JGR Atmospheres

RESEARCH ARTICLE
10.1029/2019JD031884

Key Points:
- The EnKF-based data synergy technique greatly reduced satellite AOD uncertainties
- The impact of ground observation is extended according to its representativeness
- We provide an improved monthly AOD data set for climate and environmental research

Supporting Information:
- Supporting Information S1

Correspondence to:
J. Li,
jing-li@pku.edu.cn

Citation:
Li, J., Kahn, R. A., Wei, J., Carlson, B. E., Lacis, A. A., Li, Z., et al. (2020). Synergy of satellite- and ground-based aerosol optical depth measurements using an ensemble Kalman filter approach. *Journal of Geophysical Research: Atmospheres*, *125*, e2019JD031884. https://doi.org/10.1029/2019JD031884

Received 24 OCT 2019
Accepted 17 FEB 2020
Accepted article online 19 FEB 2020

Synergy of Satellite- and Ground-Based Aerosol Optical Depth Measurements Using an Ensemble Kalman Filter Approach

Jing Li[1], Ralph A. Kahn[2], Jing Wei[3,4], Barbara E. Carlson[5], Andrew A. Lacis[5], Zhanqing Li[4], Xichen Li[6], Oleg Dubovik[7], and Teruyuki Nakajima[8]

[1]Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China, [2]Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA, [3]State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [4]Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, [5]NASA Goddard Institute for Space Studies, New York, NY, USA, [6]International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [7]Laboratoire d'Optique Atmosphérique, CNRS/Universite Lille, Villeneuve d'Ascq, France, [8]Japan Aerospace Exploration Agency, Tsukuba Space Center, Tsukuba, Japan

Abstract Satellite- and ground-based remote sensing are two widely used techniques to measure aerosol properties. However, neither is perfect in that satellite retrievals suffer from various sources of uncertainties, and ground observations have limited spatial coverage. In this study, focusing on improving estimates of aerosol information on large scale, we develop a data synergy technique based on the ensemble Kalman filter (EnKF) to effectively combine these two types of measurements and yield a monthly mean aerosol optical depth (AOD) product with global coverage and improved accuracy. We first construct a 474-member ensemble using 11 monthly mean AOD data sets to represent the variability of the AOD field. Then Moderate Resolution Imaging Spectroradiometer AOD retrievals are selected as the background field into which ground-based measurements from 135 Aerosol Robotic Network sites are assimilated using the EnKF. Compared with satellite data, the bias and root-mean-square errors of the combined field are greatly reduced, and correlation coefficients are greatly improved. Moreover, cross validation shows that at locations where surface observations were not assimilated, the reduction in root-mean-square error and bias and the increase in correlation can still reach ~20%. Locations where the spatial representativeness of AOD is large or the site density is high are where the greatest changes are typically found. This study shows that the EnKF technique effectively extends the information obtained at surface sites to a larger area, paving the way for combining information from different types of measurements to yield better estimates of aerosol properties as well as their space-time variability.

1. Introduction

Information about atmospheric aerosol loading and its variability is critical when assessing the anthropogenic forcing of climate change as well as for monitoring environmental pollution. However, because aerosols have complicated chemical compositions and short atmospheric lifetimes, their properties are highly variable in space and time, rendering the accurate measurement difficult. Remote sensing, either from satellite platforms or from the Earth's surface, has been widely used to retrieve aerosol properties. Satellite-based aerosol remote sensing, such as that done using the Moderate Resolution Imaging Spectroradiometer (MODIS, (Salomonson et al., 2002)) and the Multi-angle Imaging SpectroRadiometer (MISR, (Diner et al., 1998)), measures the backscattered radiation from the Earth-atmosphere system in the visible to near-infrared bands. These instruments retrieve column aerosol loading, namely, aerosol optical depth (AOD), by making assumptions about aerosol vertical distributions, their complex refractive indices, and size distributions and by parameterizing surface reflectance. Satellite retrievals have the advantage of extensive spatial coverage. However, the above-mentioned assumptions made in the retrieval algorithms often introduce uncertainties in the products. On the other hand, surface-based remote sensing of AOD from direct sun measurements avoids making many of these assumptions and can be

more accurate. In fact, ground retrievals of AOD are conventionally used to validate satellite data. However, the spatial coverage of surface sites is minuscule, making these data inadequate for large-scale studies.

Numerous studies have compared or validated different satellite-based AOD retrievals against ground observations (Kahn et al., 2007; Kahn et al., 2009; Levy et al., 2013; Mishchenko et al., 2010; Sayer et al., 2015; Wei et al., 2019). There is general agreement on large scales, but regionally, differences can still be significant (Li et al., 2009; Prasad & Singh, 2007; Boiyo et al., 2017; Kahn et al., 2010; Tao et al., 2015; Li et al., 2019; Wang et al., 2017; Wei et al., 2018; Wei, Peng, et al., 2019). Discrepancies are also observed among different satellite data sets (Banks et al., 2013; Bibi et al., 2015; de Leeuw et al., 2015; de Leeuw et al., 2018). Moreover, even for the monthly mean product in which random errors have been largely averaged out, there are still disagreements. For example, (Li et al., 2014a, 2014b) applied a series of different spectral decomposition techniques to compare the space-time variabilities of different satellite data sets against surface measurements and revealed large disagreements over regions with high aerosol loadings, such as Africa, South America, East Asia, and India. (Wei, Li, et al., 2019) also showed unsatisfactory performances of 11 satellite-derived monthly mean AOD products by comparing with Aerosol Robotic Network (AERONET)) AOD retrievals. The message conveyed by these works is that current satellite AOD data sets still have relatively large uncertainties and that improvements are needed to meet the accuracy requirements of climate forcing and air pollution research.

Although improving satellite products fundamentally requires refining the retrieval algorithm and/or instrument design, it is also important to make more effective use of existing products, tapping into their respective strengths and shortcomings. In particular, a promising approach would be to take advantage of the better spatial coverage of satellite data and the better accuracy of ground-based measurements to produce a data set with global coverage and improved accuracy. Previously, (Tang et al., 2016) used a Bayesian maximum entropy method to merge different satellite data sets and AERONET data over East Asia with spatial autocorrelation and the uncertainties of the different products taken into consideration. (Fu et al., 2018) combined MODIS and AERONET retrievals to estimate surface $PM_{2.5}$ concentrations based on the spatial correlation between the two data sets. Such data fusion-type studies were also conducted using a probabilistic approach (Nirala, 2008; Xu et al., 2015) and a geostatistical approach (Nguyen et al., 2012; Wang et al., 2013).

To combine satellite- and ground-based measurements, a critical piece of information is the spatial representativeness of the surface sites because this determines how large an area the measurements at a surface site can represent. For example, if a site is representative of a large area, measurements from that site can be used to infer aerosol information at distances for away, especially when ground observations are not available at these distant locations. However, this cannot be done if a site is less representative. Although some previous work on data fusion considered the spatial correlation of AOD, none of them explicitly related the data synergy to spatial representativeness. Moreover, in their estimation of spatial correlation, usually one or two data sets were considered. This may not provide enough samples to fully represent the variability of the AOD field and its spatial correlation patterns.

In our previous studies (Li et al., 2016; Li et al., 2017), we used an ensemble-based method to objectively and quantitatively estimate the spatial representativeness of the AOD field. Specifically, we assessed how much the uncertainty, represented by the ensemble spread, could be reduced by assimilating observations obtained at a specific location. The representativeness thus indicates the spatial range over which aerosol variability can be predicted based on observations at a single site. With this information in hand, improvements can be made to large-scale AOD estimates using surface observations. We therefore further develop an ensemble Kalman filter (EnKF)-based technique to effectively combine satellite and ground observations. This work is a direct follow-up of (Li et al., 2016). In that paper, we constructed a multidata set ensemble and examined the changes in the background error covariance after assimilating ground-based observations using the EnKF technique. Here, we focus on updating the means by taking advantage of the AOD spatial representativeness. In the next section, we introduce the method in detail. Section 3 presents comparisons between the combined field and the original field and cross-validation results. The last section summarizes the study and includes some discussion.

2. Method and Data Sets

2.1. The EnKF-Based Data Synergy

We start by introducing the basics of the Kalman filter. The Kalman filter (Kalman & Bucy, 1961) is a statistical algorithm that uses a series of measurements observed over time, containing errors and noise, and produces estimates of unknown variables that tend to be more accurate than those based on single measurements alone. This is realized by estimating a joint probability distribution over the variables at each time frame. Mathematically, the Kalman filter assumes that the true state variable $\mathbf{x}$ evolves from time j to time $j+1$ according to

$$\mathbf{x}(j+1) = \mathbf{M}\mathbf{x}(j) + \eta(j), \tag{1}$$

where $\mathbf{M}$ is a linear model of system dynamics and $\eta(j)$ is the random error at time j. The latter is assumed to be drawn from a zero-mean multivariate normal distribution with covariance $\mathbf{P}$. The diagonal elements of $\mathbf{P}$ are the variances of $\mathbf{x_j}$ at each location, and the off-diagonal elements are the covariances between different locations.

At time j, an observation $\mathbf{y_j}$ of the true state $\mathbf{x_j}$ is made according to

$$\mathbf{y}(j) = \mathbf{H}\mathbf{x}(j) + \varepsilon(j), \tag{2}$$

where $\mathbf{H}$ is the observation operator that maps from the observation space to the model space. The observation error $\varepsilon(j)$ is assumed to be unbiased, that is, $\sum\varepsilon(j) = 0$. The observations have a known error covariance matrix $\mathbf{R}$. It is usually assumed that there is no correlated error in observations so that $\mathbf{R}$ is a diagonal matrix whose elements are the errors of each observation.

In the analysis stage, the state variable is updated as follows:

$$\mathbf{x^a}(j) = \mathbf{x^b}(j) + \mathbf{K}(j)(\mathbf{y}(j) - \mathbf{H}\mathbf{x^a}(j)), \tag{3}$$

Here, superscripts denote the analysis field and $\mathbf{b}$ denotes the background field.

$$\mathbf{K} = \mathbf{P}\mathbf{H}^{\mathbf{T}}\left(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathbf{T}} + \mathbf{R}\right)^{-1} \tag{4}$$

is the Kalman gain. The analysis field is thus a weighted average of the background field and observations with more weight given to estimates with higher certainties (smaller errors). The Kalman filter is a popular data assimilation technique in atmospheric and oceanic sciences. Previous studies have used this technique to assimilate satellite- and ground-based aerosol observations into chemical transport models which showed notable improvements in model results (Rubin & Collins, 2014; Rubin et al., 2017; Yumimoto & Takemura, 2011). In addition to assimilating observations into a model, it can also assimilate observations from different sources. In our realization, the state $\mathbf{x}$ is the true AOD value at a certain location. For easier implementation, we focus on monthly mean AOD at a $1° \times 1°$ resolution. Therefore, $\mathbf{x}$ is the true monthly mean AOD in each $1° \times 1°$ grid box. The observations to be assimilated, $\mathbf{y}$, are the observations obtained at surface sites, and $\mathbf{H}$ is the vector that maps the scattered observation sites to the regular $1° \times 1°$ grid. The observation error is estimated as the measurement error plus the representation error. The measurement error is set to 0.01 which is the accuracy of the CE-318 Sun photometers used by AERONET at visible wavelengths under cloud-free conditions (Eck et al., 1999; Holben et al., 1998). The representation error represents the subgrid variability, that is, the variability within each $1° \times 1°$ grid. This quantity is not easy to accurately define and will be discussed in more detail in a separate study. As an approximation here, we use the Level 2 AOD product from MODIS, average it to $0.1° \times 0.1°$ grids, and calculate the standard deviation of all $0.1°$ grids falling within the larger $1°$ grid.

The Kalman filter itself has limited value in large-scale problems because the background error covariance is usually unknown. Therefore, the EnKF was proposed (Evensen, 1994) to overcome this problem by replacing the true background covariance with the sample covariance. This means that we need to generate an ensemble to approximate the distribution of the state vector $\mathbf{x}$. With multiple samples in each grid cell, the spatial correlation between different grid cells can thus be estimated. The size of the ensemble must

be large enough to represent the bulk of the variance in **x,** and its distribution is ideally unbiased. This method is particularly suitable for AOD data because we have multiple sensors retrieving this variable, allowing for construction of a large enough ensemble. In our case, the ensemble is constructed as

$$\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_N}] = [\mathbf{x_i}], \tag{5}$$

where each **x** is one set of satellite data. We use a total of 11 data sets that are described in section 2.3. The background error covariance matrix can then be calculated as

$$\mathbf{P} = \frac{1}{N-1}\left(\mathbf{X}-\overline{\mathbf{X}}\right)\left(\mathbf{X}-\overline{\mathbf{X}}\right)^T, \tag{6}$$

where N is the ensemble size. Using the above expression of **P** in equation (4), analysis AODs can be obtained.

It is important to note that in the traditional EnKF, the background error covariance matrix also evolves with time. However, this is not possible for the multisensor problem here because we have only 11 observations (from the 11 data sets) at each time frame. As a result, we construct a "stable" ensemble using all monthly mean AODs from all data sets (with some quality control, as described in section 2.3). Considering that aerosol properties have distinct seasonal features at many locations, we also construct four seasonal ensembles, one for each season. (Li et al., 2016) have shown that the "stable" ensemble works well in capturing the spatial variability and covariability.

2.2. Covariance Localization

In practice, a number of issues can arise when implementing the EnKF. A typical problem is the existence of spurious correlation (Anderson, 2001). This refers to the correlation between remote locations that are not physically related, possibly caused by insufficient or biased sampling. For example, if the ensemble size is too small or does not cover all seasons sufficiently. The consequence of spurious correlation is that the state variable may be incorrectly impacted by a distant observation, deteriorating the analysis results. Covariance localization is an effective method proposed to combat this problem (Hamill et al., 2001; Houtekamer & Mitchell, 2001) by cutting off the longer-range correlations in the error covariance matrix beyond a specified distance. In our study, we do note that localization can help improve the results, a topic further discussed in section 3. We, therefore, multiply the background error covariance matrix by the localization function suggested by Gaspari and Cohn (Gaspari & Cohn, 1999) as follows:

$$\rho = \begin{cases} -\frac{1}{4}\left(\frac{|z|}{c}\right)^5 + \frac{1}{2}\left(\frac{|z|}{c}\right)^4 + \frac{5}{8}\left(\frac{|z|}{c}\right)^3 - \frac{5}{3}\left(\frac{|z|}{c}\right)^2 + 1, & 0 \le |z| \le c \\ \frac{1}{12}\left(\frac{|z|}{c}\right)^5 - \frac{1}{2}\left(\frac{|z|}{c}\right)^4 + \frac{5}{8}\left(\frac{|z|}{c}\right)^3 + \frac{5}{3}\left(\frac{|z|}{c}\right)^2 - 5\left(\frac{|z|}{c}\right) + 4 - \frac{2}{3}\left(\frac{c}{|z|}\right), & c \le |z| \le 2c \\ 0, & 2c \le |z| \end{cases}, \tag{7}$$

where z is the Euclidean distance between either two grid cells or a grid cell and an observation and c is a length scale such that the correlation decreases to zero at $2c$. The length scale is generally set as $c = \sqrt{10/3}l$, where l is a predefined cutoff distance (Lorenc, 2003). Here, we let l be equal to 3,000 km. We tested cutoff distances from 1,000 to 5,000 km and found that the results stabilized once the distance was for distances below 3,000 km. Note that this distance is only globally representative, whereas the optimal localization scale can vary with location.

In section 3, the results with and without localization are compared and find that localization indeed yields better estimates in our case.

2.3. Evaluation Approaches

We use three statistical parameters, namely, absolute bias, root-mean-square error (RMSE), and Pearson's correlation coefficient to evaluate the merged data set. In addition, it is more important to examine the effect of our EnKF approach at places where ground observation is not assimilated. For this purpose, we develop two cross-validation (CV) schemes: a regional threefold scheme and a leave-one-out (LOO) scheme. The

second scheme is easier to explain and is more conventional for CV, that is, for each iteration of LOO, we select data from one site as the validation set and use those from the remaining sites to train the model, until all sites have been validated once. The first approach is a regional adaptation of the traditional K-fold CV. Specifically, we divide the globe into 13 regions based on the spatial distribution of AERONET sites (will be further described in section 2.4), namely, Western North America, Eastern North America, South America, Europe, North Africa, Sahel, South Africa, Central Asia, India, East Asia, Southeast Asia, Oceania, and Islands. In this CV procedure, we first randomly divide the sites within each region into three groups (regions with fewer than three sites are skipped). Then, for each iteration, we select one group as the validation set and the remaining two groups as the training set, until every group has been validated once. The reason for not using a global K-fold CV is that our data synergy takes advantage of the spatial representativeness of surface sites, and the impact is more regional. Therefore, it is not appropriate to use remote sites to validate the results, as they will not be affected. For example, if we use a global K-fold CV, it might be the case that the training sites are mostly located in the Americas and Europe, whereas the validation sites are in Asia. If this is the case, no improvements can be observed as based on our localization process, sites more distant than 3,000 km will not affect the result. It is not expected that improvements would be observed in Asia by assimilating American and European sites, as they have weak physical relationships.

Moreover, we also selected additional 54 sites in the AERONET network that are not used in data synergy and 8 sites from the SKYNET Sun photometer network ((Nakajima et al., 2007); Takamura & Nakajima, 2004) for independent validation. Detailed data selection procedure is described in section 2.5.

2.4. Multisensor Satellite Data Sets

To construct an ensemble with a relatively large size and spread, we use 11 satellite-derived monthly AOD products, including four products from the following European Space Agency's Climate Change Initiative (ESA-CCI) products: Advanced Along-Track Scanning Radiometer (AATSR) Dual View (Veefkind et al., 1998), AATSR Swansea University (North, 2002), AATSR-Oxford-RAL Retrieval of Aerosol and Cloud (Sayer et al., 2010), and AATSR-ENSEMBLE (Holzer-Popp et al., 2013) products, which cover the period from 2002 to 2012, and the AVHRR (2006–2011, (Sayer et al., 2017)), MISR (2000–2017, (Martonchik et al., 2002)), Terra MODIS (2000–2017, (Levy et al., 2013)), Aqua MODIS (2002–2017, (Levy et al., 2013)), Polarization and Directionality of the Earth's Reflectances (POLDER, 2005–2013, (Dubovik et al., 2011)), Sea-Viewing Wide Field-of-View Sensor (SeaWiFS, 1997–2010, (Hsu et al., 2012)), and Visible Infrared Imaging Radiometer Suite (2012–2017, (Hsu et al., 2019)) products. The number ranges in parentheses are the period of data used. These are the same 11 products used in the data intercomparison work by (Wei, Li, et al., 2019). Detailed descriptions of the data sets can be found in that paper so are skipped here. All data sets except MISR and SeaWiFS are provided at 1° × 1° resolutions. The 0.5° × 0.5° MISR and SeaWiFS data sets were regridded to 1° × 1° and finally used are those 1° × 1° grid boxes constructed from more than two 0.5° × 0.5° grids. All AODs are provided at 550 nm, except for POLDER which reports AOD at 565 nm.

Since the data ensemble can be viewed as a sample of data obtained at each 1° × 1° grid box, having more samples in each grid box is desirable. In the monthly mean data sets, missing data are frequently observed over regions such as higher latitudes, bright surfaces, and where marine stratocumulus cloud decks are present. The data availability can also differ among the different products. Therefore, for better spatial coverage, we only keep monthly mean AODs that cover more than two thirds of the globe. This reduces the ensemble size to 474, but it is still much larger than using a single data set. We further remove multiyear averaged monthly means from each respective data set and form an ensemble of monthly anomalies that represents the variability at each location. As EnKF assumes that the samples at each location follow a normal distribution, we use a one-sample Kolmogorov-Smirnov test to test this assumption and find that almost all grid cells from 70°S to 70°N meet this requirement (Figure S1 in the supporting information).

Moreover, because the purpose of the ensemble is to examine the spatial representativeness, which exhibits distinct seasonal patterns (Li et al., 2016), we also construct an ensemble for each season with ~120 members. These four seasonal ensembles also meet the normal distribution requirement well (Figure S2). Section 3 compares the effects of using global and seasonal ensembles.

2.5. AERONET Ground Measurements

The ground-based AOD observations are from the Version 3 Level 2.0 AERONET products. This is currently the most extensive global surface aerosol remote sensing network, with more than 800 sites covering most of the world's continental areas and major islands. Data availability and spatial coverage are the two most important factors to consider when selecting AERONET sites. To ensure near-global spatial coverage, we divide the globe into 13 regions/groups based on site density: Western North America, Eastern North America, South America, Europe, North Africa, the Sahel, South Africa, Central Asia, India, East Asia, Southeast Asia, Oceania, and island sites. To ensure data availability, for each region, we count the number of sites with at least 60 monthly mean AOD observations during the 2000 to 2017 period and select at most 20 sites from each region. This results in a total of 135 sites: 15 sites in Western North America, 20 sites in Eastern North America, 11 sites in South America, 20 sites in Europe, 7 sites in North Africa, 6 sites in the Sahel, 2 sites in South Africa, 6 sites in Central Asia, 7 sites in India, 19 sites in East Asia, 7 sites in Southeast Asia, 4 sites in Australia, and 11 sites in Oceania. Additionally, we also selected 54 sites for result validation, including 5 sites in Western North America, 7 sites in Eastern North America, 5 sites in South America, 7 sites in Europe, 4 sites in North Africa, 2 sites in the Sahel, 2 sites in South Africa, 5 sites in Central Asia, 5 sites in East Asia, 4 sites in Southeast Asia, 4 sites in Australia, and 4 sites in Oceania. Note that the data availability of the validation set is lower than the assimilation set but still have at least 36 monthly mean AOD observations. SKYNET is another ground-based Sun photometer network that provides AOD retrievals, which mainly covers Asia and Europe. We selected eight SKYNET sites that have least 60 monhtly mean AODs from 2000 to 2017 for result validation, namely, Chiba, Etchujima, Fukue, Hedo, Miyako, Saga, Phimai, and Lauder. The first six are in Japan, Phimai is in Thailand, and Lauder is a New Zealand site. The spatial range of the regions and the distribution of the assimilation sites can be found in Figures 4–7, and the distirbuiton of the validation sites can be seen in Figures 8g–8i. AERONET AODs are interpolated to 550 nm using measurements at 440, 675, 870, and 1,020 nm through a second-order polynomial fit on a logarithmic scale (Eck et al., 1999).

3. Results

In this study, we use MODIS Collection 6.1 monthly mean AODs at 550 nm from the Terra platform (Levy et al., 2013) as the background AOD field and assimilate observations from the selected AERONET sites using the EnKF approach. The reason of choosing MODIS is that these data compare the best against surface observations (Wei, Li, et al., 2019). We have compared the results of assimilating AERONET data into each of the 11 satellite data sets individually and assimilating all 11 satellite data sets and AERONET using the reported error of each satellite data set and found that assimilating AERONET into MODIS has the best overall performance, that is, the MODIS-AERONET merged data set agrees best with AERONET. The data used here are the dark target-deep blue combined product, spanning the period from February 2000 to December 2017.

After selecting the satellite- and ground-based observations to be used, the EnKF method is ready for implementation. In short, we merge MODIS AOD with AERONET AOD using the EnKF method described in section 2.1. The background error covariance is constructed using multisensor AOD anomalies that capture the spatial correlation patterns, and the errors of surface observations are estimated as the measurement error plus the representation error estimated according to section 2.1.

3.1. Example Case—South America

Before showing the final results, we first examine a representative case of South America, to demonstrate how our method works and to show the expected effect. We assimilate observations at Alta Floresta to see the changes in the absolute bias, RMSE, and correlation coefficient at this site, as well as at nearby sites. All statistical parameters are calculated against AERONET.

Figure 1 shows the comparison of the merged field and the original field against AERONET for the Alta Floresta site. Although overall, the original MODIS data (blue line in Figure 1a) agree reasonably well with AERONET retrievals at this site, MODIS tends to overestimate AOD at a few peaks in 2008, 2011, 2012, and 2013 and to underestimate AOD in 2002, 2005, 2009, 2014, and 2015. By contrast, these biases are largely corrected in the merged data set (red line in Figure 1a). Figure 1b also shows that the merged data set and

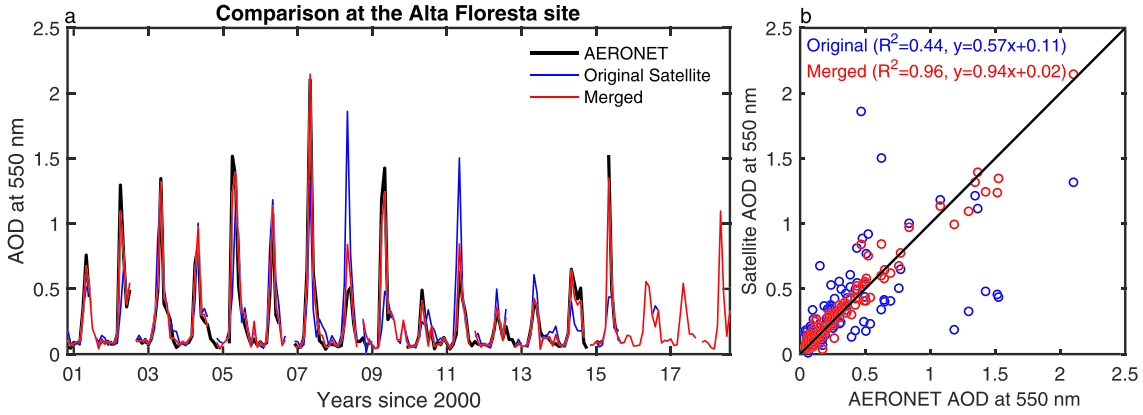

Figure 1. Example of data synergy at the Alta Floresta site located in South America. Panel a shows the time series of monthly mean AODs, and panel b shows the scatter plot of original and merged data against AERONET. It is clearly seen that the merged data set yields better agreement with AERONET.

AERONET retrievals agree better. However, this result is not surprising, as the merged data are weighted means of the original satellite data and AERONET retrievals, as per equation (3).

What is more important is a similar improvement at nearby sites whose observations were not assimilated. This is also the major benefit of the EnKF approach, that is, extending the spatial extent of single-point observations. We first examine the spatial representativeness of the Alta Floresta site (Figure 2a). The representativeness is determined using the method described by (Li et al., 2016). The representativeness of this site is fairly large, spanning almost the entire Amazon Basin, and to its south. This result is reasonable because the aerosol type and its seasonality over the region are relatively uniform. The aerosol loading usually peaks during the biomass burning season from August to October and remains low during other times of the year. Figures 2b–2d show differences in absolute bias, RMSE, and the correlation coefficient between the merged and original data sets, respectively, expressed as the relative change in percentage at all selected sites in the Amazon region. Clearly, as much as 50% bias and RMSE reductions, as well as correlation increases, can be observed at Alta Floresta. Furthermore, the biases and RMSEs (correlations) at five other sites to the west and south of Alta Floresta also decreased (increased), although the amplitude of the change is lower than that at Alta Floresta. The only exception is the site on the West Coast. However, this is expected because this site is not located in the area that Alta Floresta can represent (Figure 2a). These results suggest the effectiveness of our approach by confirming that it can reduce uncertainties both locally and regionally, meaning that it can extend the impact of measurements at single surface sites according to their spatial representativeness.

3.2. Evaluation of the Merged Data Set

In this section, we compare the merged data set and the original data set. We also evaluate the use of global or monthly ensembles with or without assimilation of the local measurements.

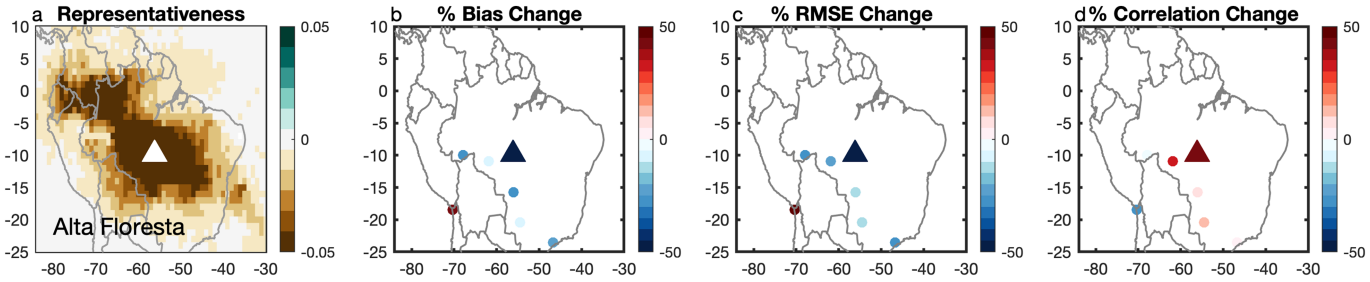

Figure 2. The range of spatial representativeness of the Alta Floresta site (marked by triangle), defined as the reduction of uncertainty (a), the change of absolute bias (b), root-mean-square error (RMSE, c), and correlation coefficient (d), relative to nearby AERONET sites, after assimilating observation at the Alta Floresta site. After assimilating data at Alta Floresta, the performance at many nearby sites also improved.

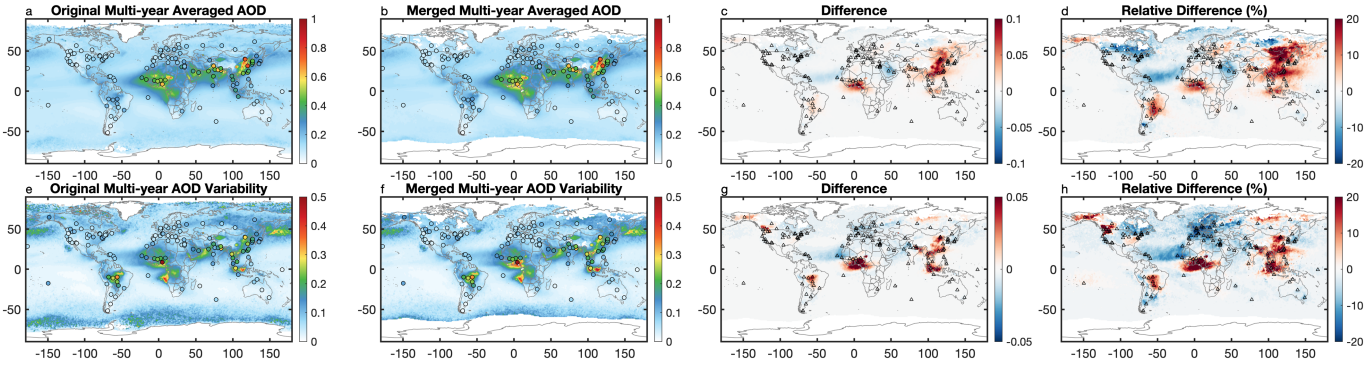

Figure 3. First row: multiyear averaged monthly mean AOD for the original MODIS data (a) and the merged data (b), their differences (merged −original, c), and relative differences (d). Second row: monthly mean AOD variability for the original MODIS data (e) and the merged data (f), their differences (merged − original, g), and relative differences (h). Circles mark the AERONET sites, with interior colors indicating the corresponding AERONET parameter values. The important feature is that the changes are not just confined to the AERONET site but can spread out over a much larger area.

We first focus on results using the global 474-member ensemble with localization. Figure 3 compares the means (panels a–d) and variabilities (panels e–h) of the original and merged data sets to give an overview of the effect of the EnKF technique. The two left-hand columns show the means and standard deviations of the original and merged fields, respectively, and the right-hand column shows their relative differences. The original MODIS data tend to underestimate the mean and standard deviation over places such as the Sahel and East and Southeast Asia. These biases have been largely corrected in the merged data set. The changes in the mean AOD are ~0.1 or 20% over these places, and the changes in the standard deviation are ~0.05. Some differences are still observed in the spatial patterns of mean and standard deviation changes. For example, the largest changes in the mean AOD lie in East Asia, whereas standard deviation change shows strong and extensive signals over Europe. More importantly, the changes shown in Figures 3e, 3d, 3g, and 3h are not only local but can extend far beyond the ground stations, such as the Saharan dust transport over the North Atlantic and the pollution transport from Asia to the Pacific.

A more detailed seasonal analysis is conducted to examine the bias, RMSE, and correlation of the original data set with reference to AERONET retrievals and their changes in the merged data set (Figures 4–6). Figure 4 shows that MODIS has relatively large biases over East Asia, India, and the Sahel during most seasons. In particular, East Asia shows underestimations greater than 0.1. AODs are also underestimated in India except during June-July-August (JJA), when AODs are overestimated in the northern part (bias > 0.1). The Sahel has the strongest underestimation in the spring (March-April-May or MAM) and summer (JJA). A few other places also show seasonally varying biases. For example, Southeast Asia is a major source of biomass burning aerosols (Duncan et al., 2003). Higher negative biases are typically found during the burning peak seasons of MAM and JJA. Biases are generally low in Europe; however, in winter (December-January-February), a negative bias exceeding 0.05 is seen. For the Americas, the biases change signs with season but are small overall. In North America, biases shift from positive in MAM/JJA to negative in September-October-November (SON)/December-January-February. In South America, negative biases are found in JJA and SON, and positive biases appear in the other two seasons. The bias in MODIS retrieval can result from multiple causes, such as surface reflectance parameterization, aerosol model assumption, and cloud contamination. For example, the bias in East Asia and Western United States is also noted by several previous studies (Levy et al., 2010; Tao et al., 2015) and is found to be associated with incorrect surface reflectance parameterization. Problem over India can be related to inappropriate aerosol models used in the retrieval (Misra et al., 2008; Misra et al., 2015). For the Sahel region, MODIS underestimates AOD because the assumed SSA in the coarse mode aerosol model is biased low. The uncertainty over South America is due to the combined effect of incorrect assumptions of aerosol model and aerosol vertical distribution (Levy et al., 2010). In addition, another important reason here is the sampling mismatch between the monthly mean products of satellite and AERONET. Satellite only samples once or twice daily, whereas AERONET has near-continuous measurements. However, because this is not a data evaluation study, we use the straight monthly mean products from each database as this includes more samples and is a closer representation to the true monthly mean AOD.

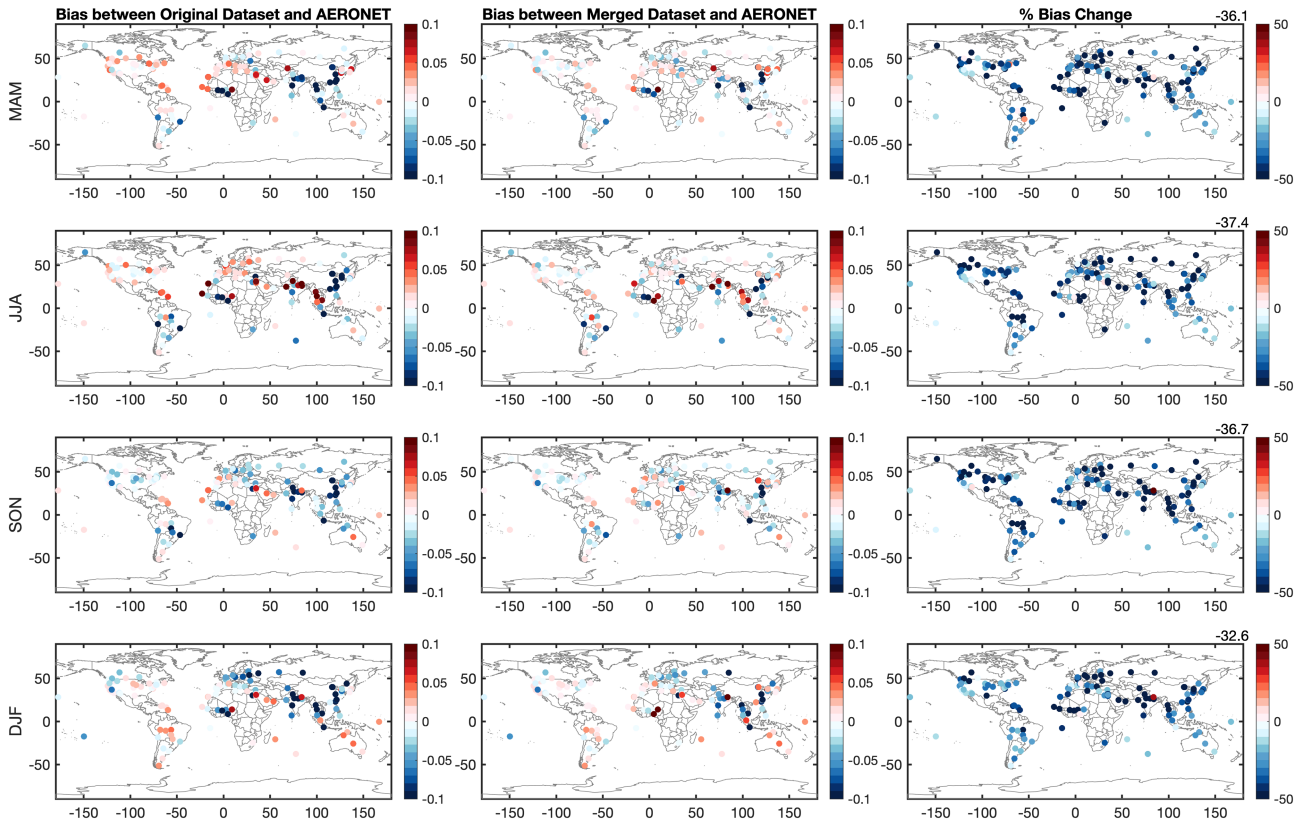

Figure 4. Seasonal bias with respective to AERONET for the original MODIS monthly mean AOD (left column) and the merged AOD (middle column) and their differences expressed as percentage relative to the original bias (right column). The number on the upper right corner of each panel in the right column is the globally averaged value.

Regions with consistently large biases, including East Asia, India, and the Sahel, also appear to have large RMSEs (Figure 5). This is reasonable because a large bias can result in a large RMSE. However, the converse is not true, because positive and negative biases can cancel out. The seasonality in the RMSE pattern is more obvious than that for the bias. For example, large RMSEs are only found in MAM and JJA in the Sahel. RMSEs are only significant in JJA in India. In South America, high RMSEs appear in SON, the peak biomass burning peak season, whereas the bias is less distinct in this season. The canceling of positive and negative biases likely explains why. As seen in the AOD time series for the Alta Floresta site (Figure 1a), MODIS overestimates the September peak in 2008, 2011, 2012, and 2013 but underestimates AODs in 2002, 2005, 2009, 2014, and 2015. The scatter plot comparison (Figure 1b) also shows a few outliers. These underestimations/overestimations appear to result from a combined issue of systematic high bias under high AOD conditions (Levy et al., 2010) and sampling mismatch between MODIS and AERONET monthly mean products. It also highlights the importance of examining the bias and RMSE together to obtain a more accurate comparison.

The correlations between the original MODIS retrievals and AERONET retrievals are low globally in all seasons, especially in regions with high biases and RMSEs, as seen in Figure 6. Eastern North America, North Africa, the Sahel, East Europe, and a few sites in India and East Asia have high correlations (>0.6) in MAM and JJA only. The low seasonal correlation is partly due to the small sample size in each season, especially over high latitudes (Figure S3). Annual correlations are generally much higher (Figure S4). These comparison results show that monthly mean MODIS AOD retrievals deviate significantly from AERONET AOD retrievals.

Merged MODIS-AERONET data set using the EnKF algorithm appears to agree better with AERONET retrievals. In many regions, the biases have been greduced to below 0.05 and even to below 0.02. RMSEs have also decreased to below 0.05 (Figure 5). The correlation coefficients have changed the most, increasing to

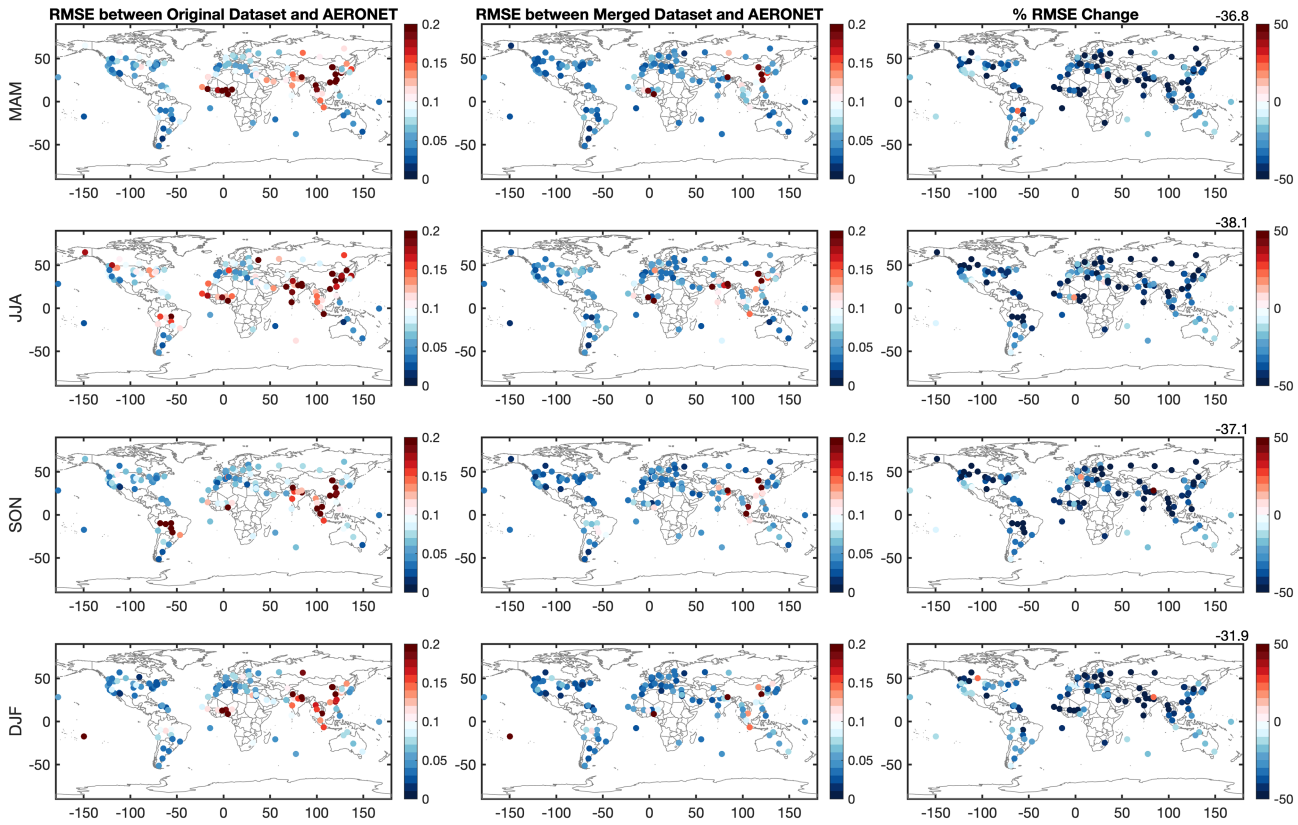

Figure 5. The same as Figure 6 but for the root-mean-square error.

mostly above 0.6 and sometimes above 0.8 (Figure 6). Figures 4–6 also show the changes from the original data sets to the merged data sets expressed as percentages relative to the original data sets. On average, the reduction in biases and RMSEs is by ~35%, and the increase in correlation is by ~500%. The greatest improvements occur in JJA, which is also the season with the highest bias/RMSE and the lowest correlation in most regions. Note that deterioration of the results is still possible at several sites, including Thompson Farm in Northeast United States, Ji Parana Se in Brazil, OHP OBSERVAROIRE in France, and Pokhara in Nepal. Further examination shows that observations at these sites have low spatial representativeness and do not agree with nearby sites (Figure S5). Possible reasons are (1) the spatial variability of AOD surrounding these sites are relatively high or (2) this site may suffer from data quality issues which make its measurements disagree with those at nearby sites. For example, the site may be located close to local emission sources, so that its aerosol properties have strong small-scale features. Nonetheless, this does not happen often and does not affect the global results. Additional data screening and a better site weighting scheme would reduce such problems.

Thus far, we used the data set produced with the global ensemble and covariance localization. Compared next is this data set with data sets using a seasonal ensemble or without localization to determine the optimal setting. Figure 7 shows the percentage change in the bias, RMSE, and correlation coefficient for the global ensemble with localization, the seasonal ensemble with localization, and the global ensemble without localization. Note that for the seasonal ensemble results each season uses a different ensemble, that is, the ensemble constructed for that season as described in section 2.4. Also note that we did not perform seasonal ensemble analysis without localization, because the sample sizes for the seasonal ensembles are considerably smaller (a quarter of the global ensemble size) and covariance localization must be applied because spurious correlations can be prominent. Results arising from using global and seasonal ensembles are similar. Using a global ensemble leads to slightly better results, with greater reductions in biases and RMSEs, and increases in the correlations, implying that on global scale, the size of the ensemble matters more than the spatial pattern of representativeness. By contrast, results arising from with-and-without covariance

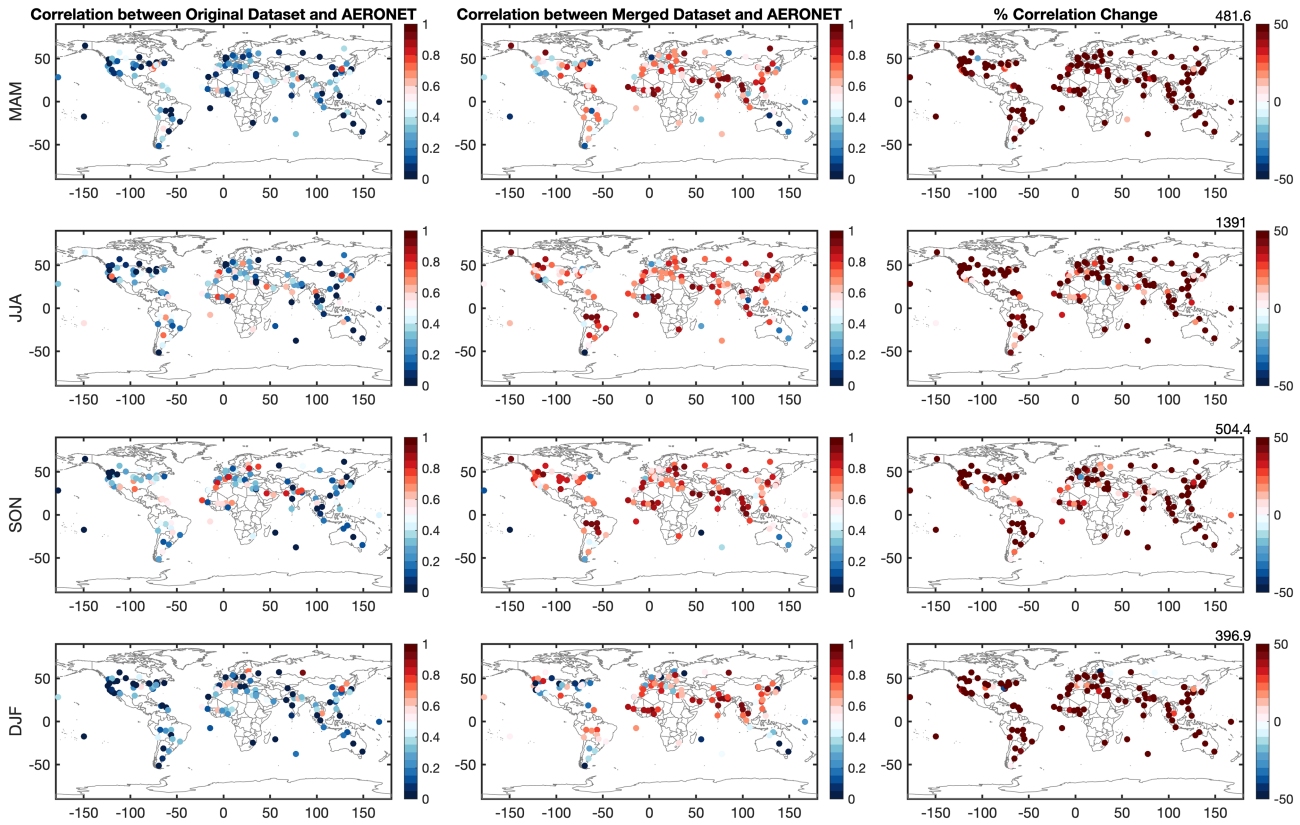

Figure 6. The same as Figure 6 but for the correlation coefficient.

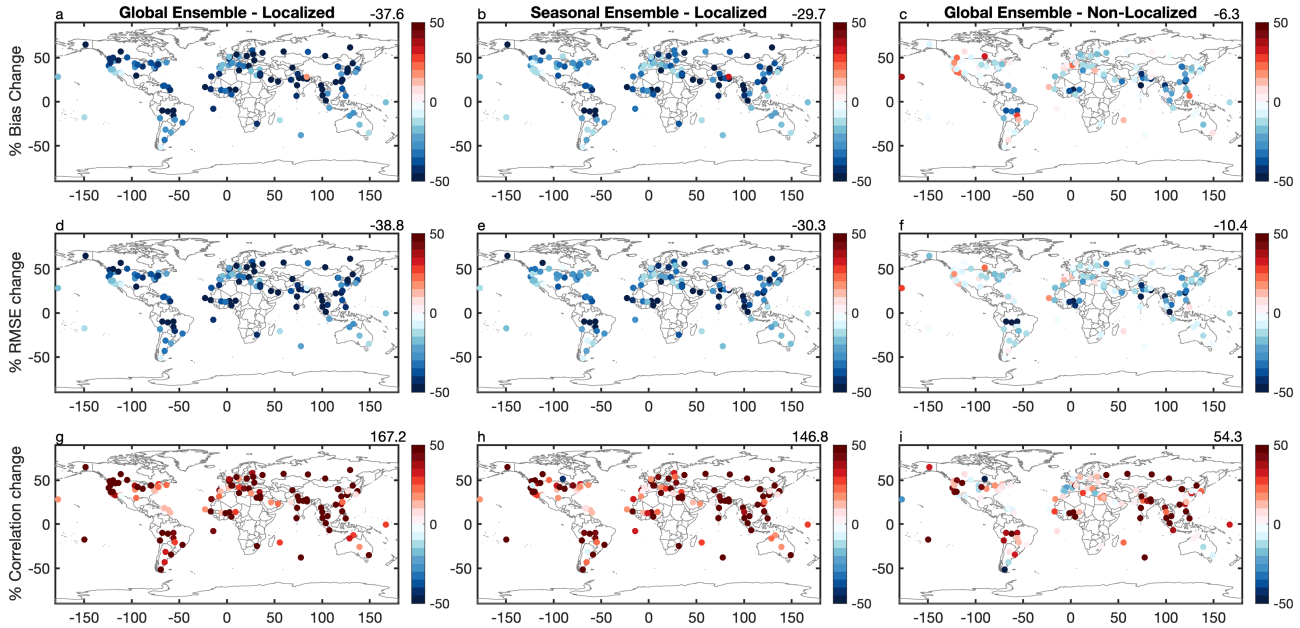

Figure 7. Overall percentage changes in bias (top row), RMSE (middle row), and correlation coefficient (bottom row) for the global ensemble with covariance localization (left column), the seasonal ensemble with covariance localization (middle column), and the global ensemble without covariance localization (right column). The number on the upper right corner of each panel indicates the global mean values.

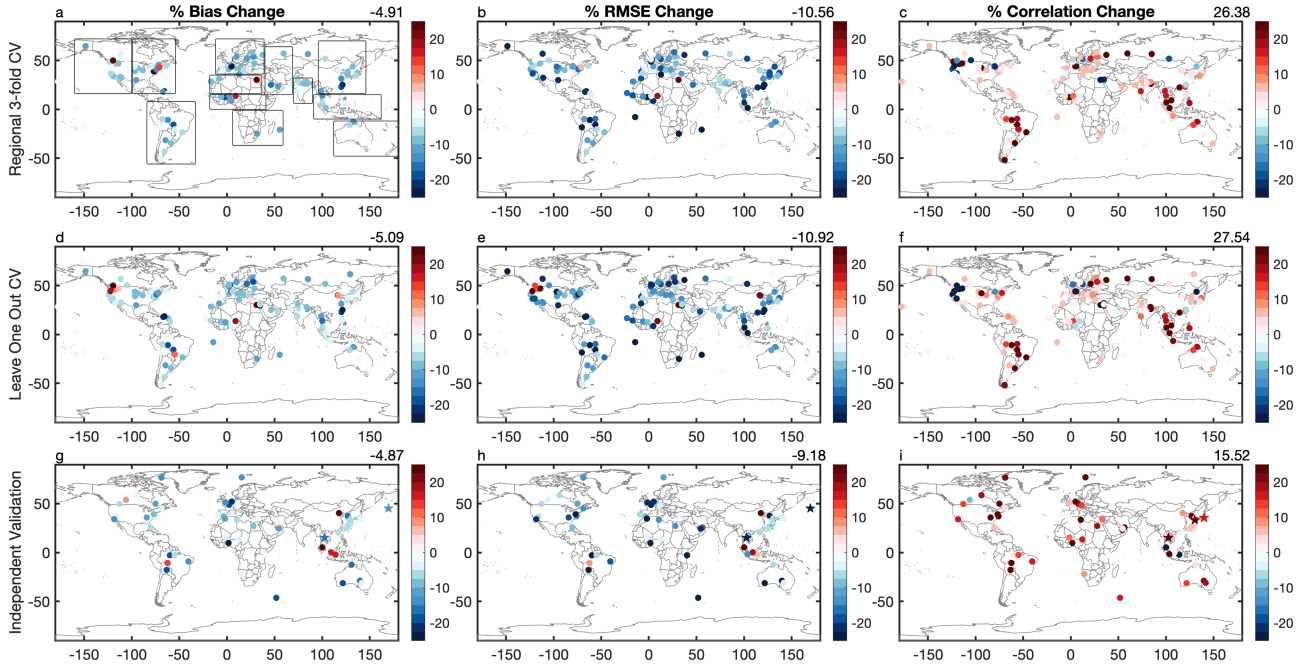

Figure 8. Overall percentage change in bias (a, d, g), RMSE (b, e, h), and correlation coefficient (c, f, i) for the regional threefold cross validation (upper row), leave one out cross validation (middle row), and validation using independent AERONET and SKYNET sites (bottom row). On panels g–i, the stars indicate SKYNET stations. The numbers on the upper right corner of each panel indicate the global mean values.

localization differ greatly. Without covariance localization, the bias and RMSE reductions and the correlation increase are much smaller than those with localization (about a quarter the magnitude). This suggests that spurious correlation is a serious problem here. Although the size of the ensemble is large (474 members), the aerosol loading is also highly variable and spatially inhomogeneous. An even larger sample is likely needed to capture its variability better at each location, as well as the correlation between different locations. The covariance matrix localization scheme helps reduce this problem.

3.3. Validation Results

Improvements in statistical metrics are expected from using the Kalman filter algorithm at sites whose data are assimilated (see section 3.1). It is more important to examine the effect at places where ground observations are not assimilated or not available. Because we do not have accurate information about aerosol properties where no ground observations are available, we need to validate the results with data from existing sites using the procedure described in section 2.3.

Figures 8a–8f show the CV results, with the first row corresponding to the regional threefold scheme and the second row corresponding to the LOO scheme. These two CV schemes produce similar spatial patterns. For the majority of the sites, biases and RMSEs decrease, and correlations increase, but the amplitudes of the changes are lower than those shown in Figure 7. Note that the color bar scale in Figure 8 is reduced compared to Figure 7 to better highlight the spatial patterns. On average, the bias and RMSE reductions and correlation increases of the CV results account for ~15% of the results by assimilating all sites. For some sites, this can reach above 20%. On average, out of the 135 sites, 122 sites show bias decreases, 128 show RMSE decreases, and 110 show correlation increases. The greatest changes are found in South America and Southeast Asia, corresponding well to places with high spatial representativeness ((Li et al., 2016)). The United States and Europe also show moderate changes, mainly due to the high site density there. Because the EnKF technique makes estimates based on spatial representativeness, it is reasonable that sites with higher spatial representativeness have greater impact. For the United States and Europe, although individual sites have lower representativeness than those in South America and Southeast Asia (Li et al., 2016), the site density is higher meaning that a site is very likely affected by nearby sites. However, in the Sahel region where representativeness is high, the CV results are less satisfactory. The reductions in bias and

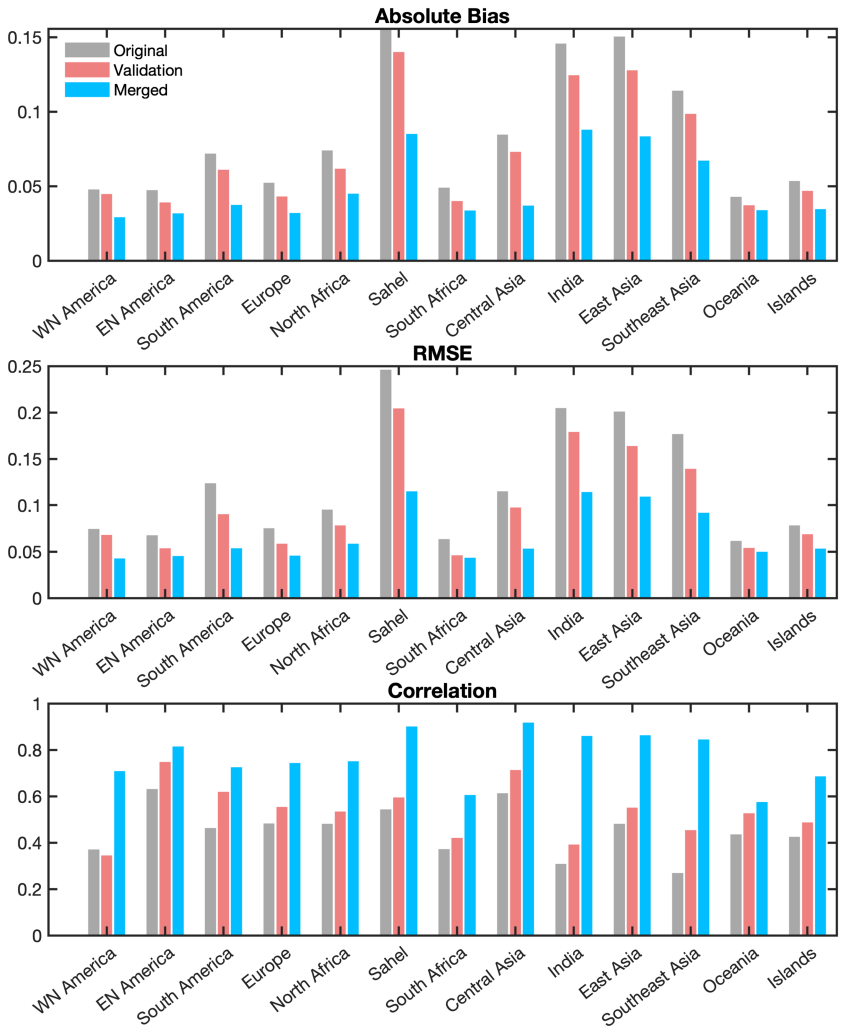

Figure 9. Bar plot comparison between the absolute bias (upper panel), RMSE (middle panel), and correlation coefficient (lower panel) for the original data set, validation data set, and merged data set, all calculated against AERONET.

RMSE here are comparable to those in highly representative regions; the change in the correlation is minimal or even negative. We examined this problem in more detail and found that the correlations between sites in this region are generally low. For example, the Banizoumbou site has a long data record and high spatial representativeness (Li et al., 2016). However, its data are highly correlated with only two other sites (Ouagadougou and Zinder Airport), while correlations with the other sites in this area are all below 0.2. Note that (Li et al., 2016) used satellite data to explore spatial representativeness, whereas here correlations were calculated between surface measurements. The latter can be more affected by local conditions such as emission sources and topography, while for the gridded satellite data, much of the local variability has been smoothed out. Nonetheless, given the overall successful performance of the CV, our proposed method effectively reduces the uncertainty of AOD estimates, at least on scales >~100 km.

The validation results using the independent set (54 AERONET sites and 8 SKYNET sites) are presented in Figures 8g–8i. Overall, the performance for the independent set is comparable to the CV results. The improvments are slightly lower, which we found are mainly due to the worsened results at three Southeast Asian sites. We further looked at the time series of these three sites and found that a common peak in October 2016 with AOD exceeding 2 is primarily responsible for the large deviation. In contrast, measurements at most sites used for data assimilation in this area do not show this peak (either due to lack of record or data quality issues such as cloud contamination).

Finally, for clearer demonstration of the spatial features, we summarize the regional performance of the original, merged, and validation set in Figure 9. It is clear that the merged data set shows significant improvements compared to the original data set, with the greatest changes in absolute bias, RMSE, and correlation coefficients over the Sahel, India, East and Southeast Asia. The validation set also exhibits noticeable improvements over most regions. For Eastern North America, Europe, South America and Oceania, the changes in the statistics are comparable between the merged data set and the validation set. The reason relevant to the former three regions has been discussed above. For Oceania, the bias/RMSE reduction and correlation increments appear low for both the merged and validation set. This is attributed mainly to the sparse site distribution and high AOD variability (relative to the mean) here. Also note that for East Asia and India, where anthropogenic pollution is typically quite high, there is still much room to improve for the validation set. Over these two regions, the various emission sources and complex aerosol composition make aerosol properties highly variable, and the deployment of surface sites seems insufficient. For example, there are very few sites over West China and Central India. This also reveals another benefit of our method, that is, identifying regions for future ground site deployments.

4. Summary and Discussion

In this study, we present a data synergy technique based on the EnKF and apply it to merge monthly mean MODIS and AERONET AOD data sets. The idea is to extend the impact of surface observations to larger areas, according to their spatial representativeness, which can be estimated using satellite measurements, to obtain a global AOD estimate with improved accuracy at the scale of a few hundred kilometers. The key to Kalman filtering is determining the errors associated with the background field and the observations. For the former, we construct a 474-member ensemble using monthly mean AOD anomalies from 11 satellite data sets, so that the AOD variability at each location and its covariability with other locations are better resolved. The latter is the sum of two terms, the AERONET AOD error (equal to its measurement error) and the representation error, which we estimate as the standard deviation of AOD within each $1° \times 1°$ grid box.

As expected, the merged data set agrees much better with AERONET than the original satellite data set. Globally averaged reductions in the bias and RMSE reach ~38%, and the increase in correlation is greater than 150%. Seasonal changes are even greater. We also evaluate the effect of using seasonal ensembles and using a global ensemble without covariance localization and find that the global ensemble with covariance localization yields the best estimates overall. More importantly, regional threefold and LOO CVs further indicate that our approach can reduce uncertainties in places where surface observations are not assimilated. This is because, in many cases, observations at a single site can represent a larger area, which is a major advantage of our data synergy technique. The merged data set likely provides more accurate information about aerosol loading over regions where no ground site retrievals are available.

Because the effect of our data synergy approach essentially depends on the spatial representativeness of ground sites, using data from more sites with greater representativeness is desired. In our study, we did not select sites solely based on this property because data availability is also an important consideration. However, our research does indicate that more consideration should be given to places with higher representativeness when planning the future deployment of surface sites. This can be realized using the method developed by (Li et al., 2017) to identify optimal ground observation locations. The data synergy results will improve with the assimilation of more measurements in highly representative areas. Ideally, we can design a network that would provide global aerosol information using a limited number of sites, at least over land. We also plan to continue refining our results by incorporating additional sites. Moreover, the spatial pattern of representativeness is typically related to physical processes, such as aerosol transport by winds. More detailed examination of such processes will also benefit the selection of sites and the interpretation of the results.

Other improvements include incorporation of more satellite data sets because they all provide independent information about aerosol properties, and the essence of data assimilation is to integrate information from different sources. In the current study, multisensor data are only used to construct the background ensemble for estimating the background error covariance, and only the MODIS data set and ground observations are combined. We have tried to combine several different data sets but find that the major difficulty is estimating

their uncertainties. When we used the same uncertainty for all grid boxes (e.g., the reported measurement accuracy), for some regions the results improve, but globally, they were worse than when using MODIS data alone. This is because on average, MODIS retrievals still agree the best with AERONET. One reason is that the MODIS algorithm uses aerosol models that are constructed using AERONET measurements. Nonetheless, over some regions, other satellite data sets, such as MISR, may perform better. To effectively combine different data sets, we also need to fully assess their performance over different regions and for different aerosol types to obtain a spatially and temporally varying error matrix for each data set. This will be explored in future work.

Finally, our study focuses on 1° × 1° monthly mean data sets and aims to improve AOD estimates over greater area. This is mainly due to the higher spatial coverage and ease of implementation. Therefore, the combined data set is more suitable for large-scale applications, such as validating monthly general circulation model outputs and studying seasonal, interannual, or decadal aerosol variations. There can still be many small-scale features that are not resolved by the merged data set, such as those smaller than the ~100 km scale. Technically, the method can be extended to higher temporal and spatial resolutions. However, to resolve smaller-scale variability, we need more observations in regions with high aerosol variability. This might be difficult as on smaller scales, for example, on daily scales, the sampling is usually much lower for many data sets, so there may not be enough ensembles at some locations. The availability of daily AERONET data is also much lower than on a monthly scale. Plans to carry out a more detailed study to extend our technique to higher spatial and temporal resolutions are currently underway.

Acknowledgments
The ESA-CCI AATSR monthly products are obtained from the ICARE Data and Services Centre (http://www.icare.univ-lille1.fr/cci, last access: 1 October 2018). The MODIS, MISR, AVHRR, and SeaWiFS monthly products are available at https://search.earthdata.nasa.gov/ (last access: 1 October 2018). The POLDER product is available at https://www.grasp-open.com/ products/polder-data-release/ (last access: 1 October 2018). AERONET AOD data are downloaded from the NASA Goddard Space Flight Center (https://aeronet.gsfc.nasa.gov/, last access: 1 October 2018), and the SKYNET AOD data are downloaded from the SKYNET data center (https://www.skynet-isdc.org/aboutSKYNET.php, last access: 26 January, 2020). J. Li is funded by the National Natural Science Foundation of China (NSFC 41975023) and the National Key Research and Development Program of China (2017YFC0212803). X. Li is funded by the National Natural Science Foundation of China (NSFC 41976193 and 41676190). O. Dubovik was supported by the CaPPA Project (Chemical and Physical Properties of the Atmosphere) that is funded by the French National Research Agency (ANR) through the PIA (Programme d'Investissement d'Avenir) under contract "ANR-11-LABX-0005-01" and by the Regional Council "Nord Pas de Calais-Picardie" and the European Funds for Regional Economic Development. Z. Li is supported by the National Science Foundation (AGS1837811). J. Wei by the National Key R&D Program of China (2017YFC1501702).

References

Anderson, J. L. (2001). An Ensemble Adjustment Kalman Filter for data assimilation. *Monthly Weather Review*, *129*, 2884–2903.

Banks, J. R., Brindley, H. E., Flamant, C., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., & Sayer, A. M. (2013). Intercomparison of satellite dust retrieval products over the west African Sahara during the Fennec campaign in June 2011. *Remote Sensing of Environment*, *136*, 99–116.

Bibi, H., Alam, K., Chishtie, F., Bibi, S., Shahid, I., & Blaschke, T. (2015). Intercomparison of MODIS, MISR, OMI, and CALIPSO aerosol optical depth retrievals for four locations on the Indo-Gangetic plains and validation against AERONET data. *Atmospheric Environment*, *111*, 113–126.

Boiyo, R., Kumar, K. R., & Zhao, T. (2017). Statistical intercomparison and validation of multisensory aerosol optical depth retrievals over three AERONET sites in Kenya, East Africa. *Atmospheric Research*, *197*, 277–288.

de Leeuw, G., Holzer-Popp, T., Bevan, S., Davies, W. H., De-scloitres, J., Grainger, R. G., et al. (2015). Evaluation of seven European aerosol optical depth retrieval algorithms for climate analysis. *Remote Sensing of Environment*, *162*, 295–315.

de Leeuw, G., Sogacheva, L., Rodriguez, E., Kourtidis, K., Georgoulias, A. K., Alexandri, G., et al. (2018). Two decades of satellite observations of AOD over mainland China using ATSR-2, AATSR and MODIS/Terra: Data set evaluation and large-scale patterns. *Atmospheric Chemistry and Physics*, *18*, 1573–1592. https://doi.org/10.5194/acp-18-1573-2018

Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., et al. (1998). Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1072–1087.

Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J. L., et al. (2011). Statistically optimized inversion algorithm for enhanced re-trieval of aerosol properties from spectral multi-angle polari-metric satellite observations. *Atmospheric Measurement Techniques*, *4*, 975–1018. https://doi.org/10.5194/amt-4-975-2011

Duncan, B. N., Martin, R. V., Staudt, A. C., Yevich, R., & Logan, J. A. (2003). Interannual and seasonal variability of biomass burning emissions constrained by satellite observations. *Journal of Geophysical Research*, *108*(D2), 4040. https://doi.org/10.1029/2002JD002378

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., et al. (1999). Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *Journal of Geophysical Research*, *104*, 31,333–31,349. https://doi.org/10.1029/1999JD900923

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, *99*(C5), 10,143–10,162.

Fu, D., Xia, X., Wang, J., Zhang, X., Li, X., & Liu, J. (2018). Synergy of AERONET and MODIS AOD products in the estimation of PM 2.5 concentrations in Beijing. *Scientific Reports*, *8*(1), 10174.

Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, *125*, 723–757.

Hamill, T., Whitaker, J. S., & Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, *129*, 2776–2790.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., et al. (1998). AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, *66*(1), 1–16. https://doi.org/10.1016/S0034-4257(98)00031-5

Holzer-Popp, T., de Leeuw, G., Griesfeller, J., Martynenko, D., Klüser, L., Bevan, S., et al. (2013). Aerosol retrieval experiments in the ESA Aerosol_cci project. *Atmospheric Measurement Techniques*, *6*, 1919–1957. https://doi.org/10.5194/amt-6-1919-2013

Houtekamer, P. L., & Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, *129*, 123–137.

Hsu, N. C., Gautam, R., Sayer, A. M., Bettenhausen, C., Li, C., Jeong, M. J., et al. (2012). Global and regionaltrends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010. *Atmospheric Chemistry and Physics*, *12*, 8037–8053. https://doi.org/10.5194/acp-12-8037-2012

Hsu, N. C., Lee, J., Sayer, A. M., Kim, W., Bettenhausen, C., & Tsay, S. C. (2019). VIIRS Deep blue aerosol products over land: Extending the EOS long-term aerosol data records. *Journal of Geophysical Research: Atmospheres*, *124*, 4026–4053. https://doi.org/10.1029/2018JD029688

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T., Smirnov, A., & Holben, B. N. (2010). Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. *Journal of Geophysical Research*, *115*, D23209. https://doi.org/10.1029/2010JD014601

Kahn, R. A., Garay, M. J., Nelson, D. L., Yau, K. K., Bull, M. A., Gaitley, B. J., & Levy, R. C. (2007). Satellite-derived aerosol optical depth over dark water from MISR and MODIS: Comparisons with AERONET and implications for climatological studies. *Journal of Geophysical Research*, *112*, D18205. https://doi.org/10.1029/2006JD008175

Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., & Remer, L. A. (2009). MISR aerosol product attributes and statistical comparisons with MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(12), 4095–4114.

Kalman, R., & Bucy, K. (1961). New results in linear prediction filtering theory. *Transactions of ASME, Series D, Journal of Basic Engineering*, *83D*, 95–108.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., & Hsu, N. C. (2013). The Collection 6 MODIS aerosol products over land and ocean. *Atmospheric Measurement Techniques*, *6*, 2989–3034. https://doi.org/10.5194/amt-6-2989-2013

Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., & Eck, T. F. (2010). Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmospheric Chemistry and Physics*, *10*, 10,399–10,420. https://doi.org/10.5194/acp-10-10399-2010

Li, C., Li, J., Xu, H., Li, Z., Xia, X., & Che, H. (2019). Evaluating VIIRS EPS aerosol optical depth in China: An Intercomparison Against Ground-based Measurements and MODIS. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *224*(2019), 368–377.

Li, J., Carlson, B. E., & Lacis, A. A. (2014a). Application of spectral analysis techniques in the inter-comparison of aerosol data: Part III. Using combined PCA to compare spatio-temporal variability of MODIS, MISR and OMI aerosol optical depth. *Journal of Geophysical Research: Atmospheres*, *119*, 4017–4042. https://doi.org/10.1029/2013JD020538

Li, J., Carlson, B. E., & Lacis, A. A. (2014b). Application of spectral analysis techniques in the inter-comparison of aerosol data. Part 4: Combined Maximum Covariance Analysis to bridge the gap between multi-sensor satellite retrievals and ground-based measurements. *Atmospheric Measurement Techniques*, *7*, 3503–3547. https://doi.org/10.5194/atmd-7-3503-2014

Li, J., Li, X., Carlson, B. E., Kahn, R. A., Lacis, A. A., Dubovik, O., & Nakajima, T. (2016). Reducing multi-sensor satellite monthly mean aerosol optical depth uncertainty: 1. Objective assessment of current AERONET locations. *Journal of Geophysical Research: Atmospheres*, *121*, 13,609–13,627. https://doi.org/10.1002/2016JD025469

Li, J., Li, X., Carlson, B. E., Kahn, R. A., Lacis, A. A., Dubovik, O., & Nakajima, T. (2017). Reducing multisensor monthly mean aerosol optical depth uncertainty: 2. Optimal locations for potential ground observation deployments. *Journal of Geophysical Research: Atmospheres*, *122*, 3920–3928. https://doi.org/10.1002/2016JD026308

Li, Z., Zhao, X., Kahn, R., Mishchenko, M., Remer, L., Lee, K. H., et al. (2009). Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: A review and perspective. *Annales de Geophysique*, *27*(7), 2755–2770.

Lorenc, A. C. (2003). The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Quarterly Journal of the Royal Meteorological Society*, *129*, 3183–3203.

Martonchik, J. V., Diner, D. J., Crean, K. A., & Bull, M. A. (2002). Regional aerosol retrieval results from MISR. *IEEE Transactions on Geoscience and Remote Sensing*, *40*, 1520–1531.

Mishchenko, M. I., Liu, L., Travis, L. D., Cairns, B., & Lacis, A. A. (2010). Toward unified satellite climatology of aerosol properties: 3. MODIS versus MISR versus AERONET. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *111*, 540–552. https://doi.org/10.1016/j.jqsrt.2009.11.003

Misra, A., Jayaraman, A., & Ganguly, D. (2008). Validation of MODIS derived aerosol optical depth over Western India. *Journal of Geophysical Research*, *113*, D04203. https://doi.org/10.1029/2007JD009075

Misra, A., Jayaraman, A., & Ganguly, D. (2015). Validation of version 5.1 MODIS aerosol optical depth (Deep Blue Algorithm and Dark Target Approach) over a semi-arid location in Western India. *Aerosol and Air Quality Research*, *15*(1), 252–262.

Nakajima, T., Yoon, S. C., Ramanathan, V., Shi, G. Y., Takemura, T., Higurashi, A., et al. (2007). Overview of the Atmospheric Brown Cloud East Asian Regional Experiment 2005 and a study of the aerosol direct radiative forcing in east Asia. *Journal of Geophysical Research*, *112*, D24S91. https://doi.org/10.1029/2007JD009009

Nguyen, H., Cressie, N., & Braverman, A. (2012). Spatial statistical data fusion for remote sensing applications. *Journal of the American Statistical Association*, *107*(499), 1004–1018. https://doi.org/10.1080/01621459.2012.694717

Nirala, M. (2008). Multi-sensor data fusion of aerosol optical thickness. *International Journal of Remote Sensing*, *29*(7), 2127–2136.

North, P. R. J. (2002). Estimation of aerosol opacity and land surface bidirectional reflectance from ATSR-2 dual-angle imagery: Operational method and validation. *Journal of Geophysical Research*, *107*(D12), 4149. https://doi.org/10.1029/2000JD000207

Prasad, A. K., & Singh, R. P. (2007). Comparison of MISR-MODIS aerosol optical depth over the Indo-Gangetic basin during the winter and summer seasons (2000–2005). *Remote Sensing of Environment*, *107*(1-2), 109–119.

Rubin, J. I., & Collins, W. D. (2014). Global simulations of aerosol amount and size using MODIS observations assimilated with an Ensemble Kalman Filter. *Journal of Geophysical Research: Atmospheres*, *119*, 12,780–12,806. https://doi.org/10.1002/2014JD021627

Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Holben, B. N., Xian, P., et al. (2017). Assimilation of AERONET and MODIS AOT observations using variational and ensemble data assimilation methods and its impact on aerosol forecasting skill. *Journal of Geophysical Research: Atmospheres*, *122*, 4967–4992. https://doi.org/10.1002/2016JD026067

Salomonson, V. V., Barnes, W., Xiong, J., Kempler, S., & Masuoka, E. (2002). An overview of the Earth Observing System MODIS instrument and associated data systems performance. In IEEE International Geoscience and Remote Sensing Symposium (Vol. 2, pp. 1174-1176). IEEE.

Sayer, A. M., Hsu, N. C., Bettenhausen, C., Jeong, M. J., & Meister, G. (2015). Effect of MODIS Terra radiometric calibration improvements on Collection 6 Deep Blue aerosol products: Validation and Terra/Aqua consistency. *Journal of Geophysical Research: Atmospheres*, *120*, 12,157–12,174. https://doi.org/10.1002/2015JD023878

Sayer, A. M., Hsu, N. C., Lee, J., Carletta, N., Chen, S. H., & Smirnov, A. (2017). Evaluation of NASA Deep Blue/SOAR aerosol retrieval algorithms applied to AVHRR measurements. *Journal of Geophysical Research: Atmospheres*, *122*, 9945–9967. https://doi.org/10.1002/2017JD026934

Sayer, A. M., Thomas, G. E., & Grainger, R. G. (2010). A sea surface reflectance model for AATSR, and application to aerosol retrievals. *Remote Sensing of Environment*, *3*, 813–838. https://doi.org/10.5194/amt-3-813-2010

Takamura, T., & Nakajima, T. (2004). Overview of SKYNET and its activities. *Optica Pura y Aplicada*, *37*, 3303–3308.

Tang, Q., Bo, Y., & Zhu, Y. (2016). Spatiotemporal fusion of multiple-satellite aerosol optical depth (AOD) products using Bayesian maximum entropy method. *Journal of Geophysical Research: Atmospheres*, *121*, 4034–4048. https://doi.org/10.1002/2015JD024571

Tao, M., Chen, L., Wang, Z., Tao, J., Che, H., Wang, X., & Wang, Y. (2015). Comparison and evaluation of the MODIS Collection 6 aerosol data in China. *Journal of Geophysical Research: Atmospheres*, *120*, 6992–7005. https://doi.org/10.1002/2015JD023360

Veefkind, J. P., de Leeuw, G., & Durkee, P. A. (1998). Retrieval of aerosol optical depth over land using two-angle view satellite radiometry during TARFOX. *Geophysical Research Letters*, *25*(3135–3138), 1998.

Wang, J., Brown, D. G., & Hammerling, D. (2013). Geostatistical inverse modeling for super-resolution mapping of continuous spatial processes. *Remote Sensing of Environment*, *139*, 205–215. https://doi.org/10.1016/j.rse.2013.08.007

Wang, W., Mao, F., Pan, Z., Du, L., & Gong, W. (2017). Validation of VIIRS AOD through a Comparison with a Sun Photometer and MODIS AODs over Wuhan. *Remote Sensing*, *9*(5), 403.

Wei, J., Li, Z., Peng, Y., & Sun, L. (2019). MODIS collection 6.1 aerosol optical depth products over land and ocean: Validation and comparison. *Atmospheric Environment*, *201*, 428–440.

Wei, J., Peng, Y., Mahmood, R., Sun, L., & Guo, J. (2019). Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products. *Atmospheric Chemistry and Physics*, *19*, 7183–7207.

Wei, J., Sun, L., Huang, B., Bilal, M., Zhang, Z., & Wang, L. (2018). Verification, improvement and application of aerosol optical depths in China. Part 1: Inter-comparison of NPP-VIIRS and Aqua-MODIS. *Atmospheric Environment*, *175*, 221–233.

Xu, H., Guang, J., Xue, Y., de Leeuw, G., Che, Y. H., Guo, J., et al. (2015). A consistent aerosol optical depth (AOD) dataset over mainland China by integration of several AOD products. *Atmospheric Environment*, *114*, 48–56. https://doi.org/10.1016/j.atmosenv.2015.05.023

Yumimoto, K., & Takemura, T. (2011). Direct radiative effect of aerosols estimated using ensemble-based data assimilation in a global aerosol climate model. *Geophysical Research Letters*, *38*, L21802. https://doi.org/10.1029/2011GL049258