TU Delft

Delft University of Technology

Aerosol Absorption over Land Derived from the Ultra-Violet Aerosol Index by Deep Learning

Sun, Jiyunting; Veefkind, Pepijn; Van Velthoven, Peter; Levelt, Pieternel F.

DOI
10.1109/JSTARS.2021.3108669

Publication date
2021

Document Version
Final published version

Published in
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

Citation (APA)
Sun, J., Veefkind, P., Van Velthoven, P., & Levelt, P. F. (2021). Aerosol Absorption over Land Derived from the Ultra-Violet Aerosol Index by Deep Learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 9692-9710. Article 9525298. https://doi.org/10.1109/JSTARS.2021.3108669

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Aerosol Absorption Over Land Derived From the Ultra-Violet Aerosol Index by Deep Learning

Jiyunting Sun, Pepijn Veefkind, Peter van Velthoven, and Pieternel F. Levelt

***Abstract*—Quantitative measurements of aerosol absorptive properties, e.g., the absorbing aerosol optical depth (AAOD) and the single scattering albedo (SSA), are important to reduce uncertainties of aerosol climate radiative forcing assessments. Currently, global retrievals of AAOD and SSA are mainly provided by the ground-based aerosol robotic network (AERONET), whereas it is still challenging to retrieve them from space. However, we found the AERONET AAOD has a relatively strong correlation with the satellite retrieved ultra-violet aerosol index (UVAI). Based on this, a numerical relation is built by a deep neural network (DNN) to predict global AAOD and SSA over land from the long-term UVAI record (2006–2019) provided by the ozone monitoring instrument (OMI) onboard Aura. The DNN predicted aerosol absorption is satisfying for samples with AOD at 550 nm larger than 0.1, and the DNN model performance is better for smaller absorbing aerosols (e.g., smoke) than larger ones (e.g., mineral dust). The comparison of the DNN predictions with AERONET shows a high correlation coefficient of 0.90 and a root mean square of 0.005 for the AAOD, and over 80% of samples are within the expected uncertainty of AERONET SSA ($\pm$0.03).**

***Index Terms*—Absorbing aerosol optical depth (AAOD), deep neural network (DDN), machine learning, ozone monitoring instrument (OMI), single scattering albedo (SSA), ultra-violet aerosol index (UVAI).**

I. Introduction

ATMOSPHERIC aerosols are solid or liquid particles suspended in the air. Most aerosols have a dominating cooling effect on the Earth's climate (-0.9W/m^2 [-1.9 to 0.2 W/m^2]), however, the presence of light absorbing aerosols poses a partially offsetting warming effect [1]. Absorbing aerosols are mainly produced by combustion processes and dust uplifts [2]. They affect the Earth's climate directly by absorbing the solar radiation or indirectly by altering the cloud and surface properties [3]–[6]. Quantifying the aerosol absorption is an important task to determine the aerosol radiative forcing and its uncertainties [7]–[9].

Quantitative aerosol absorptive properties, e.g., the aerosol single scattering albedo (SSA), which presents the ratio of scattering efficiency to total extinction efficiency, and the absorbing aerosol optical depth (AAOD), which describes the fraction of total columnar extinction due to particle absorption, are usually provided by the ground-based aerosol robotic network (AERONET) [10] [11], but the global distribution of AERONET stations is sparse and unbalanced. Observations from space have the advantage to have a better spatial coverage, but only few satellite sensors are capable to retrieve aerosol absorptive properties. Currently, satellite retrievals of AAOD and/or SSA use either multiangle measurements, optionally combined with polarimetry [2], [12], e.g., the multi-angle imaging spectroradiometer (MISR) onboard Terra (1999-present) [13], [14], and the third polarization and directionality of the earth's reflectances mission (POLDER-3) onboard PARASOL (2005–2013) [15]– [17].

Another parameter holding the information of the qualitative aerosol absorption is the ultra-violet aerosol index (UVAI) [18]. Compared with the AAOD and SSA retrieved by MISR and POLDER-3, UVAI is easier to derive without multiangular and polarimetric techniques. Moreover, its calculation does not rely on *a priori* assumptions of aerosol properties (e.g., the size distribution function and the complex refractive index). By contrast, the aerosol properties retrieved by MISR are based on look-up tables (LUTs) produced for 74 predefined aerosol mixtures [19]. The POLDER retrievals, either based on LUTs [16], [20], [21] or the recently developed generalized retrieval of aerosol and surface properties (GRASP) algorithm that online calculates radiative transfer for multiple pixels simultaneously [22], [23], also require predefined information on aerosol properties. On the other hand, the calculation of UVAI does not need to make such assumptions, because UVAI is defined as the change of the spectral contrast at two UV channels ($\lambda_1 < \lambda_2$) between a measured and a calculated radiance due to the presence of absorbing aerosols [18], [24]:

$$\mathrm{UVAI} = -100\left[\log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{\mathrm{obs}} - \log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{\mathrm{Ray}}\right] \quad (1)$$

Manuscript received December 3, 2020; revised March 13, 2021, May 3, 2021, and July 6, 2021; accepted August 22, 2021. Date of publication August 30, 2021; date of current version October 8, 2021. This work was supported by the KNMI Multiannual Strategic Research (MSO). *(Corresponding author: Jiyunting Sun.)*

Jiyunting Sun and Pepijn Veefkind are with the Department of R&D Satellite Observations, Royal Netherlands Meteorological Institute, 3731 De Bilt, The Netherlands, and also with the Department of Geoscience, and Remote Sensing (GRS), Civil Engineering, and Geosciences, Delft University of Technology, 2628 Delft, The Netherlands (e-mail: jiyunting.sun@gmail.com; pepjin.veefkind@knmi.nl).

Peter van Velthoven is with the Department of R&D Weather, and Climate Modeling, Royal Netherlands Meteorological Institute, 3731 De Bilt, The Netherlands (e-mail: peter.van.velthoven@knmi.nl).

Pieternel F. Levelt is with the Atmospheric Chemistry Observations & Modeling Laboratory (ACOM), National Center for Atmospheric Research (NCAR), Boulder, CO 80301 USA, and also with the Department of Geoscience, and Remote Sensing (GRS), Civil Engineering, and Geosciences, Delft University of Technology, 2628 Delft, The Netherlands (e-mail: p.f.levelt@tudelft.nl).

Digital Object Identifier 10.1109/JSTARS.2021.3108669

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

where Ray indicates the radiance calculated by radiative transfer simulations assuming no aerosol present. By assuming $I^{\text{obs}}_{\lambda_2} = I^{\text{Ray}}_{\lambda_2}(a_{s,\lambda_2})$ (a_{s,λ_2} is the surface albedo at the longer wavelength λ_2), the above can be rewritten as

$$\text{UVAI} = -100\log_{10}\frac{I^{\text{obs}}_{\lambda_1}}{I^{\text{Ray}}_{\lambda_1}}. \quad (2)$$

UVAI is a qualitative measure of aerosol absorption. A positive value indicates the presence of absorbing aerosols, whereas nonabsorbing aerosols and clouds yield neutral or negative values [18].

The UVAI has been continuously produced on a global scale by multiple satellites since 1978. It is of interest to quantify the aerosol absorption from such a long-term record. We have attempted to derive AAOD and SSA from UVAI based on radiative transfer simulations [25] and the support vector regression (SVR) [26] for specific cases. Following previous studies, this work generates a database of quantitative aerosol absorptive properties over land from the UVAI between 2006 and 2019 provided by the ozone monitoring instrument (OMI) onboard Aura using a deep neural network (DNN) model. We create a training dataset using both observations and simulations, employ a filter as well as a wrapper method for feature selection, and apply the cross validation for hyperparameter tuning. We provide an error analysis of the DNN predictions and validate the outputs with the AERONET observations. The result of this study leads to an aerosol climatology (2006–2019) of quantitative absorptive properties over land. Section II introduces the training data and the DNN algorithm. Section III presents the assessments of the derived aerosol absorption and the climatology for the period from 2006 to 2019. Section IV summarizes the major findings and potential improvements in the future.

II. Datasets and Methodology

A. Datasets

The datasets used in this study include the OMI/Aura level 2 version 3 aerosol product OMAERUV,[1] the moderate resolution imaging spectroradiometer (MODIS)/Aqua Collection 6 level 3 daily gridded aerosol product MYD08,[2] and the ground-based network AERONET level 1.5 inversion almucantar product.[3] Besides, the modern-era retrospective analysis for research and applications, Version 2 (MERRA-2) aerosol reanalysis is used to calculate the aerosol vertical distributions (MERRA-2 inst3_3d_aer_Nv, 10.5067/LTVB4GPCOTK2) and to provide AAOD and SSA (MERRA-2 tavg1_2d_aer_Nx, 10.5067/KLICLTZ8EM9D) for comparisons with the predicted results. All the data are collected globally from 2006-01-01 to 2019-12-31. All the observational data (i.e., satellite and ground-based data) are regarded to be cloud-free and all the satellite data are regarded to be not affected by large viewing angles and sun-glint effect after preprocessing. The detailed introduction of datasets and preprocessing are presented in Appendix A.

[1][Online]. Available: http://dx.doi.org/10.5067/Aura/OMI/DATA2004
[2][Online]. Available: http://dx.doi.org/10.5067/MODIS/MYD08_M3.006
[3][Online]. Available: https://aeronet.gsfc.nasa.gov

TABLE I
Summary of Datasets and Relevant Variables Used in This Study

OMI	**Solar zenith angle** (SZA), **viewing zenith angle** (VZA), **relative azimuth angle** (RAA), **surface albedo** (a_s, at 388 nm), **surface pressure** (P_s), **UVAI** (calculated by radiance at 354 and 388 nm), AOD (at 500 nm), AAOD (at 500 nm), SSA (at 500 nm), **latitude** (°), **longitude** (°)
MODIS	**AOD** (at 550 nm)
AERONET	AOD (at 550 nm), **AAOD** (at 550 nm), SSA (at 440 and 550 nm), Extinction Ångström Exponent (EÅE, between 440 and 870 nm), Absorption Ångström Exponent (AÅE, between 440 and 870 nm), **time** (day of year)
MERRA-2	AOD (at 550 nm), AAOD (at 550 nm), SSA (at 550 nm), **aerosol layer height** (ALH, derived from aerosol profiles)

All the data are collected globally from 2006-01-01 to 2019-12-31. Parameters used as input features in DNN are bold.

The relevant parameters used in this work are listed in Table I, including satellite-solar geometries (SZA, VZA, RAA), surface properties (a_s and P_s) and UVAI from OMI; AOD from MODIS; and AAOD and SSA from AERONET. Extinction Ångström Exponent (EÅE) and Absorption Ångström Exponent (AÅE) calculated between 440 and 870 nm from AERONET are used for AOD and AAOD spectral conversion according to the power law. Except for UVAI, all aerosol optical properties in this study is at 550 nm by default, unless other wavelengths are explicitly mentioned. The AOD, AAOD, and SSA provided by OMI and MERRA-2 are also included for discussions and comparisons. Note that the OMAERUV product contains two aerosol index variables, and this study uses the term "residue" [27]. For more detailed explanations on this parameter, one can refer to Appendix A.

Since UVAI is highly sensitive to the aerosol vertical distribution [18], [24], [25], [28], whereas the observation on global aerosol vertical distribution is limited in space and time [29]. Instead, we derive the geometric top boundary height of aerosol layers as ALH from the aerosol profiles provided by MERRA-2. For more detailed explanations on this parameter, one can refer to Appendix A.

B. Construction of the Training Dataset

All the variables in Table I are merged into one hybrid dataset. First, the daily OMI and MODIS satellite data are projected onto the MERRA-2 grid ($0.5^\circ \times 0.625^\circ$). Next, the satellite-model data is colocated to the ground-based AERONET observations. According to previous studies [30]–[35], the satellite-model joint data passing a time window of ± 3 h and a spatial distance $\leq$50 km is allocated to an AERONET record after averaging. To ensure the consistency between different measurement techniques (i.e., space-borne versus ground-based), we apply the following filtering criteria, based on the expected errors of the MODIS AOD ($\pm(0.05 + 15\%)$ over land and from $-0.04 + 10\%$ to $+0.02 + 10\%$) over ocean [36], [37]), the OMAERUV AOD (0.1 or 30% [38], [39]) and the AERONET SSA ($\pm$0.03 [11])

1) $|\text{AOD}^M_{550} - \text{AOD}^A_{550}| \leq 0.05 + 15\% \times \text{AOD}^A$ (over land).

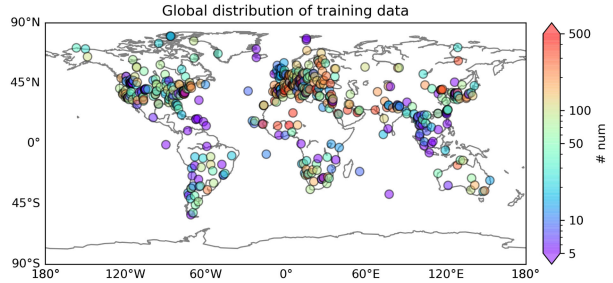

Fig. 1. Global distribution of the training data. The color indicates the number of observations. Note that the temporal coverage of each AERONET site varies.

2) $-0.04-10\% \times \text{AOD}^A \leq \text{AOD}^M_{550}-\text{AOD}^A_{550} \leq 0.02+10\% \times \text{AOD}^A$ (over ocean).
3) $|\text{AOD}^O_{500}-\text{AOD}^A_{500}| \leq 0.1$ or $|\text{AOD}^O_{500}-\text{AOD}^A_{500}|/\text{AOD}^A_{500} \leq 30\%$).
4) $|\text{SSA}^O_{500}-\text{SSA}^A_{500}| \leq 0.03$.

where the upper-script M, A, and O indicate MODIS, AERONET, and OMAERUV. Note that OMAERUV only provides aerosol properties on 354, 388, and 500 nm. Therefore, the corresponding AERONET AOD and AAOD are converted to 550 nm according to the power law

$$\tau_{\lambda_1} = \tau_{\lambda_2}\left(\frac{\lambda_1}{\lambda_2}\right)^{-\alpha} \tag{3}$$

where τ_{λ_1} is the AOD at wavelength λ_1 to be estimated, and τ_{λ_2} is the AOD at wavelength λ_2 that is known; α is the EÅE listed in Table I. The same conversion method is applied to AAOD using corresponding AÅE.

The final hybrid dataset after quality filtering has 48 080 coincidences. Fig. 1 shows its distribution. Most samples are located in North America and Western Europe. On the contrary, few sites measure aerosol absorptive properties over ocean. As a result, this study only focus on deriving the aerosol absorptive properties over land.

C. Feature Selection

1) Feature Selection by Domain Knowledge: In machine learning, features are treated as explanatory variables used to predict the target variable. In our algorithm, the target variable is the quantitative aerosol absorptive parameter AAOD and/or SSA. From our previous study [26], we have proven that compared with SSA, AAOD is better correlated with UVAI, because both AAOD and UVAI are sensitive to aerosol loading and aerosol absorption. Therefore, deriving AAOD from UVAI is expected to give better results than deriving SSA from UVAI.

UVAI is calculated from the satellite measured radiance. It does not only depend on the aerosol properties (e.g. AOD, ALH, aerosol absorption), but also on other parameters, including SZA, VZA, RAA, surface albedo (a_s), and surface pressure (P_s) [18] [24], [25],[28]. The above variables are thus listed as candidate features. Note that the influence of clouds has been eliminated by discarding pixels with cloud fraction larger than 0.3 (see Appendix A) and will not be considered in this work.

Machine learning algorithms solve the numerical relationship between given variables, but they neglect the fact that environmental variables are correlated in space and time [40]. Recently, there is an increasing trend involving the spatial and/or temporal autocorrelation in machine learning applications. This can be done in many ways, for example, Li *et al.* [41] and Li *et al.* [42] directly used the measurements in the past and/or the measurements from surrounding sites. In other studies, geo-statistical methods were used in combination with machine learning algorithms (e.g., [43]–[48]). There is also an increasing trend using the convolutional neural network (CNN) and/or the recurrent neural network (RNN) to present the temporal and spatial correlation (e.g., [49]–[51]), as these techniques can share weight parameters in space and/or time domain. However, applying above techniques in this study is not feasible due to the limitation of training datasets (limited spatial coverage and different time spans, see Fig. 1). Instead, we just directly add geo-coordinates (longitude and latitude) and time information (day of the year) into the feature space as that done in other studies (e.g. [52]–[58]). In total, we selected 11 variables as the candidate features based on our domain knowledge.

2) Feature Selection by Filter and Wrapper Methods: We have chosen 11 features according to our experience and knowledge, but whether those features are favorable to derive AAOD from UVAI needs further investigations. Feature selection is the process of selecting a subset of features that is most relevant to the target variable. This process is important to enhance the model interpretability, computational efficiency, generalization performance, etc. [59].

In our previous study, we only selected three features that have relatively high Spearman's correlation coefficients (R^2_s) with the target variable (the AERONET AAOD) [26]. However, R^2_s only measures the monotonic relationship between features and the target variable. In this work, we apply two independent feature selection methods: 1) the maximum information coefficients (MIC) [60] and 2) the recursive feature elimination (RFE) [59].

MIC is a filter method that was first introduced in [60]. It measures the dependence between two variable X and Y, no matter the relation is linear or nonlinear. It uses a constrained adaptive bin method to apply mutual information (I) on continuous variables. The mutual information measures the mutual dependence between two random variables, which is defined as

$$I(X;Y) = \int p(x,y)\log_2\frac{p(x,y)}{p(x)p(y)}dxdy \tag{4}$$

where $p(x,y)$ is the joint probability density function of X and Y, and $p(x)$ and $p(y)$ are the marginal probability density function of X and Y. Then MIC uses binning to compute the normalized mutual information on continuous variables

$$\text{MIC}(X,Y) = \max_{n_x * n_y < N^z}\frac{I(X;Y)}{\log_2\min(n_x,n_y)} \tag{5}$$

where n_x and n_y are the bin numbers of X and Y, respectively. Their product is supposed to be smaller than a number N^z, where N is the size of the data and z usually takes 0.6 [60]. Last, MIC is the maximum of the normalized mutual information values calculated by different combinations of n_x and n_y (MIC ranges between 0 and 1).

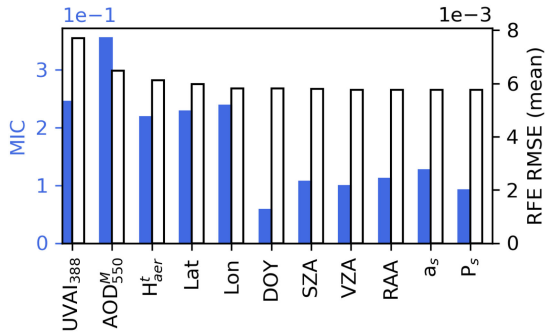

Fig. 2. Feature selection: the (MIC), blue bars and the (RFE), white bars. The higher MIC values, the stronger correlation between a feature and the target variable (AAOD). RFE presents how the model performance (in format of averaged RMSE from 100 RF experiments) varies with an eliminated feature. The higher the RMSE (the worse the model performance), the more important the corresponding feature.

MIC is a filter method, which uses statistical measures to independently evaluate the correlation between the target variable and the input features, and to filter out the least relevant features. However, filter methods only look at one individual feature at a time, thus ignoring relationship between features. Wrapper methods like RFE can also detect the interactions between features [61]. RFE recursively removes the least relevant feature by a certain metric and aims to find the feature combination that leads to the optimal model performance. The feature ranking metrics of RFE is given by an external estimator that assigns weights to each feature, such as the coefficients of the linear regression or the support vector machines [59], or the feature importance of the random forests (RF) [62]. In this work, we use the RF-based RFE, as RF can deal with nonlinear problems and it does not have many hyperparameters to tune. Since RF predictions will not always be the same, we set up 100 RFE experiments and use the average value to select the features of interest.

The MIC and REF evaluation on the 11 candidate features are presented in Fig. 2. MIC (blue bars) indicates the UVAI, AOD, ALH are the most relevant features to AAOD, which is in agreement with our previous study [26]. The geo-coordinates (Lat and Lon) are more important compared to the rest parameters. The average of 100 RFE experiments (white bars) shows how the model performance changes (in terms of the averaged root mean square error, RMSE) with an eliminated feature. It is clear that the most important feature is UVAI, followed by AOD and ALH. The remaining features do neither significantly improve nor hurt the model performance. As there is no solid reason to exclude them, we decide to keep all 11 features.

The basic statistics of the selected features are shown in Fig. 3. But to show the impact of the spatial and temporal information (Lat, Lon, DOY) on predictions, we also build a feature space without this information to allow a comparison. Consequently, we have the following two feature spaces.

1) Feature space with 11 features (F11): UVAI, AOD, ALH, Lat, Lon, DOY, SZA, VZA, RAA, a_s, P_s.
2) Feature space with 8 features (F8): UVAI, AOD, ALH, SZA, VZA, RAA, a_s, P_s.

D. Deep Neural Network (DNN)

Machine learning algorithms learn the underlying behavior of a system from a set of training data [63]. They have been widely applied in geosciences and remote sensing [63], [64]. Recently, driven by the increasing size of geo-data, algorithms such as deep learning becomes more and more popular [65], [66]. Deep learning is characterized by neural networks with no less than two hidden layers [67].

We use a feed-forward DNN with multiple hidden layers (see Fig. 4). A DNN consists of an input layer, n intermediate hidden layers and an output layer. The input layer contains input features, and the output layer gives the predicted variable(s). A hidden layer consists of m neurons. The jth neuron in lth hidden layer (a_l^j) is calculated by neurons in the previous layer

$$z_l^j = \sum_{i=1}^{n} w_l^{ji} a_{l-1}^i + b_l^j \tag{6}$$

where w_l^{ji} is the weight of the ith neuron in $(l-1)$th layer (a_{l-1}^j) given to the a_l^j; b_l^j is a bias term. Then, the computed z_l^j is fed into an activation function σ

$$a_l^j = \sigma(z_l^j). \tag{7}$$

The activation function is used to add the nonlinear properties to a neural network. In this work, the rectified linear unit (ReLU) is used as the activation function

$$\sigma(x) = \max(0, x). \tag{8}$$

The DNN is to find the optimal weight matrix W containing all w_l^{ji} that can minimized the loss function (L). Here, we use the RMSE as the error metric

$$L = \left[\frac{1}{k} \sum_{i=1}^{k} (\hat{y}_i - y_i) \right]^{1/2} \tag{9}$$

where k is the number of samples; $\hat{y}_i$ and y_i are the DNN-predicted and the true values. In this work, the optimization of loss function is realized by the Adam solver [68] with a fixed learning rate of 10^{-4}. The Adam solver is more computationally efficient than classical optimization algorithms and performs well for large datasets [69].

The structure of DNN (the number of hidden layer n and the number of neurons in each layer m) highly varies with applications. Therefore, we applied an exhaustive grid search over specified hyperparameter values. The process is based on tenfold cross validations: the dataset is randomly split into 10 sets. For each fold, 9 out of 10 sets are used to train the DNN model and the remaining set is used for validation.

Tables II and III present the training results and the validation of the model performance for different layer and neuron configurations. Compared to the model performance of DNN-F8, the DNN-F11 predictions show a better consistency with the true values (R^2 =0.9). The RMSE and the mean absolute error (MAE) of the DNN-11 predictions are also smaller than that of DNN-8. The comparison reveals that the additional spatial and temporal information indeed improves the DNN model

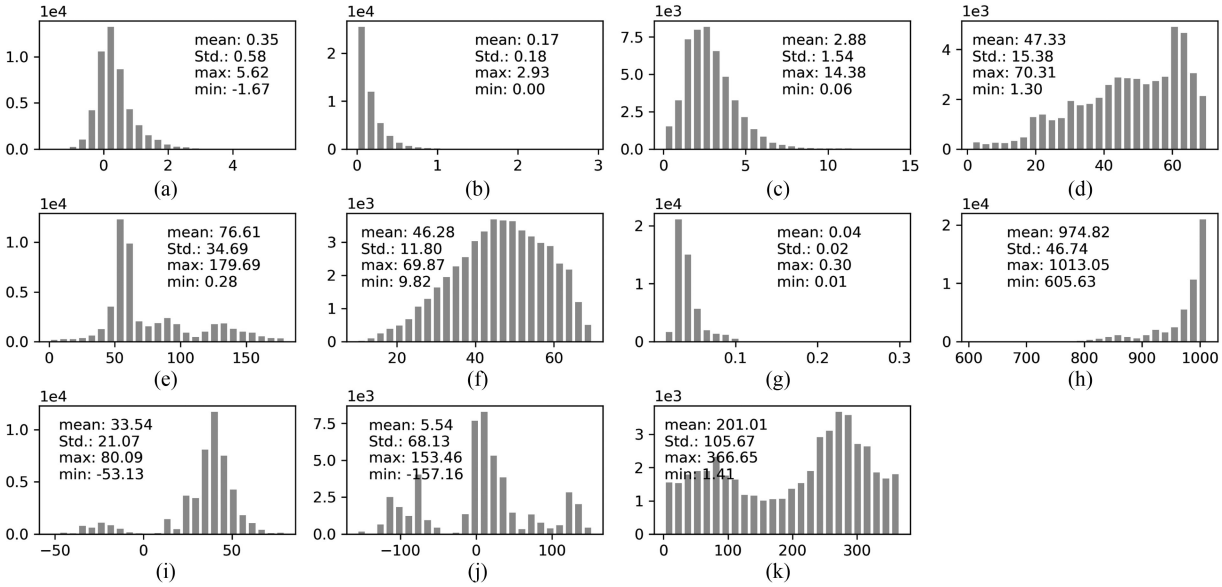

Fig. 3. Histogram and basic statistics of the selected features.

TABLE II
MODEL PERFORMANCE OF DNN WITH 11 FEATURES UNDER DIFFERENT DNN CONFIGURATIONS IN TERMS OF THE LINEAR FITTING SLOPE (k), THE INTERCEPT (b), THE CORRELATION COEFFICIENT (R^2), THE RMSE, AND THE MAE

		Training					Validation				
Layer	Neuron	k	b	R^2	RMSE	MAE	k	b	R^2	RMSE	MAE
1	64	0.78	0	0.88	5.34E-03	3.42E-03	0.78	0	0.88	5.47E-03	3.48E-03
	128	0.79	0	0.89	5.19E-03	3.34E-03	0.79	0	0.88	5.46E-03	3.44E-03
	256	0.79	0	0.9	5.16E-03	3.29E-03	0.77	0	0.88	5.34E-03	3.40E-03
2	64	0.85	0	0.9	4.92E-03	3.19E-03	0.82	0	0.89	5.21E-03	3.36E-03
	128	0.84	0	0.91	4.71E-03	3.05E-03	0.82	0	0.89	5.18E-03	3.31E-03
	256	0.85	0	0.92	4.51E-03	2.92E-03	0.83	0	0.89	5.06E-03	3.23E-03
3	64	0.83	0	0.91	4.71E-03	3.06E-03	0.82	0	0.89	5.25E-03	3.33E-03
	128	0.86	0	0.93	4.19E-03	2.69E-03	0.82	0	0.9	4.96E-03	3.18E-03
	256	0.85	0	0.93	4.10E-03	2.61E-03	0.82	0	0.9	4.96E-03	3.13E-03
4	64	0.83	0	0.92	4.53E-03	2.94E-03	0.81	0	0.89	5.10E-03	3.28E-03
	128	0.84	0	0.94	3.92E-03	2.51E-03	0.81	0	0.9	4.86E-03	3.10E-03
	256	0.84	0	0.93	4.11E-03	2.56E-03	0.8	0	0.9	4.94E-03	3.10E-03
5	64	0.84	0	0.92	4.39E-03	2.84E-03	0.81	0	0.89	5.03E-03	3.23E-03
	128	0.86	0	0.94	3.99E-03	2.56E-03	0.84	0	0.9	4.81E-03	3.09E-03
	256	0.86	0	0.94	3.79E-03	2.36E-03	0.82	0	0.91	4.76E-03	3.03E-03
6	64	0.82	0	0.92	4.52E-03	2.94E-03	0.79	0	0.9	4.98E-03	3.23E-03
	128	0.86	0	0.94	3.78E-03	2.37E-03	0.82	0	0.9	4.82E-03	3.05E-03
	256	0.83	0	0.94	3.87E-03	2.35E-03	0.79	0	0.9	4.88E-03	3.03E-03

performance. However, the improvement is not significant as that in other studies (e.g., [42]), because we only provide coordinate and time information in a straightforward way, instead of giving the explicit spatial and temporal autocorrelation calculated by geo-statistics methods (e.g., deterministic interpolation, Kriging, or Gaussian process regression), or sharing weights in space and time domain by CNN or RNN.

The optimal model of DNN-F11 is constructed by 3 hidden layers with 64 neurons in each layer. The final trained model has a prediction accuracy of 0.0045 [see Fig. 5(a)]. This model performance is significantly better than the SVR model used in the previous work ([26], RMSE = 0.01). Considering the representative uncertainty of AERONET AAOD is at the level between 0.004 and 0.006 (calculated by the expected errors of AERONET AOD and SSA via the error propagation equation, see Fig. 23 in Appendix A), the model performance is encouraging. The optimal model of DNN-F8 is constructed by 2 hidden layers with 128 neurons in each layer. Without the information on coordinate and time, the final trained model has a prediction accuracy of 0.0056, slightly higher than that of DNN-F11. Consequently, we only use the DNN-F11 in the following applications.

TABLE III
MODEL PERFORMANCE OF DNN WITH 8 FEATURES UNDER DIFFERENT DNN CONFIGURATIONS IN TERMS OF THE LINEAR FITTING SLOPE (k), THE INTERCEPT (b), THE CORRELATION COEFFICIENT (R^2), THE RMSE, AND THE MAE

		Training					Validation				
Layer	Neuron	k	b	R^2	RMSE	MAE	k	b	R^2	RMSE	MAE
1	64	0.76	0	0.86	5.84E-03	3.65E-03	0.75	0	0.85	5.88E-03	3.68E-03
	128	0.76	0	0.86	5.70E-03	3.60E-03	0.75	0	0.86	5.84E-03	3.65E-03
	256	0.76	0	0.87	5.54E-03	3.55E-03	0.75	0	0.86	5.73E-03	3.62E-03
2	64	0.75	0	0.88	5.51E-03	3.52E-03	0.74	0	0.86	5.76E-03	3.63E-03
	128	0.8	0	0.88	5.35E-03	3.42E-03	0.79	0	0.86	5.77E-03	3.59E-03
	256	0.79	0	0.89	5.15E-03	3.28E-03	0.77	0	0.87	5.63E-03	3.52E-03
3	64	0.78	0	0.89	5.17E-03	3.32E-03	0.76	0	0.87	5.63E-03	3.56E-03
	128	0.83	0	0.91	4.82E-03	3.01E-03	0.8	0	0.87	5.61E-03	3.44E-03
	256	0.84	0	0.91	4.64E-03	2.94E-03	0.81	0	0.87	5.45E-03	3.40E-03
4	64	0.77	0	0.9	4.94E-03	3.16E-03	0.74	0	0.87	5.51E-03	3.48E-03
	128	0.82	0	0.92	4.49E-03	2.84E-03	0.77	0	0.87	5.41E-03	3.37E-03
	256	0.8	0	0.91	4.71E-03	2.91E-03	0.76	0	0.87	5.48E-03	3.38E-03
5	64	0.82	0	0.9	4.92E-03	3.13E-03	0.79	0	0.87	5.53E-03	3.45E-03
	128	0.82	0	0.92	4.49E-03	2.85E-03	0.78	0	0.88	5.33E-03	3.36E-03
	256	0.81	0	0.92	4.43E-03	2.72E-03	0.77	0	0.88	5.30E-03	3.28E-03
6	64	0.77	0	0.9	4.92E-03	3.14E-03	0.73	0	0.87	5.58E-03	3.47E-03
	128	0.81	0	0.92	4.39E-03	2.76E-03	0.77	0	0.88	5.34E-03	3.33E-03
	256	0.81	0	0.92	4.42E-03	2.67E-03	0.78	0	0.88	5.31E-03	3.26E-03

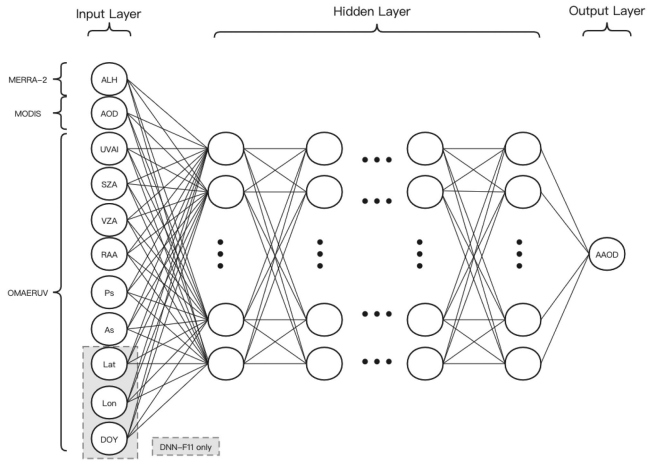

Fig. 4. Conceptual structure of the DNN in this study.

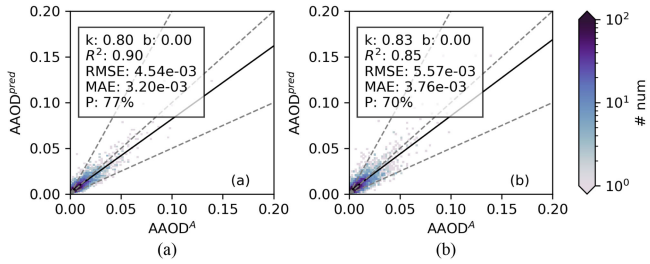

Fig. 5. Performance of the optimal model selected by grid search. The linear fitting slope (k), intercept (b), the correlation coefficient (R^2), the RMSE and the MAE between the observed and predicted values, and the percentage of AAOD difference within AERONET AAOD estimated error in the total samples (P) are provided: (a) DNN with 11 features (DNN-F11); (b) DNN with 8 features (DNN-F8). The dashed lines are 1:1 and 50% difference lines.

III. RESULTS AND DISCUSSION

A. Assessments of the DNN-Predicted Aerosol Absorption

The DNN-F11 model is applied to predict the aerosol absorption from 2006 to 2019. The AAOD estimated by the DNN and further derived SSA (SSA = 1 - AAOD/AOD) are validated by the colocated AERONET records and compared with the aerosol absorption provided by OMAERUV and MERRA-2. To ensure the consistency between different products, we apply the same quality filtering described in Section II-B before analysis. This filtering is also applied to MERRA-2 AOD and SSA. There are in total 39 504 coincidences for validation.

1) Influence of Input Features on the DNN Predictions: First, we investigate the relationship between the errors of the DNN predictions and input features. According to Fig. 6, the difference between the DNN-predicted AAOD and the AERONET-retrieved AAOD shows little dependence on the input features. However, it is slightly related to the AOD difference between MODIS and AERONET [see Fig. 6 (l)]. The positive bias of MODIS AOD leads to overestimation of the DNN-predicted AAOD. On the other hand, Fig. 7(b) reflects that low AOD may cause large biases in the DNN-derived SSA, because MODIS observation become less sensitive to low aerosol loading [37]. This indicates that a quality filtering on AOD is necessary to ensure the quality of the DNN predictions. As a result, based on the sensitivity study of the DNN-derived SSA accuracy [see Fig. 24 in Appendix B], a threshold of 0.1 on the MODIS AOD at 550 nm is applied to the DNN predictions. The same threshold is also applied to the OMAERUV AOD (converted to 550 nm) and the MERRA-2 AOD for direct comparisons. There are total 21 600 samples after AOD-screening. The following discussions are based on the filtered data.

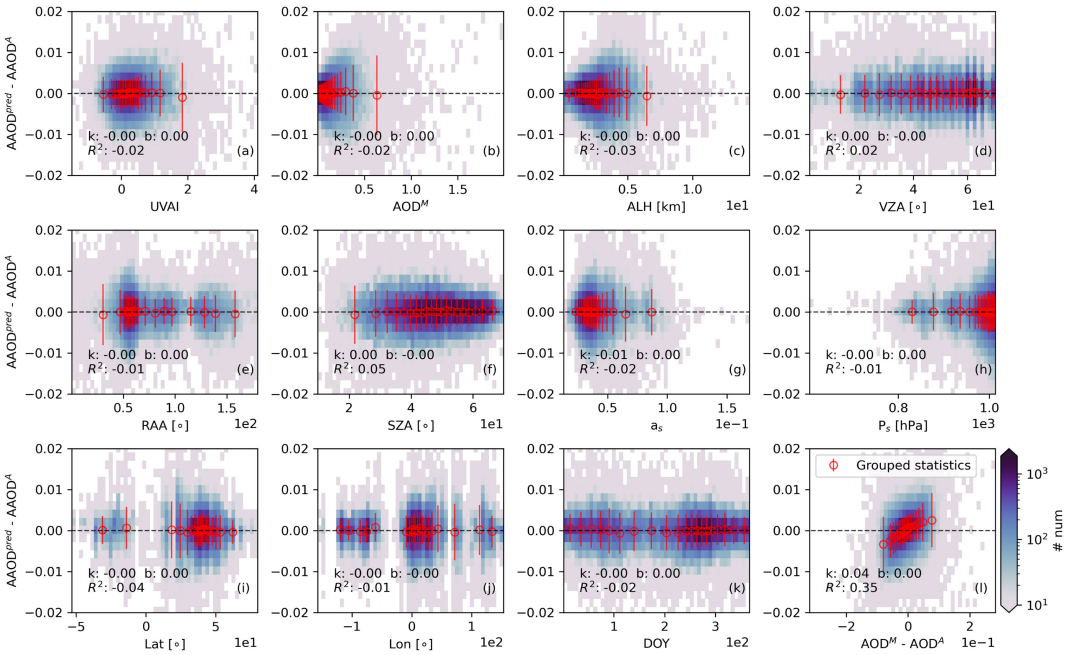

Fig. 6. AAOD difference between DNN and AERONET against input features. The gray dashed line indicates the neutral. The color of density plots indicates the number of samples. The data is further grouped into 20 sets, and the red circles and error bars are the mean and standard deviation of each group.

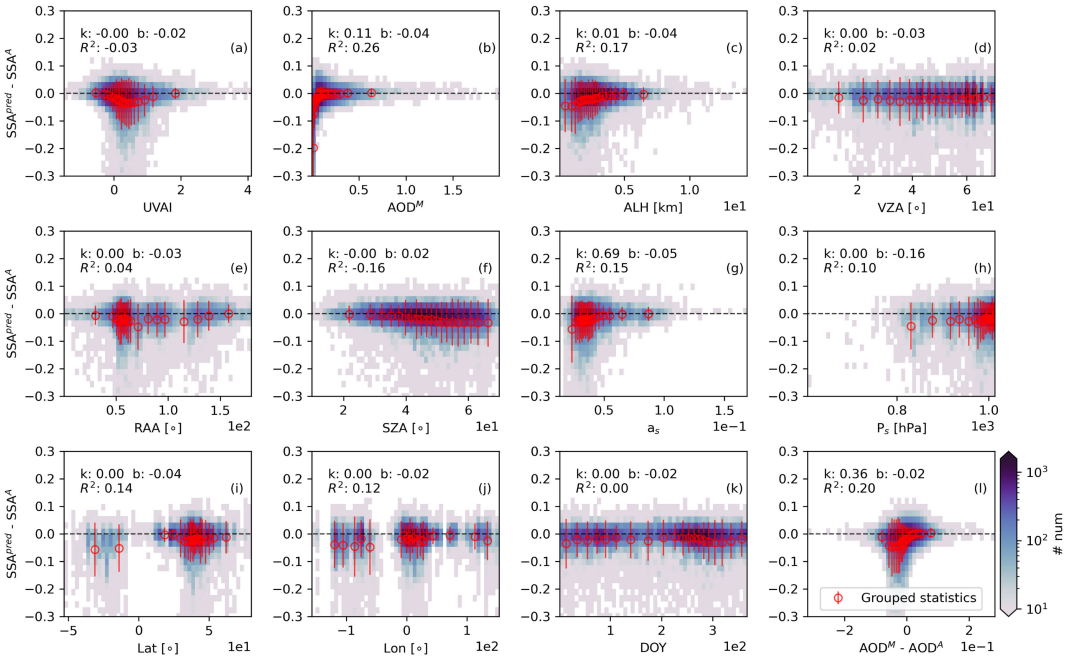

Fig. 7. SSA difference between DNN and AERONET against input features. The gray dashed line indicates the neutral. The color in density plots indicates the number of samples. The data is further grouped into 20 sets, and the red circles and error bars are the mean and standard deviation of each group.

2) Validating the DNN Predictions With AERONET: According to Fig. 8(a), the filtered DNN-predicted AAOD shows a high correlation with the AERONET retrievals (R^2=0.89). The linear fitting shows that the DNN predictions have a slight negative bias, which is mainly caused by the underestimation for cases with high aerosol absorption. Both RMSE and MAE are at level around 0.005, and about 83% samples falling within the AERONET AAOD estimated error. The corresponding SSA is also positively biased [see Fig. 9(a)]. Over 80% of the samples are within the AERONET SSA typical uncertainty of ±0.03. Compared with the SVR predictions in our previous study with only 66% data is inside the ±0.03 confidence interval [26], the DNN-predictions show a significant improvement. Despite of the larger bias for high absorption cases, the averaged error of DNN-derived SSA is at level around 0.02.

Aerosol absorptive properties of OMAERUV (at 500 nm) and MERRA-2 (at 550 nm) are also compared with the AERONET retrievals. Although both the OMAERUV AOD and SSA are

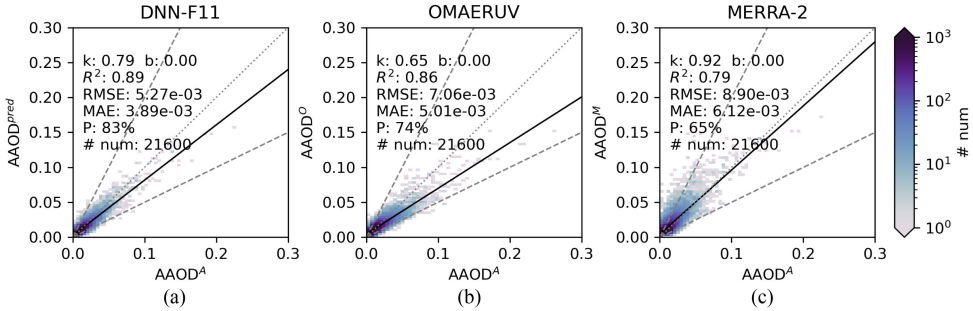

Fig. 8. AERONET AAOD against the AAOD of (a) DNN-F11 predictions, (b) OMAERUV and (c) MERRA-2 after the AOD-screening. The black lines are the linear fittings. The dotted and dashed lines are 1:1 and 50% difference lines, respectively. P is the percentage of the difference between the DNN-predicted AAOD and AERONET AAOD smaller than the expected error of the AERONET AAOD calculated by (23) in the total samples.

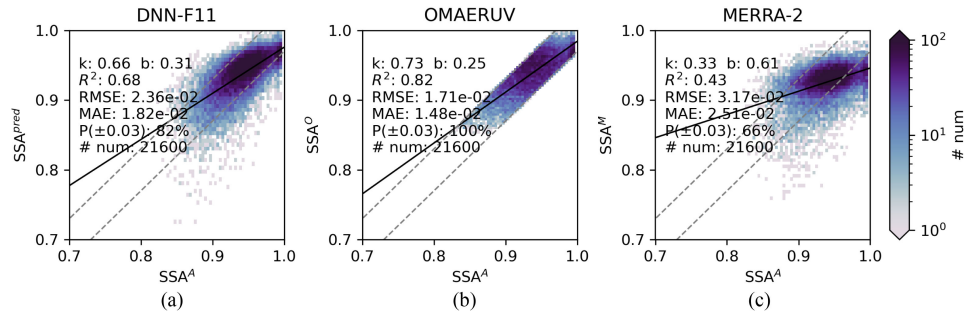

Fig. 9. AERONET SSA against the SSA of (a) DNN-F11 predictions, (b) OMAERUV and (c) MERRA-2 after the AOD-screening. The black lines are linear fittings. The gray dashed lines are the ± 0.03 uncertainty of AERONET SSA. P is the percentage of the difference between the DNN-predicted SSA and AERONET SSA smaller than the expected error of the AERONET SSA (± 0.03) in the total samples.

TABLE IV
AEROSOL CLASSIFICATION BASED ON THE EXTINCTION ÅNGSTRÖM EXPONENT (EÅE), THE ABSORPTION ÅNGSTRÖM EXPONENT (AÅE), AND THE SSA AT 440 NM (ADAPTED FROM [70] AND [71])

Aerosol types	$E\mathring{A}E_{440-870}$	$A\mathring{A}E_{440-870}$	SSA_{440}	Comment
Smoke	$E\mathring{A}E \geq 1.5$	$A\mathring{A}E \geq 0$	$SSA \leq 0.95$	Small absorbing aerosols
Dust	$E\mathring{A}E \leq 0.5$	$A\mathring{A}E \geq 1$	$SSA \leq 0.95$	Large absorbing aerosols
Mixed	$0.5 \leq E\mathring{A}E \leq 1.5$	$A\mathring{A}E \geq 1$	$SSA \leq 0.95$	Mixture of small and large absorbing aerosols
Other		All other conditions		Non-absorbing aerosols

quality assured by AERONET, its AAOD is still considerably lower than the corresponding AERONET retrievals, resulting in a RMSE larger than the expected magnitude [see Fig. 8(b)]. Similar as the DNN predictions, the gap is mainly due to the underestimation of the high aerosol absorption cases. The MERRA-2 AAOD tends to be overestimated for lower values and underestimated for higher values, leading to the highest error level (both RMSE and MAE) among the three data sets [see Fig. 8(c)]. In most cases, the MERRA-2 SSA is also higher than that of AERONET, with only 66% of samples within the ± 0.03 confidence range [see Fig. 9(c)].

3) Validating the DNN Predictions by Aerosol Types: We further validate the DNN predictions according to aerosol types. We categorize the aerosols into four types (i.e., smoke, dust, smoke-and-dust-mixed, and other nonabsorbing aerosols) by the EÅE between 440 and 870 nm, the AÅE between 440 and 870 nm, and the SSA at 440 nm reported in the AERONET inversion product (see Table IV). This classification method is adapted from [70], [71].

The comparisons of the DNN, OMAERUV, and MERRA-2 against the AERONET AAOD and SSA for the four aerosol types are presented in Fig. 10 and Table V. All three datasets show their best consistency with the AERONET retrievals for smoke aerosols (R^2 >0.90). R^2 slightly decreases for larger absorbing aerosols (mixed and dust aerosols), but most of the samples is still within the $\pm 50\%$ difference range (dashed lines). On the other hand, all three datasets are not well correlated with AERONET for nonabsorbing aerosols (R^2 <0.8), and tend to overestimate the aerosol absorption.

Compared to OMAERUV and MERRA-2, the DNN predictions have higher R^2 and lower error in general. The agreement between the DNN predictions and the AERONET retrievals varies with aerosol types. The DNN is best at predicting aerosol absorption for smaller absorbing aerosols. This is because the DNN-predicted AAOD is most sensitive to the input MODIS AOD and OMAERUV UVAI. The uncertainty of the MODIS AOD in the visible band is higher for dust aerosols due to the stronger surface reflectance [36], [37], [72], [73]. Consequently,

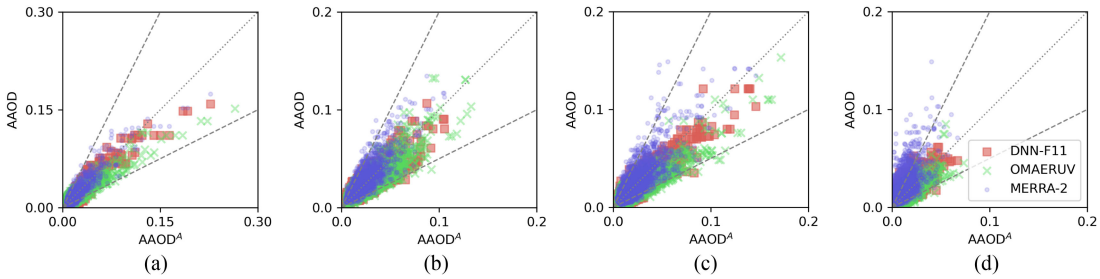

Fig. 10. AERONET AAOD against that of DNN-F11 (red square), OMAERUV (green cross) and MERRA-2 (blue circle) for four aerosol types based on Table IV: (a) smoke, (b) dust, (c) mixed and (d) other aerosols. The dashed lines are 1:1 and 50% difference lines.

TABLE V
AAOD VALIDATION BY AEROSOL TYPES

Type	Data	RMSE	MAE	k	b	R^2	P[%]	N	ME$_{AAOD}$	ME$_{AOD}$	ME$_{SSA}$
Smoke	DNN-F11	5.46E-03	3.94E-03	0.9	0	0.93	79	3533	-0.003	-0.02	0.00
	MERRA-2	6.09E-03	4.17E-03	0.96	0	0.91	77	3533	-0.001	-0.01	0.00
	OMAERUV	7.51E-03	4.74E-03	0.65	0	0.91	72	3533	-0.002	0.00	0.01
Dust	DNN-F11	6.50E-03	4.75E-03	0.73	0.01	0.86	89	2410	0.00	0.00	0.00
	MERRA-2	1.01E-02	7.42E-03	1.02	0	0.83	74	2410	0.005	0.01	-0.01
	OMAERUV	8.92E-03	6.66E-03	0.73	0	0.85	81	2410	-0.003	-0.01	0.00
Mixed	DNN-F11	5.21E-03	3.86E-03	0.82	0	0.9	82	7322	-0.002	-0.01	0.00
	MERRA-2	7.92E-03	5.21E-03	1.02	0	0.82	73	7322	0.001	-0.01	0.00
	OMAERUV	7.77E-03	5.47E-03	0.6	0	0.83	69	7322	-0.003	-0.01	0.01
Other	DNN-F11	4.83E-03	3.64E-03	0.73	0.00	0.73	84	8335	0.002	0.00	-0.01
	MERRA-2	1.02E-02	7.37E-03	1.04	0.01	0.63	50	8335	0.007	0.00	-0.03
	OMAERUV	5.41E-03	4.24E-03	0.67	0.00	0.7	76	8335	0.002	0.01	-0.01

The columns are statistics between AAOD of each data set and AERONET: The RMSE, the MAE, the linear fitting slope (k), intercept (b) and correlation coefficient (R^2), the percentage of difference within the expected AERONET error level in the total samples (P), the number of samples (N), and the mean error of AAOD, AOD and SSA (ME_{AAOD}, ME_{AOD}, and ME_{SSA}) of each aerosol type.

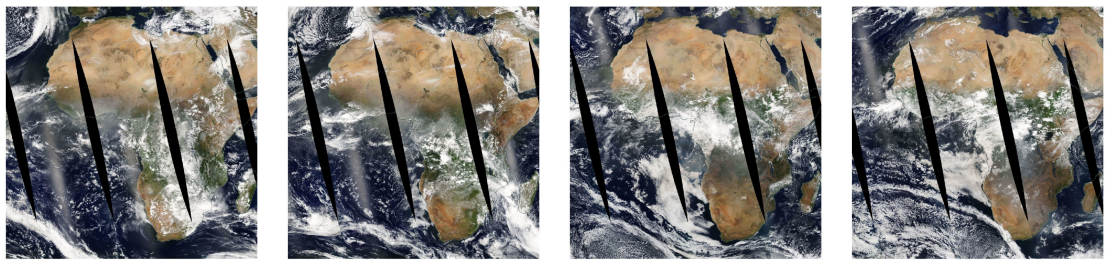

Fig. 11. MODIS/Aqua true color maps of the selected cases: (a) 2019-02-14; (b) 2019-02-18; (c) 2019-07-10; (d) 2019-08-07. Source: https://worldview.earthdata.nasa.gov/.

the DNN predictions for dust aerosols are less consistent with the AERONET retrievals than smoke aerosols. The gap between DNN-predictions and AERONET is largest for nonabsorbing aerosols. It is because that the OMAERUV UVAI is more sensitive to absorbing aerosols than scattering aerosols by its definition [18], [24].

The OMAERUV AAOD is overall lower than the AERONET retrievals, with k <0.8 for all aerosol types. This is caused by the higher OMAERUV SSA compared to AERONET for the majority of the data. This agrees with findings in [34]. According to Jethva *et al.*, the potential reasons behind the SSA difference to AERONET could come from the clouds contamination, the assumptions on ALH, the high surface albedo (in desert and arid areas), and the assumed aerosol microphysical properties. On the contrary, the MERRA-2 AAOD shows slightly positive biases (except for smoke), particularly for nonabsorbing aerosols, reflecting that the MERRA-2 aerosol models may overestimate the aerosol absorption.

4) Assessments of the DNN-Predicted Aerosol Absorption by Cases: Here, we selected four cases to further investigate the performance of the DNN predictions, covering dust storms and biomass burning events in Africa. Figs. 11 and 12 present the MODIS/Aqua true color images and the OMAERUV UVAI for the selected cases. Case 2019-02-14 and 2019-02-18 show the biomass burning events in the central Africa. Case 2019-07-10 and 2019-08-07 mainly present the dust storms in the

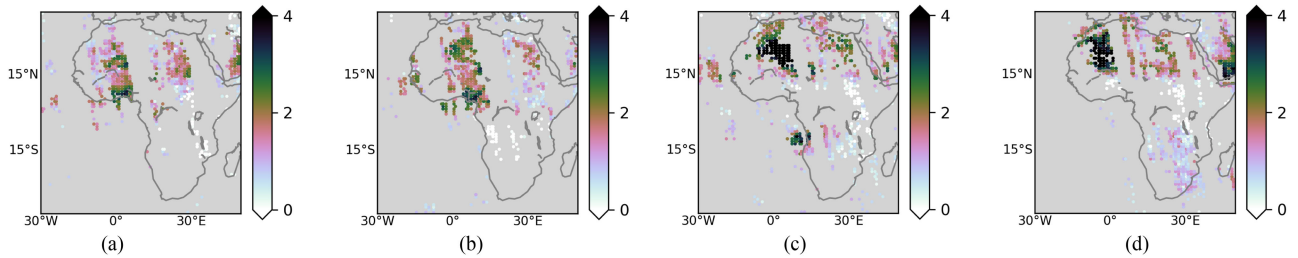

Fig. 12. OMAERUV UVAI of the selected cases: (a) 2019-02-14; (b) 2019-02-18; (c) 2019-07-10; (d) 2019-08-07.

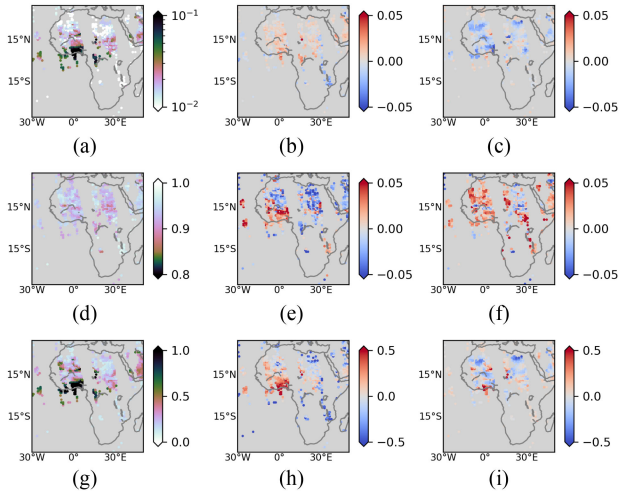

Fig. 13. Case study of 2019-02-14. The upper-script O and M indicate OMAERUV and MERRA-2 respectively. First row: DNN-predicted AAOD and its difference with the OMAERUV and MERRA-2 AAOD; second row: DNN-derived SSA and its difference with the OMAERUV and MERRA-2 SSA; third row: input UVAI, the MODIS AOD difference with OMAERUV, and MODIS AOD difference with MERRA-2.

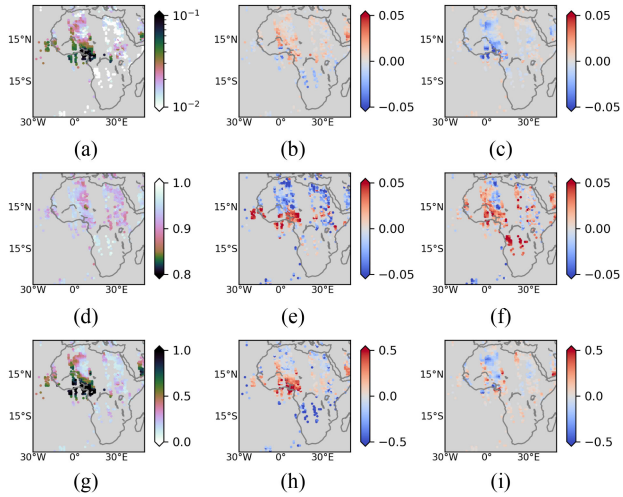

Fig. 14. Case study of 2019-02-18. The upper-script O and M indicate OMAERUV and MERRA-2 respectively. First row: DNN-predicted AAOD and its difference with the OMAERUV and MERRA-2 AAOD; second row: DNN-derived SSA and its difference with the OMAERUV and MERRA-2 SSA; third row: input UVAI, the MODIS AOD difference with OMAERUV, and MODIS AOD difference with MERRA-2.

northern Africa with some biomass burning events in the southern Africa.

Figs. 13–16 show the DNN-predicted AAOD, SSA, and the input MODIS AOD (left column), and their difference with the corresponding parameters of OMAERUV (middle column) and MERRA-2 (right column). In general, the DNN-derived AAOD can reflect the distribution of absorbing aerosols (cloud-free parts) as that shown in Figs. 11 and 12, whereas it is difficult to tell the dust or smoke plumes from the derived-SSA alone (same for SSA of OMAERUV and MERRA-2). The distribution of the DNN-derived AAOD is dominant by the input MODIS AOD, whereas the effect of UVAI is weaker. For example, in case 2019-07-10 and 2019-08-07, both AOD and the DNN-derived AAOD show high values in biomass burning regions, whereas the corresponding UVAI is very low.

The difference plots of AAOD and SSA (see Figs. 13–16 (b, c, e, f)) show that the DNN-predicted aerosol absorption is generally higher than that of OMAERUV, whereas it is lower than that of MERRA-2. This finding is consistent with that found in previous sections. The difference in aerosol absorption is mainly associated with the difference in AOD. For example, in case 2019-02-18, the MODIS AOD of the smoke plume is

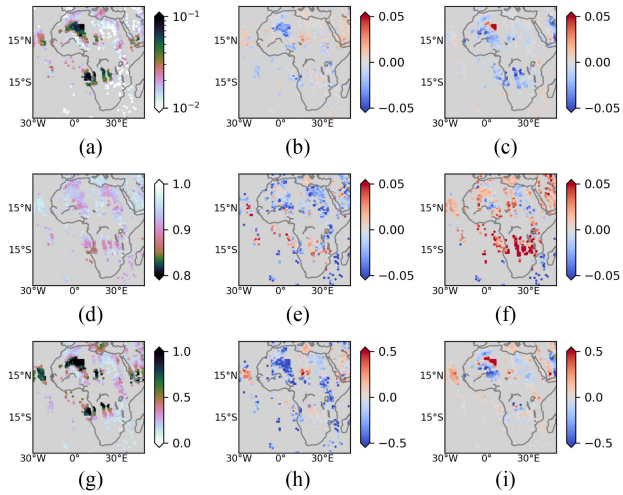

Fig. 15. Case study of 2019-07-10. The upper-script O and M indicate OMAERUV and MERRA-2 respectively. First row: DNN-predicted AAOD and its difference with the OMAERUV and MERRA-2 AAOD; second row: DNN-derived SSA and its difference with the OMAERUV and MERRA-2 SSA; third row: input UVAI, the MODIS AOD difference with OMAERUV, and MODIS AOD difference with MERRA-2.

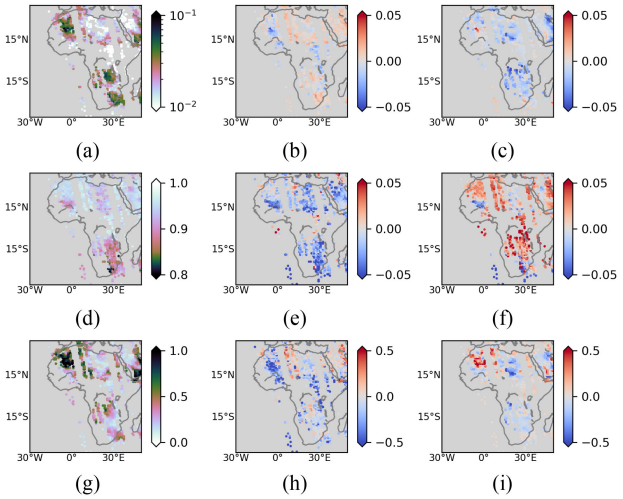

Fig. 16. Case study of 2019-08-07. The upper-script O and M indicate OMAERUV and MERRA-2 respectively. First row: DNN-predicted AAOD and its difference with the OMAERUV and MERRA-2 AAOD; second row: DNN-derived SSA and its difference with the OMAERUV and MERRA-2 SSA; third row: input UVAI, the MODIS AOD difference with OMAERUV, and MODIS AOD difference with MERRA-2.

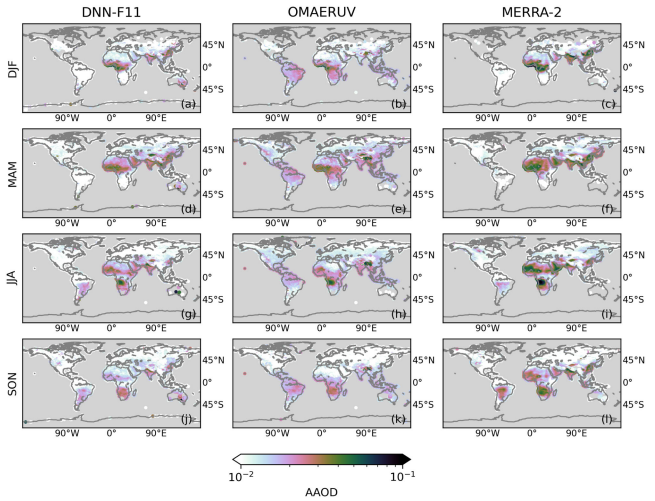

Fig. 17. Seasonal AAOD of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over land.

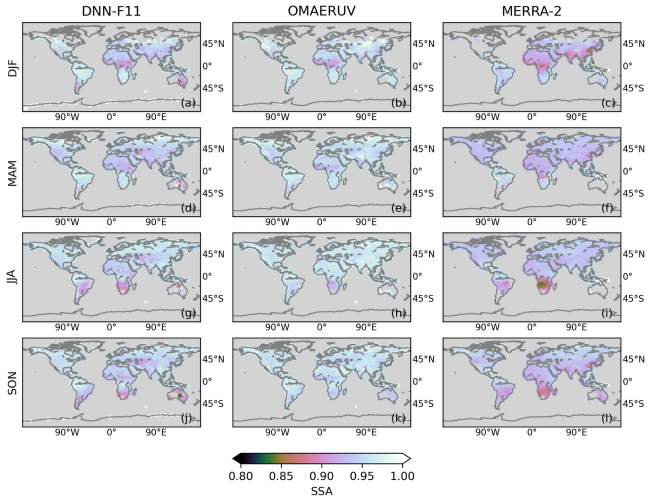

Fig. 18. Seasonal SSA of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over land.

significantly higher than that of OMAERUV [see Fig. 14(h)], and the corresponding AAOD and SSA difference shows the same pattern [see Fig. 14 (b, e)].

B. Aerosol Absorption Climatology

Here, we discuss the aerosol absorption climatology over land derived by DNN for the period from 2006 to 2019. Although it is applicable for aerosols over ocean, the DNN model is mainly trained for aerosols over land, as the most of AERONET sites measure aerosol absorptive properties of continental aerosols (see Fig. 1). The aerosol properties over ocean may have a different distribution from that of continental aerosols, which could bias the predictions. Therefore, we only discuss the aerosol absorption over land in this section, and provide a tentative analysis on the predicted aerosol absorption over ocean in Appendix C.

1) Global Aerosol Absorption Climatology: Figs. 17 and 18 present the seasonal aerosol absorption of DNN, OMAERUV, and MERRA-2 over land averaged between 2006 and 2019. All three datasets show the major absorbing aerosol sources in each season, for instance, the biomass burning events in Central Africa in DJF and MAM, and dust storms from Sahara in MAM and JJA, etc. Compared with MERRA-2, both the DNN and OMAERUV AAOD have lower magnitudes, especially over desert regions. It is because both MODIS and OMI have difficulty in retrieving AOD over bright surfaces in the visible band. Besides, in OMAERUV, dust particles are assumed to be spherical, which is not a proper assumption to calculated the dust optical properties [74]. On the contrary, MERRA-2 assimilates the MISR and AERONET AOD observations over bright surfaces [75], [76]. Owing to MISR's multiangle measurement technique providing more constraints on the inversion, the MISR AOD is better associated with AERONET than MODIS over desert regions [77], [78]. Both MERRA-2 and DNN show that the northern India is a considerable source in all seasons, although the latter has a lower magnitude. However, this is not observed in OMAERUV. By contrast, the OMAERUV AAOD is higher in South America during all seasons, whereas the other two data only show significant AAOD during the fire season (SON). These disagreements are associated with the difference in AOD climatology (see Fig. 27 in Appendix D).

In terms of SSA, the correlation between three datasets is not explicit. Because the OMAERUV and MERRA-2 SSA are calculated from independent *a priori* aerosol models, whereas the DNN derives SSA also using observational data. The OMAERUV SSA shows the highest magnitude and the MERRA-2 SSA is lower than others, which agrees with findings in the validation section. The spatial patterns of the DNN and MERRA-2 SSA are more similar to each other than that of OMAERUV over the Southern Africa and South America biomass burning regions, although the DNN predictions are slightly higher than that of MERRA-2.

2) Regional Aerosol Absorption Climatology: We further select the major absorbing aerosol sources for regional analysis,

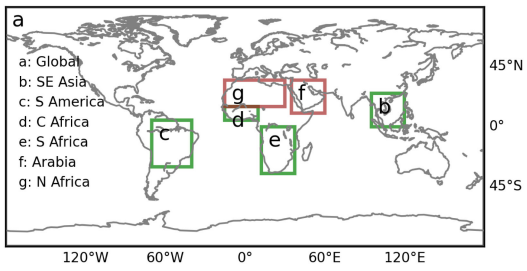

Fig. 19. Regions of interest. Green and red indicates regions dominated by smoke aerosols and dust aerosols, respectively.

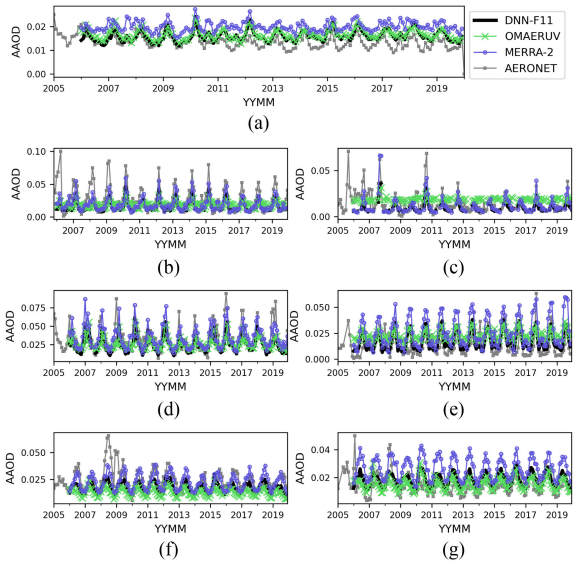

Fig. 20. AAOD time series of DNN (black), OMAERUV (green), MERRA-2 (blue) and AERONET (gray) for regions of interest.

as shown in Fig. 19. We analyze the monthly averaged aerosol absorption time series of the selected regions of interest with the mean of the corresponding AERONET measurements as a reference. According to Fig. 20, the DNN-predicted AAOD and MERRA-2 AAOD show better consistency with AERONET, whereas in terms of SSA (Fig. 21), the fluctuation of the AERONET SSA is larger than other datasets, and generally shows a stronger aerosol absorption in almost all regions. But note that the selected AERONET sites can hardly represent the whole features of the selected regions, and thus we only provide statistical analysis between DNN predictions, OMAERUV and MERRA-2. All the three datasets show similar seasonal cycles of aerosol absorption. Overall, the DNN-predicted AAOD are better associated with the other two datasets than the DNN-predicted SSA.

In biomass burning regions, e.g., Southeast Asia, Central Africa, and Southern Africa, the DNN-predicted AAOD is highly correlated with OMAERUV and MERRA-2 (R^2 >0.8), but the magnitude of the DNN predictions is more similar to the latter. An exception is South America, where the OMAERUV AAOD has the highest magnitude but with little seasonal variations, showing very low correlations with other two datasets. This is also reflected in the climatology maps (see Fig. 17). However, the corresponding SSA of DNN and MERRA-2 are in

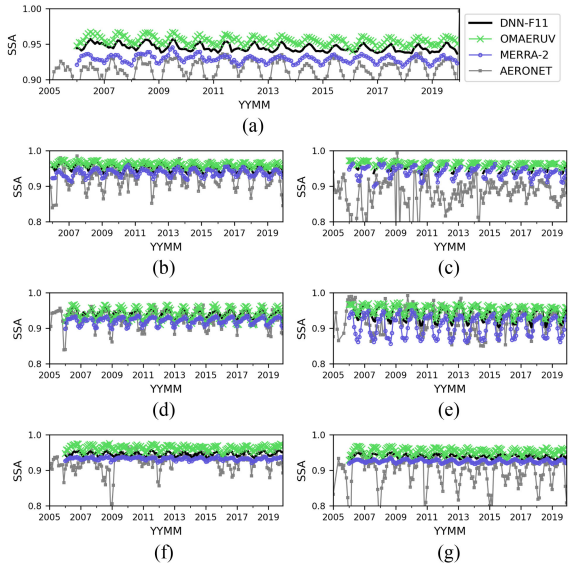

Fig. 21. SSA time series of DNN (black), OMAERUV (green), MERRA-2 (blue) and AERONET (gray) for regions of interest.

good agreement over this region (R^2 =0.93). The DNN-derived SSA is also well correlated with that of MERRA-2 in Africa biomass burning regions (R^2 =0.9), whereas lower correlations are found in Southeast Asia and the Amazonian regions (R^2 <0.8). The reason could be that the satellite observations are less sensitive to these regions as the aerosol loading is lower compared to that in Africa.

Compared with biomass burning regions, the AAOD and SSA of all datasets show less correlations in the desert regions of Arabia and Northern Africa. MERRA-2 presents the strongest absorption, whereas OMAERUV presents the lowest absorption, which is consistent with the findings in the validation section. The DNN predictions are between the two datasets, as the low measurement sensitivity of AOD in the visible band due to high surface albedo is partially compensated by the advantage of UVAI that can detect the presence of absorbing aerosols even over bright surfaces [18], [24].

IV. Conclusion

A global aerosol absorption database is important to reduce uncertainties of aerosol radiative forcing assessments. In this study, we introduced a DNN model to retrieve quantitative aerosol absorptive properties (AAOD and SSA) from the long-term OMI UVAI product. The input features are selected by both filter and wrapper methods. The hyperparameters of the DNN model are determined by grid search with tenfold cross validation. The final trained model has an accuracy of 0.0045 for AAOD prediction, which is within the expected AAOD error (0.004–0.006).

The trained DNN model has been applied to predict aerosol absorption over land for the period between 2006 and 2019. It is recommended to use the DNN-predicted AAOD and SSA with AOD at 550 nm above 0.1, as the low aerosol loading will lead to significant biases in the DNN predictions. The point-to-point validation shows that the DNN AAOD is negatively biased,

but still highly associated with the AERONET retrievals (R^2=0.89 and RMSE=0.0050). There are 82% samples fall in the excepted uncertainty of the AERONET SSA ($\pm$0.03). The analysis based on aerosol types shows that our DNN model is better at predicting fine absorbing aerosols (e.g., smoke) rather than coarse ones (e.g., mineral dust). This is mainly caused by the property of the MODIS AOD in the visible band. The consistency with ground-based measurements is lowest for nonabsorbing aerosols, since UVAI is only sensitive to absorbing aerosols according to its definition.

The DNN predictions are also compared with the OMAERUV retrievals and the MERRA-2 aerosol reanalysis. It should be noted that for both OMAERUV and MERRA-2, the SSA is calculated from *a priori* assumptions on aerosol microphysical properties, whereas no such constraints in our DNN model. Using AERONET as the reference, the DNN AAOD outperforms that of OMAERUV and MERRA-2 for all aerosol types. Specifically, OMAERUV underestimates the aerosol absorption particularly for absorbing aerosols, whereas MERRA-2 has the tendency to overestimate aerosol absorption in general. The AAOD spatial distribution and temporal variation of the DNN predictions and MERRA-2 are highly similar, particularly in the biomass burning regions. The spatial pattern of SSA is less comparable.

Our DNN model shows encouraging results in aerosol absorptive properties retrieval over land. However, the result is only satisfying for AOD above a certain level, since the low measurement sensitivity to low aerosol loading is a general problem of aerosol remote sensing. Besides, the satellite retrieved AOD in the visible band is still challenging for dust particles over bright surfaces. The application to aerosols over ocean is still restricted by the availability of reliable aerosol absorption sources used for training machine learning models. Future applications should pay attention to the above aspects.

Currently, due to the limitation of the training dataset (a hybrid datasets of satellite observations, model simulations and ground-based network measurements), our study did not properly account for spatial and temporal autocorrelation of geo-parameters. With more efforts put on the global monitoring of aerosol properties and vertical distribution, and better resolution of satellite instruments, it is expected to solve this problem either using geo-statistics that explicitly provide the spatial and temporal autocorrelation, or using advanced deep learning techniques such as CNN and/or RNN that can share weight parameters in space and/or time domain.

APPENDIX A
DATASETS

A. OMI Observations

OMI is a UV/Visible spectrometer onboard Aura (2004-present) [79]. The instrument has a large swath of 2600 km, providing around 14 orbits per day to complete a global coverage (overpass on 13:45 local time). The spatial resolution at nadir is 13 Km $\times$ 24 km (13 km $\times$ 48 km in the UV-1 band).

We take UVAI, the solar zenith angle (SZA), the satellite viewing zenith angle (VZA), the solar-satellite relative azimuth angle (RAA), the surface reflectance (a_s) and the surface pressure (P_s) from the level 2 OMAERUV version 3 product.[4] Pixels are excluded if they have SZA larger than 70°, or they are over bright surface (a_s at 388 nm higher than 0.3), contaminated by clouds (cloud fraction larger than 0.3) or sun-glint (glint angle larger than 20° over water).

OMI data suffers from the so-called row anomaly issue since 2008, which affects the quality of the level 1B radiance data at all wavelengths and consequently the level 2 products [80]. The row anomaly is caused by a partial blockage of the Earth telescope, affecting part of the across track swath. More information on row anomaly can refer to,[5] [80]. Even after filtering the data using the quality flag provided in OMAERUV product, the row anomaly can still be observed in some pixels. Therefore, we apply a moving standard deviation method to further eliminate the effects due to the row anomaly: for a given pixel, the standard deviation of UVAI (σ_{UVAI}) is calculated over a sliding window of its eight neighboring elements. Pixels with σ_{UVAI} larger than a threshold are discarded. The threshold is determined by sensitivity studies based on colocated data of OMAERUV and AERONET.

According to Fig. 22, the final determined σ_{UVAI} threshold is 0.5. This value on the one hand retains a large data size, meanwhile ensuring the consistency between the satellite and the ground-based AERONET measurements in terms of AAOD. We use AAOD as the criteria is because our study is based on the relationship between UVAI and AAOD. Both parameters contain information on the aerosol loading and the aerosol absorption. Moreover, AAOD is the common aerosol absorption parameter provided by both OMAERUV and AERONET.

The OMAERUV data product contains two UVAI parameters. One is called "residue," where $I_{\lambda}^{\mathrm{cal}}$ is calculated by a lambert equivalent reflectivity (LER) model as that done in UVAI products provided by other sensors [18], [24]. The other one is "UV index," where $I_{\lambda}^{\mathrm{cal}}$ is reprocessed by considering the effects of clouds on scattering angular variability [27]. Although the "UV index" shows a reduced across-scan bias and is better associated with the AERONET AAOD (see Fig. 22), it is overall higher than the "residue," especially over ocean where neural or negative should have been found (see [27, Fig. 12]). The reason behind is not explicitly explained. On the other hand, the method used to calculate "residue" is widely used in other satellite products, whereas "UV Index" is exclusively used for the OMAERUV product. Therefore, the term UVAI in this work refers to the "residue."

B. MODIS Observations

MODIS is a multispectral radiometer covering 36 wavelengths from 0.4 to 14.4 μm. A large swath of 2330 km allows it to complete a global coverage within 1 or 2 days. The spatial resolution ranges from 0.25 to 1 km. Currently, two MODIS instruments are operational: one is on EOS-Terra satellite launched in 1999 (descending node, overpass on 10:30 local time) and the

[4][Online]. Available: http://dx.doi.org/10.5067/Aura/OMI/DATA2004

[5][Online]. Available: http://projects.knmi.nl/omi/research/product/rowanomaly-background.php

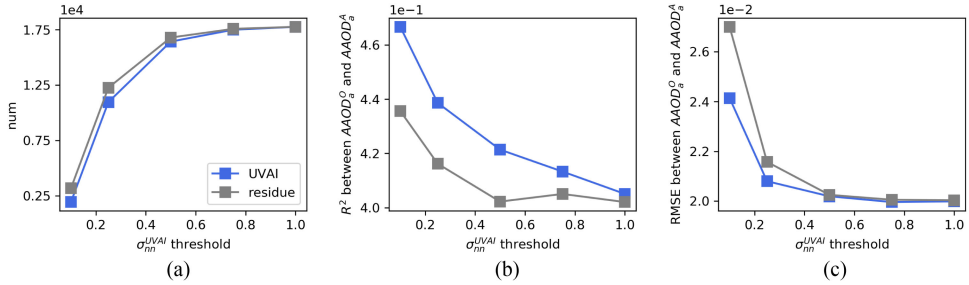

Fig. 22. Sensitivity studies to determine the σ_{UVAI} threshold. (a) the number of OMAERUV-AERONET coincidences as a function of the σ_{UVAI} threshold; (b) the correlation coefficient (R^2) between the OMAERUV AAOD (AAOD^O) and the AERONET AAOD (AAOD^A) as a function of the σ_{UVAI} threshold; (c) the root mean square error (RMSE) between AAOD^O and AAOD^A as a function of the σ_{UVAI} threshold.

other one is on EOS-Aqua launched in 2002 (ascending node, overpass on 13:30 local time) [36].

MODIS has two aerosol retrieval algorithms for 1) aerosols over ocean and vegetated dark surface ("Dark Target," DT) [37], [81], [82] and 2) aerosols over desert and arid regions ("Deep Blue," DB) [72], [83], [84]. The retrieved AOD uncertainty of DT is $\pm(0.05 + 15\%)$ over land and from $-0.04 - 10\%$ to $+0.02 + 10\%$ over ocean, and the expected uncertainty of DB is estimated better than $\pm(0.05 + 20\%)$. A hybrid AOD is also reported in a combination of DT and DB retrievals, where the method is based on the the normalized difference vegetation index [37].

In this work, we use the DT and DB combined AOD at 550 nm from Collection 6.1 level 3 daily gridded data (MYD08_D3) of MODIS/Aqua[6] to provide information of aerosol loading. The resolution of the level 3 data is $1^\circ \times 1^\circ$. Pixels with geometric cloud fraction larger than 0.3 are excluded.

C. AERONET Observations

AERONET is a ground-based remote sensing network to retrieve aerosol optical, microphysical and radiative properties for aerosol research and characterization, validation of satellite retrievals, and synergism with other databases [10].

The AERONET radiometer takes two types of measurements: 1) direct sun irradiance and 2) sky radiance. AOD is retrieved by the Beer–Bouguer Law from the direct sun measurements at a wide spectrum from 340 to 1020 nm [85]. The bias of the near-real time AOD data is estimated as 0.02 with a standard deviation of 0.02 [86]. The sky radiance measurements scan an aerosol profile at multiple scattering angles to retrieve aerosol absorption (SSA and AAOD) and other aerosol microphysics (e.g., volume size distribution, complex refractive index, and the aerosol scattering phase function, etc.) at 440, 670, 870, and 1020 nm [87]. The AERONET level 2 (quality assured) SSA has a typical uncertainty of ± 0.03 (for AOD at 440 nm larger than 0.2 [11],[88], or 0.3 [89]).

In this work, we use the AOD, AAOD, and SSA from the AERONET version 3 level 1.5 inversion almucantar product.[7] Although the level 2 data is recommended (cloud screened and quality assured) [90], we use the level 1.5 product (only cloud screened) because the level 2 data availability is insufficient for our applications. Using the expected error of AOD (assuming $\sigma_{\text{AOD}} = 0.02$) and SSA ($\sigma_{\text{SSA}} = 0.03$), and the definition of AAOD (10), we can calculate the expected error level of the AERONET AAOD using the error propagation equation (11)

$$\text{AAOD} = (1 - \text{SSA}) \times \text{AOD} \tag{10}$$

$$\sigma_{\text{AAOD}} = \sqrt{\sigma_{\text{SSA}}^2 \left(\frac{\partial \text{AAOD}}{\partial \text{SSA}}\right)^2 + \sigma_{\text{AOD}}^2 \left(\frac{\partial \text{AAOD}}{\partial \text{AOD}}\right)^2}. \tag{11}$$

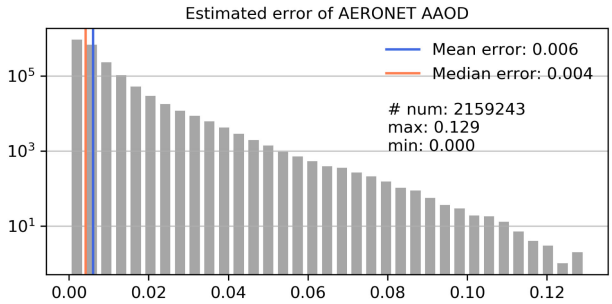

Fig. 23. Estimated error of AERONET AAOD at 550 nm calculated by (10) with AERONET AOD uncertainty of ± 0.02 and AERONET SSA uncertainty of ± 0.03.

Fig. 23 presents the distribution of the AERONET AAOD at 550 nm uncertainty for all AERONET records collected between 2006 and 2019. For more than 80% of the samples, the AAOD uncertainty is less than 0.01. The mean and median of the estimated AAOD error are 0.006 and 0.004, respectively.

D. MERRA-2 Aerosol Reanalysis

MERRA-2 is the latest modern satellite era (1980 onwards) atmospheric reanalysis [75], [76]. The model resolution is $0.5^\circ \times 0.625^\circ$ latitude by longitude with 72 hybrid-eta layers from the surface up to 0.01 hPa. MERRA-2 assimilates multiple observational AOD datasets, including MODIS, MISR, the advanced very high resolution radiometer (AVHRR) and AERONET. MERRA-2 aerosol assimilation and the total column AOD evaluation are well-documented in [75]. They elaborated that the MERRA-2 AOD constrained by observations better matches independent measurements. Improved agreement is also found for aerosol optical properties and aerosol vertical distributions [76].

[6][Online]. Available: http://dx.doi.org/10.5067/MODIS/MYD08_M3.006
[7][Online]. Available: https://aeronet.gsfc.nasa.gov

In this work, we use the MERRA-2 columnar AOD and AAOD (reported at 550 nm) provided by the MERRA-2 1-hourly time-averaged aerosol (MERRA-2 tavg1_2d_aer_Nx, 10.5067/KLICLTZ8EM9D) to compare with our derived aerosol absorption results. Furthermore, we also derive the aerosol vertical distribution information from MERRA-2. This is important because UVAI is sensitive to the aerosol layer vertical location [18], [24], [25], [28]. Many efforts have been made on measuring aerosol vertical structures [29], including ground-based lidar systems [91], [92], space-borne lidar missions [93], [94], multi-angle measurements [95], polarimetry [96], oxygen absorption at A-band [56], [97]–[103], oxygen absorption in the visible band [104] and thermal infrared [105], [106]. However, currently an aerosol vertical distribution product based on observations that has a daily global coverage as that of UVAI is still missing. Consequently, we turn to derive the ALH from the aerosol vertical profiles provided by MERRA-2.

We derive the geometric aerosol layer top height as the ALH from the MERRA-2 3-hourly instantaneous aerosol mass mixing ratio profiles (MERRA-2 inst3_3d_aer_Nv, 10.5067/LTVB4GPCOTK2). The mass mixing ratio is converted to extinction coefficients first. Then for a MERRA-2 extinction profile, we attempt to find the top boundary of an aerosol layer with the extinction coefficient lapse rate (γ_{ext}, unit: km^{-2}, [107]), which is defined as

$$\gamma_{\text{ext}}(z) = -\frac{d\beta(z)}{dz} \tag{12}$$

where $d\beta(z)$ is the extinction coefficient difference between two continuous layers, and dz is the atmospheric interval geometric thickness. Given an aerosol profile, we search upward from the surface and retain the first height at which the magnitude of γ_{ext} above this height is always smaller than a certain value. The choice of the threshold is empirical, which we select 0.01 km^{-2} based on sensitivity studies (not shown). The relation between the derived ALH and the UVAI matches our knowledge of UVAI dependence on altitude [18], [24], [25], [28].

Appendix B
Quality Filtering on DNN Predictions

The sensitivity of the DNN-derived SSA accuracy as a function of the MODIS AOD (see Fig. 24). It is found that when AOD at 550 nm is larger than 0.1, the most of the samples (83%) are within the AERONET SSA uncertainty of ± 0.03 meanwhile retaining a relative large sample size for validation.

Appendix C
Aerosol Absorption Climatology Over Ocean

Here provides a tentative analysis for the DNN-derived aerosol absorption climatology over ocean. Figs. 25 and 26 present the AAOD and SSA, respectively. All AAOD datasets show the absorbing aerosol outflows from major continental sources, e.g., the dust storms from the Sahara Desert and smoke plumes from Central and Southern Africa. The OMAERUV climatology map is more noisy compared with others due to

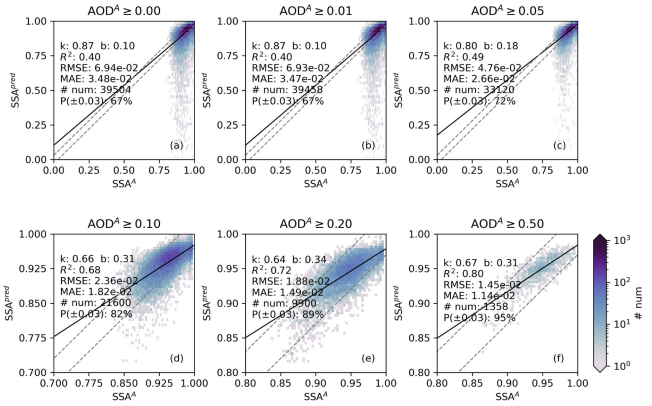

Fig. 24. Sensitivity of DNN-derived SSA accuracy to the MODIS AOD. The gray dashed lines are the ± 0.03 uncertainty of AERONET SSA.

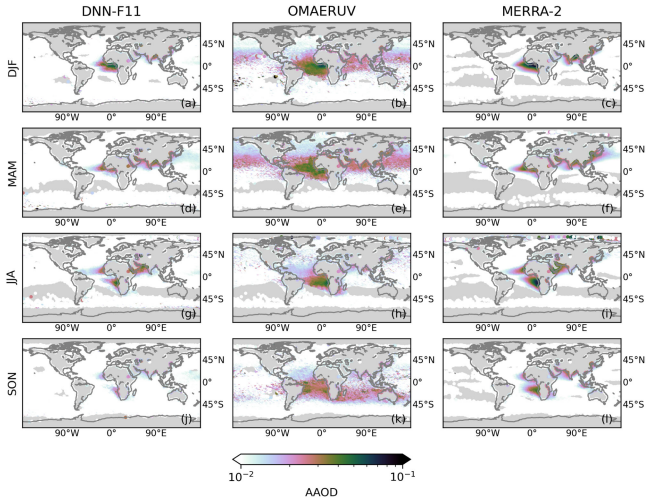

Fig. 25. Seasonal AAOD of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over ocean.

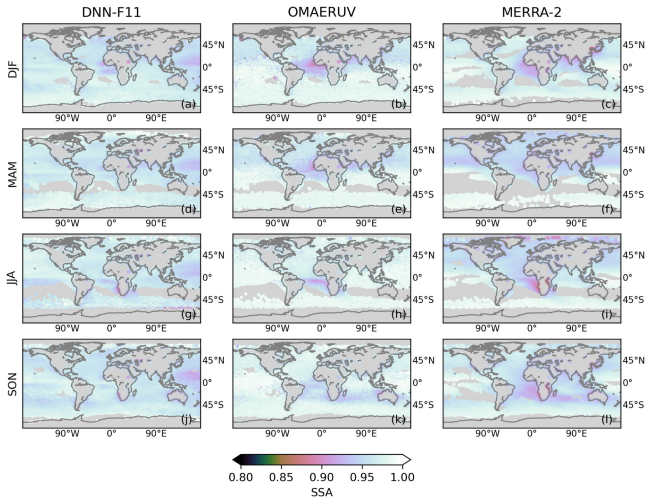

Fig. 26. Seasonal SSA of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over ocean.

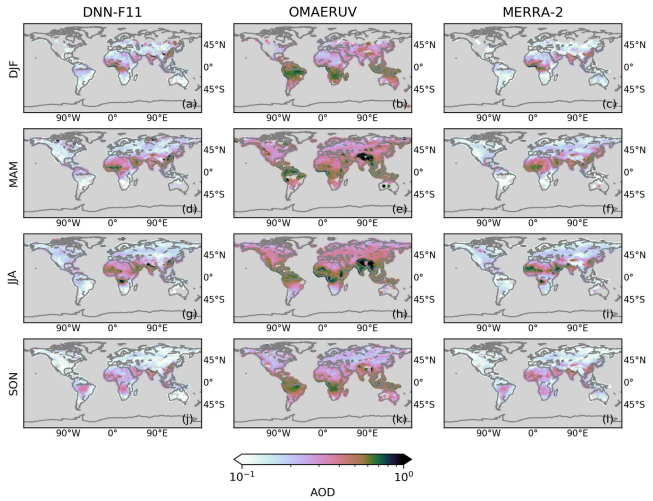

Fig. 27. Seasonal AOD of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over land.

observational and retrieval errors, missing data, etc. There are high values of the OMAERUV AAOD over remote oceans which are far away from absorbing aerosol sources, especially in the Northern Pacific during MAM and in the Southern Pacific during SON 25 (e, k).

The outflow of the Sahara dust and the Africa biomass burning plumes can be observed in the OMAERUV SSA climatology. The MERRA-2 SSA is also associated with the corresponding AAOD, particularly for smoke plumes. But the dust outflows over the northern Atlantic Ocean and the Arabian Sea can hardly be observed in the SSA maps. This may be caused by the aerosol models used in MERRA-2. The SSA of dust aerosols is between 0.77 and 0.96 (depending on particle size), whereas the absorption of black carbon is much stronger (SSA varies from 0.21 to 0.38, depending on relative humidity) [75]. On the other hand, the outflows of Africa dust and smoke are observable in the DNN-derived SSA map, but the magnitude is much less than other two datasets. The DNN predictions over ocean is overall not as good as that over land due to the limitation of spatial distribution of the training data.

Appendix D
AOD Climatology Over Land From 2006 to 2019

Fig. 27 shows the seasonal AOD climatology over land. The distribution and magnitude between DNN (i.e., MODIS) and MERRA-2 AOD are in good agreement, whereas the OMAERUV AOD (has been converted to 550 nm) is significantly higher than the other two datasets.

Acknowledgment

The authors would like to thank NASA's GES-DISC for free online access to the OMAERUV product and the MERRA-2 aerosol reanalysis, would also like to thank LAADS DAAC for free online access to the MODIS data, and also would like to thank NASA Goddard Space Flight Center AERONET project for free online access to the AERONET data.

References

[1] R. K. Pachauri *et al.*, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II, and III to the 5th Assessment Report of the Intergovernmental Panel on Climate Change*., 2014.

[2] H. Moosmüller, R. Chakrabarty, and W. Arnott, "Aerosol light absorption and its measurement: A review," *J. Quantitative Spectrosc. Radiative Transfer*, vol. 110, no. 11, pp. 844–878, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022407309000879

[3] J. Hansen, M. Sato, and R. Ruedy, "Radiative forcing and climate response," *J. Geophysical Res., Atmos.*, vol. 102, no. D6, pp. 6831–6864, 1997. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03436

[4] V. Ramanathan, P. J. Crutzen, J. T. Kiehl, and D. Rosenfeld, "Aerosols, climate, and the hydrological cycle," *Science*, vol. 294, no. 5549, pp. 2119–2124, 2001. [Online]. Available: https://science.sciencemag.org/content/294/5549/2119

[5] J. Hansen and L. Nazarenko, "Soot climate forcing via snow and Ice albedos," *Proc. Nat. Acad. Sci.*, vol. 101, no. 2, pp. 423–428, 2004. [Online]. Available: https://www.pnas.org/content/101/2/423

[6] V. S. Nair, S. S. Babu, K. K. Moorthy, A. K. Sharma, A. Marinoni, and Ajai, "Black carbon aerosols over the Himalayas: Direct and surface albedo forcing," *Tellus B, Chem. Phys. Meteorol.*, vol. 65, no. 1, 2013, Art. no. 19738. [Online]. Available: https://doi.org/10.3402/tellusb.v65i0.19738

[7] J. E. Penner *et al.*, "Aerosols, their direct and indirect effects," in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I. To the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, U.K.: Cambridge Univ. Press, 2001, pp. 289–348.

[8] A. McComiskey *et al.*, "Direct aerosol forcing: Calculation from observables and sensitivities to inputs," *J. Geophys. Res., Atmos.*, vol. 113, no. D9, 2008. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009170

[9] N. G. Loeb and W. Su, "Direct aerosol radiative forcing uncertainty based on a radiative perturbation analysis," *J. Climate*, vol. 23, no. 19, pp. 5288–5293, Oct. 2010. [Online]. Available: https://doi.org/10.1175/2010JCLI3543.1

[10] B. N. Holben *et al.*, "Aeronet-a federated instrument network and data archive for aerosol characterization," *Remote Sens. Environ.*, vol. 66, no. 1, pp. 1–16, 1998.

[11] O. Dubovik *et al.*, "Variability of absorption and optical properties of key aerosol types observed in worldwide locations," *J. Atmos. Sci.*, vol. 59, no. 3, pp. 590–608, 2002.

[12] Y. Kaufman *et al.*, "Passive remote sensing of tropospheric aerosol and atmospheric correction for the aerosol effect," *J. Geophys. Res., Atmos.*, vol. 102, no. D14, pp. 16815–16830, 1997.

[13] D. J. Diner Gerstl *et al.*, "Multi-angle imaging spectroradiometer (MISR) instrument description and experiment overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1072–1087, Jul. 1998.

[14] M. J. Garay *et al.*, "Introducing the 4.4 km spatial resolution multi-angle imaging spectroradiometer (MISR) aerosol product," *Atmos. Meas. Techn.*, vol. 13, no. 2, pp. 593–628, 2020. [Online]. Available: https://amt.copernicus.org/articles/13/593/2020/

[15] P.-Y. Deschamps *et al.*, "The polder mission: Instrument characteristics and scientific objectives," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 598–615, May 1994.

[16] P. Formenti, L. Mbemba Kabuiku, I. Chiapello, F. Ducos, F. Dulac, and D. Tanré, "Aerosol optical properties derived from polder-3/parasol (2005–2013) over the western Mediterranean Sea—Part 1: Quality assessment with aeronet and in situ airborne observations," *Atmos. Meas. Techn.*, vol. 11, no. 12, pp. 6761–6784, 2018. [Online]. Available: https://www.atmos-meas-tech.net/11/6761/2018/

[17] Y. Wei *et al.*, "Validation of polder grasp aerosol optical retrieval over China using sonet observations," *J. Quantitative Spectrosc. Radiative Transfer*, vol. 246, 2020, Art. no. 106931. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022407320300650

[18] J. R. Herman, P. K. Bhartia, O. Torres, C. Hsu, C. Seftor, and E. Celarier, "Global distribution of UV-absorbing aerosols from nimbus 7/toms data," *J. Geophys. Res., Atmos.*, vol. 102, no. D14, pp. 16911–16922, 1997. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03680

[19] R. A. Kahn and B. J. Gaitley, "An analysis of global aerosol type as retrieved by MISR," *J. Geophys. Res., Atmos.*, vol. 120, no. 9, pp. 4248–4281, 2015. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023322

[20] J. Deuzé *et al.*, "Estimate of the aerosol properties over the ocean with polder," *J. Geophys. Res., Atmos.*, vol. 105, no. D12, pp. 15329–15346, 2000.

[21] D. Tanré *et al.*, "Remote sensing of aerosols by using polarized, directional and spectral measurements within the A-Train: The PARASOL mission," *Atmos. Meas. Techn. Discuss.*, vol. 4, no. 2, pp. 2037–2069, 2011.

[22] C. Chen *et al.*, "Retrieval of desert dust and carbonaceous aerosol emissions over Africa from polder/parasol products generated by the GRASP algorithm," *Atmos. Chem. Phys.*, vol. 18, no. 16, pp. 12551–12580, 2018. [Online]. Available: https://acp.copernicus.org/articles/18/12551/2018/

[23] O. Dubovik *et al.*, "GRASP: A versatile algorithm for characterizing the atmosphere," *SPIE Newsroom*, vol. 25, 2014, Art. no. 2-1201408.

[24] O. Torres, P. Bhartia, J. Herman, Z. Ahmad, and J. Gleason, "Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis," *J. Geophys. Res., Atmos.*, vol. 103, no. D14, pp. 17099–17110, 1998.

[25] J. Sun, J. P. Veefkind, P. Van Velthoven, and P. F. Levelt, "Quantifying the single-scattering albedo for the January 2017 Chile wildfires from simulations of the OMI absorbing aerosol index," *Atmos. Meas. Techn.*, vol. 11, no. 9, 2018, Art. no. 5261.

[26] J. Sun, P. Veefkind, S. Nanda, P. Van Velthoven, and P. Levelt, "The role of aerosol layer height in quantifying aerosol absorption from ultraviolet satellite observations," *Atmos. Meas. Techn.*, vol. 12, no. 12, pp. 6319–6340, 2019.

[27] O. Torres, P. K. Bhartia, H. Jethva, and C. Ahn, "Impact of the ozone monitoring instrument row anomaly on the long-term record of aerosol products," *Atmos. Meas. Techn.*, vol. 11, no. 5, 2018, Art. no. 2701.

[28] M. de Graaf, P. Stammes, O. Torres, and R. B. A. Koelemeijer, "Absorbing aerosol index: Sensitivity analysis, application to Gome and comparison with Toms," *J. Geophys. Res., Atmos.*, vol. 110, no. D1, 2005. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005178

[29] T. Islam, Y. Hu, A. A. Kokhanovsky, and J. Wang, *Remote Sensing of Aerosols, Clouds, and Precipitation*. Amsterdam, The Netherlands: Elsevier, 2017.

[30] C. Ichoku *et al.*, "A spatio-temporal approach for global validation and analysis of modis aerosol products," *Geophys. Res. Lett.*, vol. 29, no. 12, pp. MOD1-1–MOD1-4, 2002. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013206

[31] L. A. Remer *et al.*, "Validation of MODIS aerosol retrieval over ocean," *Geophysical Res. Lett.*, vol. 29, no. 12, pp. MOD3-1–MOD3–4, 2002. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013204

[32] O. Torres, P. K. Bhartia, J. R. Herman, A. Sinyuk, P. Ginoux, and B. Holben, "A long-term record of aerosol optical depth from TOMS observations and comparison to AERONET measurements," *J. Atmos. Sci.*, vol. 59, no. 3, pp. 398–413, Feb. 2002. [Online]. Available: https://doi.org/10.1175/1520-0469(2002)059<0398:ALTROA>2.0.CO;2

[33] F.-M. Bréon, A. Vermeulen, and J. Descloitres, "An evaluation of satellite aerosol products against sunphotometer measurements," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3102–3111, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0034425711002410

[34] H. Jethva, O. Torres, and C. Ahn, "Global assessment of OMI aerosol single-scattering albedo using ground-based AERONET inversion," *J. Geophys. Res., Atmos.*, vol. 119, no. 14, pp. 9020–9040, 2014.

[35] C. Lacagnina *et al.*, "Aerosol single-scattering albedo over the global oceans: Comparing parasol retrievals with AERONET, OMI, and AEROCOM models estimates," *J. Geophys. Res., Atmos.*, vol. 120, no. 18, pp. 9814–9836, 2015. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023501

[36] L. A. Remer *et al.*, "The modis aerosol algorithm, products, and validation," *J. Atmospheric Sci.*, vol. 62, no. 4, pp. 947–973, 2005.

[37] R. C. Levy *et al.*, "The collection 6 MODIS aerosol products over land and ocean," *Atmos. Meas. Techn.*, vol. 6, no. 11, pp. 2989–3034, 2013. [Online]. Available: https://amt.copernicus.org/articles/6/2989/2013/

[38] C. Ahn, O. Torres, and H. Jethva, "Assessment of OMI near-UV aerosol optical depth over land," *J. Geophys. Res., Atmos.*, vol. 119, no. 5, pp. 2457–2473, 2014.

[39] H. Jethva and O. Torres, "Satellite-based evidence of wavelength-dependent aerosol absorption in biomass burning smoke inferred from ozone monitoring instrument," *Atmos. Chem. Phys.*, vol. 11, no. 20, pp. 10541–10551, 2011. [Online]. Available: https://acp.copernicus.org/articles/11/10541/2011/

[40] P. Legendre, "Spatial autocorrelation: Trouble or new paradigm?," *Ecology*, vol. 74, no. 6, pp. 1659–1673, 1993. [Online]. Available: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1939924

[41] X. Li, L. Peng, Y. Hu, J. Shao, and T. Chi, "Deep learning architecture for air quality predictions," *Environ. Sci. Pollut. Res.*, vol. 23, no. 22, pp. 22408–22417, 2016.

[42] T. Li, H. Shen, Q. Yuan, X. Zhang, and L. Zhang, "Estimating ground-level $PM_{2.5}$ by fusing satellite and station observations: A GEO-intelligent deep learning approach," *Geophys. Res. Lett.*, vol. 44, no. 23, pp. 11985–11993, 2017. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL075710

[43] A. M. Michalak and S. Shlomi, "A Geostatistical data assimilation approach for estimating groundwater plume distributions from multiple monitoring events," *Geophys. Monograph-Amer. Geophys. Union*, vol. 171, 2007, Art. no. 73.

[44] L. Li, R. Shi, L. Zhang, J. Zhang, and W. Gao, "The data fusion of aerosol optical thickness using universal kriging and stepwise regression in East China," in *Remote Sensing and Modeling of Ecosystems for Sustainability XI*, W. Gao, N.-B. Chang, and J. Wang, Eds., vol. 9221, International Society for Optics and Photonics, 2014, pp. 219–229. [Online]. Available: https://doi.org/10.1117/12.2061764

[45] E. Tapoglou, G. P. Karatzas, I. C. Trichakis, and E. A. Varouchakis, "A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation," *J. Hydrol.*, vol. 519, pp. 3193–3203, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S002216941400835X

[46] M. K. Singh, R. Gautam, and P. Venkatachalam, "Bayesian merging of MISR and MODIS aerosol optical depth products using error distributions from Aeronet," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 12, pp. 5186–5200, Dec. 2017.

[47] K. Ishitsuka, T. Mogi, K. Sugano, Y. Yamaya, T. Uchida, and T. Kajiwara, "Resistivity-based temperature estimation of the Kakkonda geothermal field, Japan, using a neural network and neural Kriging," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 8, pp. 1154–1158, Aug. 2018.

[48] D. Tarasov, A. Buevich, A. Sergeev, and A. Shichkin, "High variation topsoil pollution forecasting in the Russian subarctic: Using artificial neural networks combined with residual Kriging," *Appl. Geochem.*, vol. 88, pp. 188–197, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0883292717302421

[49] F. Biancofiore *et al.*, "Analysis of surface ozone using a recurrent neural network," *Sci. Total Environ.*, vol. 514, pp. 379–387, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S004896971500128X

[50] C. Wen *et al.*, "A novel spatiotemporal convolutional long short-term neural network for air pollution prediction," *Sci. Total Environ.*, vol. 654, pp. 1091–1099, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0048969718344413

[51] X. Chen *et al.*, "Joint retrieval of the aerosol fine mode fraction and optical depth using MODIS spectral reflectance over northern and eastern China: Artificial neural network method," *Remote Sens. Environ.*, vol. 249, 2020, Art. no. 112006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S003442572030376X

[52] P. Gupta and S. A. Christopher, "Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2 A neural network approach," *J. Geophys. Res., Atmos.*, vol. 114, no. D20, 2009. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD011497

[53] L. E. Olcese, G. G. Palancar, and B. M. Toselli, "A method to estimate missing AERONET AOD values based on artificial neural networks," *Atmos. Environ.*, vol. 113, pp. 140–150, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1352231015300832

[54] W. Qin, L. Wang, A. Lin, M. Zhang, and M. Bilal, "Improving the estimation of daily aerosol optical depth and aerosol radiative effect using an optimized artificial neural network," *Remote Sens.*, vol. 10, no. 7, 2018, Art. no. 1022. [Online]. Available: https://www.mdpi.com/2072-4292/10/7/1022

[55] S. Mauceri, B. Kindel, S. Massie, and P. Pilewskie, "Neural network for aerosol retrieval from hyperspectral imagery," *Atmos. Meas. Techn.*, vol. 12, no. 11, pp. 6017–6036, 2019. [Online]. Available: https://amt.copernicus.org/articles/12/6017/2019/

[56] S. Nanda *et al.*, "A neural network radiative transfer model approach applied to the tropospheric monitoring instrument aerosol height algorithm," *Atmos. Meas. Techn.*, vol. 12, no. 12, pp. 6619–6634, 2019.

[57] J. Xiong, R. Yao, W. Wang, W. Yu, and B. Li, "A spatial-and-temporal-based method for rapid particle concentration estimations in an urban environment," *J. Cleaner Prod.*, vol. 256, 2020, Art. no. 120331. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652620303784

[58] X. Zhu, Q. Zhang, C.-Y. Xu, P. Sun, and P. Hu, "Reconstruction of high spatial resolution surface air temperature data across China: A new GEO-intelligent multisource data-based machine learning technique," *Sci. Total Environ.*, vol. 665, pp. 300–313, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0048969719305571

[59] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Mar. 2003.

[60] D. N. Reshef *et al.*, "Detecting novel associations in large data sets," *Science*, vol. 334, no. 6062, pp. 1518–1524, 2011.

[61] T. M. Phuong, Z. Lin, and R. B. Altman, "Choosing SNPs using feature selection," in *Proc. IEEE Comput. Syst. Bioinf. Conf.*, 2005, pp. 301–309.

[62] P. M. Granitto, C. Furlanello, F. Biasioli, and F. Gasperi, "Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products," *Chemometrics Intell. Lab. Syst.*, vol. 83, no. 2, pp. 83–90, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169743906000232

[63] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geosci. Front.*, vol. 7, no. 1, pp. 3–10, 2016.

[64] A. Karpatne, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar, "Machine learning for the geosciences: Challenges and opportunities," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 8, pp. 1544–1554, Aug. 2019.

[65] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[66] M. Reichstein *et al.*, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.

[67] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

[68] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, 2015, pp. 1–15.

[69] A. Alfadda, S. Rahman, and M. Pipattanasomporn, "Solar irradiance forecast using aerosols measurements: A data driven approach," *Sol. Energy*, vol. 170, pp. 924–939, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038092X18305309

[70] D. M. Giles *et al.*, "An analysis of AERONET aerosol absorption properties and classifications representative of aerosol source regions," *J. Geophys. Res., Atmos.*, vol. 117, no. D17, 2012. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JD018127

[71] A. Cazorla *et al.*, "Relating aerosol absorption due to soot, organic carbon, and dust to emission sources determined from in-situ chemical measurements," *Atmos. Chem. Phys.*, vol. 13, no. 18, pp. 9337–9350, 2013. [Online]. Available: https://acp.copernicus.org/articles/13/9337/2013/

[72] N. C. Hsu *et al.*, "Enhanced deep blue aerosol retrieval algorithm: The second generation," *J. Geophys. Res., Atmos.*, vol. 118, no. 16, pp. 9296–9315, 2013. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50712

[73] J. Wei, Z. Li, Y. Peng, and L. Sun, "MODIS collection 6.1 aerosol optical depth products over land and ocean: Validation and comparison," *Atmos. Environ.*, vol. 201, pp. 428–440, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1352231018308513

[74] O. Torres, R. Decae, J. Veefkind, and G. De Leeuw, "OMI aerosol retrieval algorithm, OMI algorithm theoretical basis document: Clouds, aerosols, and surface UV irradiance, vol. 3, version 2," OMI-ATBD-03, 2002. [Online]. Available: http://eospso.gsfc.nasa.gov/eoshomepage/forscientists/atbd/docs/OMI/ATBD-OMI-03.pdf

[75] C. A. Randles *et al.*, "The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation," *J. Climate*, vol. 30, no. 17, pp. 6823–6850, Jul. 2017. [Online]. Available: https://doi.org/10.1175/JCLI-D-16-0609.1

[76] V. Buchard *et al.*, "The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies," *J. Climate*, vol. 30, no. 17, pp. 6851–6872, Jul. 2017. [Online]. Available: https://doi.org/10.1175/JCLI-D-16-0613.1

[77] O. V. Kalashnikova and R. A. Kahn, "Mineral dust plume evolution over the atlantic from MISR and MODIS aerosol retrievals," *J. Geophys. Res., Atmos.*, vol. 113, no. D24, 2008. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010083

[78] M. I. Mishchenko, L. Liu, I. V. Geogdzhayev, L. D. Travis, B. Cairns, and A. A. Lacis, "Toward unified satellite climatology of aerosol properties: 3 MODIS versus MISR versus AERONET," *J. Quantitative Spectrosc. Radiative Transfer*, vol. 111, no. 4, pp. 540–552, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022407309003343

[79] P. F. Levelt *et al.*, "The ozone monitoring instrument," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1093–1101, May 2006.

[80] V. M. E. Schenkeveld *et al.*, "In-flight performance of the ozone monitoring instrument," *Atmos. Meas. Techn.*, vol. 10, no. 5, pp. 1957–1986, 2017. [Online]. Available: https://amt.copernicus.org/articles/10/1957/2017/

[81] Y. Kaufman, D. Tanré, L. A. Remer, E. Vermote, A. Chu, and B. Holben, "Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer," *J. Geophys. Res., Atmos.*, vol. 102, no. D14, pp. 17051–17067, 1997.

[82] D. Tanré, Y. J. Kaufman, M. Herman, and S. Mattoo, "Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances," *J. Geophys. Res., Atmos.*, vol. 102, no. D14, pp. 16971–16988, 1997. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03437

[83] N. C. Hsu, S.-C. Tsay, M. D. King, and J. R. Herman, "Aerosol properties over bright-reflecting source regions," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 557–569, Mar. 2004.

[84] N. C. Hsu, S. Tsay, M. D. King, and J. R. Herman, "Deep blue retrievals of asian aerosol properties during ACE-Asia," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3180–3195, Nov. 2006.

[85] T. F. Eck *et al.*, "Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols," *J. Geophys. Res., Atmos.*, vol. 104, no. D24, pp. 31333–31349, 1999.

[86] D. M. Giles *et al.*, "Advancements in the aerosol robotic network (AERONET) version 3 database-automated near-real-time quality control algorithm with improved cloud screening for sun photometer aerosol optical depth (AOD) measurements," *Atmos. Meas. Techn.*, vol. 12, no. 1, pp. 169–209, 2019.

[87] O. Dubovik and M. D. King, "A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements," *J. Geophys. Res., Atmos.*, vol. 105, no. D16, pp. 20673–20696, 2000.

[88] O. Dubovik *et al.*, "Accuracy assessments of aerosol optical properties retrieved from aerosol robotic network (AERONET) sun and sky radiance measurements," *J. Geophys. Res., Atmos.*, vol. 105, no. D8, pp. 9791–9806, 2000.

[89] A. Sinyuk *et al.*, "The AERONET version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to version 2," *Atmos. Meas. Tech. Discuss*, vol. 2020, pp. 1–80, 2020.

[90] A. Smirnov, B. Holben, T. Eck, O. Dubovik, and I. Slutsker, "Cloud-screening and quality control algorithms for the AERONET database," *Remote Sens. Environ.*, vol. 73, no. 3, pp. 337–349, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0034425700001097

[91] E. J. Welton, J. R. Campbell, J. D. Spinhirne, and V. S. Scott III, "Global monitoring of clouds and aerosols using a network of micropulse Lidar systems," in *Lidar Remote Sens. Ind. Environ. Monit.*, vol. 4153, 2001, pp. 151–158.

[92] G. Pappalardo *et al.*, *EARLINET Atmos. Meas. Techn.*, no. 8, pp. 2389–2409, 2014.

[93] B. E. Schutz, H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio, "Overview of the ICESat mission," *Geophys. Res. Lett.*, vol. 32, no. 21, 2005. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/action/showCitFormats?doi=10.1029%2F2005GL024009

[94] D. M. Winker *et al.*, "Overview of the calipso mission and caliop data processing algorithms," *J. Atmos. Ocean. Technol.*, vol. 26, no. 11, pp. 2310–2323, 2009.

[95] D. L. Nelson, M. J. Garay, R. A. Kahn, and B. A. Dunst, "Stereoscopic height and wind retrievals for aerosol plumes with the MISR interactive explorer (MINX)," *Remote Sens.*, vol. 5, no. 9, pp. 4593–4628, 2013.

[96] O. Dubovik *et al.*, "Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations," *Atmos. Meas. Techn.*, vol. 4, no. 5, pp. 975–1018, 2011.

[97] L. Duforêt, R. Frouin, and P. Dubuisson, "Importance and estimation of aerosol vertical structure in satellite ocean-color remote sensing," *Appl. Opt.*, vol. 46, no. 7, pp. 1107–1119, 2007.

[98] P. Dubuisson *et al.*, "Estimating the altitude of aerosol plumes over the ocean from reflectance ratio measurements in the O_2 a-band," *Remote Sens. Environ.*, vol. 113, no. 9, pp. 1899–1911, 2009.

[99] S. Sanghavi, J. Martonchik, J. Landgraf, and U. Platt, "Retrieval of the optical depth and vertical distribution of particulate scatterers in the atmosphere using O_2 a-and b-band Sciamachy observations over Kanpur: A case study," *Atmos. Meas. Techn.*, vol. 5, no. 5, 2012, Art. no. 1099.

[100] P. Wang, L. Tilstra, M. de Graaf, and P. Stammes, "Interpretation of fresco cloud retrievals in case of absorbing aerosol events," *Atmos. Chem. Phys.*, vol. 12, no. 19, pp. 9057–9077, 2012.

[101] A. Sanders *et al.*, "Evaluation of the operational aerosol layer height retrieval algorithm for sentinel-5 precursor: Application to O_2 a band observations from GOME-2A, ATMOS," *Meas. Tech*, vol. 8, pp. 4947–4977, 2015.

[102] A. F. J. Sanders and J. F. de Haan, "TROPOMI ATBD of the aerosol layer height product," 2016. Accessed: Mar. 25, 2020. [Online]. Available: http://www.tropomi.eu/sites/default/files/files/S5P-KNMI-L2-0006-RP-TROPOMI_ATBD_Aerosol_Height-v1p0p0-20160129.pdf

[103] X. Xu *et al.*, "Passive remote sensing of altitude and optical depth of dust plumes using the oxygen a and b bands: First results from epic/dscovr at lagrange-1 point," *Geophys. Res. Lett.*, vol. 44, no. 14, pp. 7544–7554, 2017.

[104] J. Chimot *et al.*, "An exploratory study on the aerosol height retrieval from OMI measurements of the 477 nm O_2–O_2 spectral band using a neural network approach," *Atmos. Meas. Techn.*, vol. 10, no. 3, 2017, Art. no. 783.

[105] C. Pierangelo, A. Chédin, S. Heilliette, N. Jacquinet-Husson, and R. Armante, "Dust altitude and infrared optical depth from AIRS," *Atmos. Chem. Phys.*, vol. 4, no. 7, pp. 1813–1822, 2004.

[106] S. Vandenbussche, S. Kochenova, A. C. Vandaele, N. Kumps, and M. De Mazière, "Retrieval of desert dust aerosol vertical profiles from IASI measurements in the TIR atmospheric window," *Atmos. Meas. Techn.*, vol. 6, no. 10, pp. 2577–2591, 2013.

[107] P. Tian *et al.*, "Aerosol vertical distribution and optical properties over China from long-term satellite and ground-based remote sensing," *Atmos. Chem. Phys.*, vol. 17, no. 4, pp. 2509–2523, 2017. [Online]. Available: https://acp.copernicus.org/articles/17/2509/2017/

Jiyunting Sun received the M.S. degree in geoscience and remote sensing from Delft University of Technology, Delft, the Netherlands, in 2016, where she is currently working towards the Ph.D. degree in atmospheric remote sensing.

Her research interests focus on applications of machine learning algorithms to atmospheric remote sensing.

Pepijn Veefkind's photograph and biography not available at the time of publication.

Peter van Velthoven's photograph and biography not available at the time of publication.

Pieternel F. Levelt's photograph and biography not available at the time of publication.