the IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Brussels, Belgium, pp. 3753–3756. <https://doi.org/10.1109/IGARSS47720.2021.9554689>

Zhang, X., Zhang, Y., Bai, L., Tao, J., Chen, L., Zou, M., Han, Z., Wang, Z., 2021a. Retrieval of Carbon Dioxide Using Cross-Track Infrared Sounder (CrIS) on S-NPP. *Remote Sens.* 13, 1163.

Zhang, X., Zhang, Y., Lu, X., Bai, L., Chen, L., Tao, J., Wang, Z., Zhu, L., 2021b. Estimation of lower-stratosphere-to-troposphere ozone profile using long short-term memory (LSTM). *Remote Sens.* 13, 1374.

Zhu, S., Zhu, H., Xu, J., Zeng, Q., Zhang, D., Liu, X., 2021. Satellite remote sensing of daily surface ozone in a mountainous area. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5.