



HAL
open science

Validation of SOAR VIIRS Over-Water Aerosol Retrievals and Context Within the Global Satellite Aerosol Data Record

Andrew M. Sayer, N. Â. Christina Hsu, Jaehwa Lee, Woogyung V. Kim, Oleg Dubovik, Steven T. Dutcher, Dong Huang, Pavel Litvinov, Alexei Lyapustin, Jason L. Tackett, et al.

► To cite this version:

Andrew M. Sayer, N. Â. Christina Hsu, Jaehwa Lee, Woogyung V. Kim, Oleg Dubovik, et al.. Validation of SOAR VIIRS Over-Water Aerosol Retrievals and Context Within the Global Satellite Aerosol Data Record. *Journal of Geophysical Research: Atmospheres*, 2018, 123, pp.13,496-13,526. 10.1029/2018JD029465 . insu-03686250

HAL Id: insu-03686250

<https://insu.hal.science/insu-03686250>

Submitted on 3 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

RESEARCH ARTICLE

10.1029/2018JD029465

Key Points:

- Based on comparisons to AERONET, the AOD uncertainty is consistent with expectations
- Retrieved parameters related to aerosol particle size and type are also demonstrated to have skill
- SOAR shows similar regional and temporal patterns to other satellite data sets

Correspondence to:

A. M. Sayer,
andrew.sayer@nasa.gov

Citation:

Sayer, A. M., Hsu, N. C., Lee, J., Kim, W. V., Dubovik, O., Dutcher, S. T. et al. (2018). Validation of SOAR VIIRS over-water aerosol retrievals and context within the global satellite aerosol data record. *Journal of Geophysical Research: Atmospheres*, 123, 13,496–13,526. <https://doi.org/10.1029/2018JD029465>

Received 9 AUG 2018

Accepted 13 NOV 2018

Accepted article online 20 NOV 2018

Published online 13 DEC 2018

Author Contributions

Conceptualization: Andrew M. Sayer, N. Christina Hsu

Data curation:

Oleg Dubovik, Steven T. Dutcher, Dong Huang, Pavel Litvinov, Alexei Lyapustin, Jason L. Tackett, David M. Winker

Funding Acquisition: N. Christina Hsu

Methodology: Andrew M. Sayer

Validation: Andrew M. Sayer

Writing - Original Draft: Andrew M. Sayer

Formal Analysis: Andrew M. Sayer

Investigation: Andrew M. Sayer, Jason L. Tackett

Resources: N. Christina Hsu, Oleg Dubovik, Steven T.

Dutcher, Dong Huang, Pavel Litvinov, Alexei Lyapustin, Jason L. Tackett, David M. Winker






Supervision: N. Christina Hsu

Visualization: Andrew M. Sayer

Writing - review & editing: Andrew M. Sayer, N. Christina Hsu, Jaehwa Lee, Oleg Dubovik, Steven T. Dutcher, Dong Huang, Pavel Litvinov, Alexei Lyapustin, Jason L. Tackett, David M. Winker

©2018. American Geophysical Union.
All Rights Reserved.

Validation of SOAR VIIRS Over-Water Aerosol Retrievals and Context Within the Global Satellite Aerosol Data Record

Andrew M. Sayer^{1,2} , N. Christina Hsu² , Jaehwa Lee^{2,3} , Woogyung V. Kim^{2,3} , Oleg Dubovik⁴, Steven T. Dutcher⁵, Dong Huang^{2,6}, Pavel Litvinov⁴, Alexei Lyapustin² , Jason L. Tackett^{6,7}, and David M. Winker⁷ 

¹Goddard Earth Sciences Technology and Research, Universities Space Research Association, Greenbelt, MD, USA, ²NASA Goddard Space Flight Center, Greenbelt, MD, USA, ³Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, ⁴Laboratoire d'Optique Atmosphérique, Université de Lille, Villeneuve d'Ascq, France, ⁵Atmosphere SIPS, Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI, USA, ⁶Science Systems and Applications, Inc, Glenn Dale, MD, USA, ⁷NASA Langley Research Center, Hampton, VA, USA

Abstract This study validates aerosol properties retrieved using a Satellite Ocean Aerosol Retrieval (SOAR) algorithm applied to Visible Infrared Imaging Radiometer Suite (VIIRS) measurements, from Version 1 of the VIIRS Deep Blue data set. SOAR is the over-water complement to the over-land Deep Blue algorithm and has two processing paths: globally, 95% of pixels are processed with the full retrieval algorithm, while the 5% of pixels in shallow or turbid (mostly coastal) waters are processed with a backup algorithm. Aerosol Robotic Network (AERONET) data are used to validate and compare the midvisible (550 nm) aerosol optical depth (AOD), Ångström exponent (AE), and fine mode fraction of AOD at 550 nm (FMF). AOD uncertainty is shown to be approximately $\pm(0.03 + 10\%)$ for the full and $\pm(0.03 + 15\%)$ for the backup algorithms, with a small positive median bias around 0.02. When AOD is below about 0.2, the AE and FMF have small negative offsets from AERONET around -0.15 and -0.04 , respectively. For higher AOD, AE is less offset and the magnitudes of differences versus AERONET are about ± 0.2 and ± 0.14 , respectively. Aerosol-type classifications provided by SOAR are found to be reasonable, matching optical-based classifications from AERONET over 80% of the time. Spatial and temporal patterns of AOD and AE are also compared with those of other contemporary over-water satellite aerosol data sets; dependent on region, the satellite data sets show varying levels of consistency, with SOAR broadly in-family, and the largest discrepancies in regions with persistent heavy cloud cover.

Plain Language Summary Aerosols are small particles in the atmosphere like desert dust, volcanic ash, smoke, industrial haze, and sea spray. Understanding them is important for applications such as hazard avoidance, air quality and human health, and climate studies. Satellite instruments provide an important tool to study aerosol loading over the world. However, individual satellites do not last forever and newer satellites often have improved capabilities compared to older ones. This paper evaluates an extension of an algorithm, originally designed to monitor aerosols from an older satellite instrument, to a new satellite instrument called Visible Infrared Imaging Radiometer Suite. The evaluation is performed by comparing to ground truth data which are part of the National Aeronautics and Space Administration's global Aerosol Robotic Network, as well as to other satellite-based aerosol data sets from different spaceborne instruments.

1. Introduction

The Visible Infrared Imaging Radiometer Suite (VIIRS) was launched in late 2011 on the Suomi-National Polar-orbiting Partnership (S-NPP) satellite, alongside several other sensors, and remains in operation at the time of writing. S-NPP is the precursor to a new series of U.S. polar-orbiting satellites intended to continue observations relevant to forecasting and scientific research from previous satellite programmes, and its instrument suite was designed with these goals in mind. As a result, the VIIRS instrument has similarities to, and is being applied to carry on geophysical data records from, heritage sensor series such as the Moderate Resolution Imaging Spectroradiometers (MODIS), Sea-viewing Wide Field-of-view Sensor (SeaWiFS), and Advanced Very High Resolution Radiometers (AVHRRs).

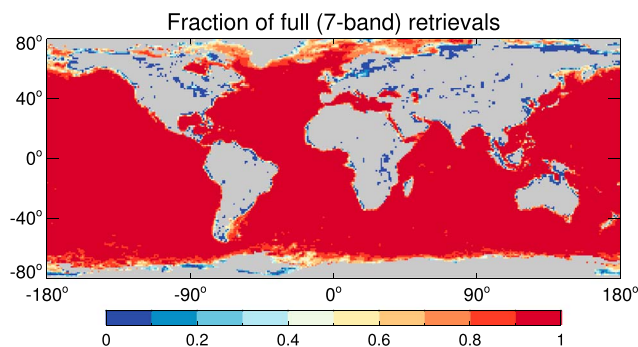


Figure 1. Fraction of SOAR VIIRS Suite retrievals from the full seven-band retrieval algorithm. Data shown for the year 2013; grid cells with fewer than 100 retrievals are shaded in gray. Continental outlines are omitted to more easily visualize behavior near coasts.

One such application is the study of atmospheric aerosols, the main remotely sensed geophysical quantity being total column midvisible aerosol optical depth (AOD). Aerosol remote sensing is a difficult problem, and many AOD retrieval algorithms have been developed for different sensors for either regional or global applications. Some reviews of basic algorithm concepts which have been applied to various sensors types are provided by Kokhanovsky and de Leeuw (2009) and Lenoble et al. (2013). This study is concerned with the Satellite Ocean Aerosol Retrieval (SOAR) algorithm, which was developed initially for application to SeaWiFS (Sayer et al., 2012) but has since been applied to AVHRR (Hsu et al., 2017) and VIIRS (Sayer, Hsu, Bettenhausen, et al., 2017; Sayer et al., 2018) measurements. The motivation behind SOAR was to provide an overwater companion to the over-land Deep Blue (DB) algorithm, which has been applied to SeaWiFS, MODIS, AVHRR, and VIIRS measurements (Hsu et al., 2013, 2017), the combination together providing near-global data coverage (for daytime scenes free from clouds and snow/ice). Recently, Sayer et al. (2018) described updates and modifications to SOAR to create the recently released VIIRS DB/SOAR version 1 data set; the purpose of the present study is to provide an evaluation of the over-water part of this new data set, which is freely available for download from the National Aeronautics and Space Administration Level 1 and Atmosphere Archive and Distribution System at <https://ladsweb.nascom.nasa.gov>.

Section 2 provides a summary of the VIIRS application of the SOAR algorithm and the resulting data set. This is validated against ground truth data from the Aerosol Robotic Network (AERONET, Holben et al., 1998) in section 3. Next, section 4 illustrates and evaluates the aerosol-type information provided by SOAR, and finally, section 5 examines the VIIRS time series in the context of some of the main other over-water satellite data products available in the modern period.

Section 2 provides a summary of the VIIRS application of the SOAR algorithm and the resulting data set. This is validated against ground truth data from the Aerosol Robotic Network (AERONET, Holben et al., 1998) in section 3. Next, section 4 illustrates and evaluates the aerosol-type information provided by SOAR, and finally, section 5 examines the VIIRS time series in the context of some of the main other over-water satellite data products available in the modern period.

2. Summary of SOAR VIIRS Data Set

A comprehensive description of the SOAR algorithm applied to VIIRS, along with relevant features of the VIIRS instrument, was provided by Sayer et al. (2018). As a result, this section omits algorithm description and focuses on the features of the data set itself. That study describes the algorithm used to produce Version 1 of the VIIRS DB/SOAR data set, which is the data version analyzed in this study. References to SOAR without a specified sensor hereafter refer to the VIIRS application; likewise, references to AOD without a specific wavelength indicate the AOD at 550 nm.

SOAR performs a multispectral fit of VIIRS top of atmosphere (TOA) measured reflectance to simultaneously retrieve the AOD, fine mode fractional contribution to AOD at 550 nm (FMF), and best fitting aerosol optical model. Together, these may be used to self-consistently derive other quantities such as spectral AOD or the Ångström exponent (AE). There are four candidate aerosol optical models with optical properties representative of aerosol columns dominated by clean maritime, dust, fine mode (e.g., smoke and haze), and mixed (e.g., dust plus smoke) particles. These models all assume a bimodal lognormal distribution with fine and coarse modes. Before the retrieval, the calibration corrections derived in Sayer, Hsu, Bettenhausen, et al. (2017) are applied to the Level 1 (L1) TOA reflectances, which brings the sensor into radiometric consistency with MODIS aboard the Aqua satellite, aside from the L1 differences expected due to different band spectral response functions.

Over open ocean, a total of seven bands are used in the inversion; over waters identified as shallow or turbid, only four near infrared and shortwave infrared (swIR) bands are used, because at the shorter visible wavelengths the surface reflectance model assumed may be in error in these cases. This four-band retrieval is often referred to as the *backup* algorithm, for distinction from the seven-band *full* algorithm. As these two paths are expected to have different error characteristics, they are considered separately herein. Globally, about 95% of pixels are processed with the seven-band full algorithm, with backup retrievals concentrated around coasts, inland waters, and in continental outflow regions (Figure 1).

The retrieval is performed on suitable (free from detected cloud, sea ice, or strong Sun glint) pixels identified as being water surfaces, at native sensor resolution (~ 740 m horizontal pixel size at nadir). The median value

Table 1
Aerosol SDS From VIIRS Data Products Used in This Study

Product level	SDS name	Description
L2	Aerosol_Optical_Thickness_550_Ocean_Best_Estimate	AOD at 550 nm, QA filtered
L2	Spectral_Aerosol_Optical_Thickness_Ocean	Spectral AOD (490–2,250 nm)
L2	Angstrom_Exponent_Ocean_Best_Estimate	AE (550–870 nm), QA filtered
L2	Fine_Mode_Fraction_550_Ocean	Fine mode fractional contribution to AOD at 550 nm
L2	Aerosol_Optical_Thickness_QA_Flag_Ocean	Indicates poor (1) or good (3) data
L2	Algorithm_Flag_Ocean	Indicates full (0), turbid/shallow (1), or mixed (2) retrieval
M3	Aerosol_Optical_Thickness_550_Ocean_Mean	Monthly mean AOD at 550 nm
M3	Angstrom_Exponent_Ocean_Mean	Monthly mean AE (550–870 nm)
M3	Aerosol_Type_Land_Ocean_Histogram	Histogram of grid cell daily modal aerosol types over a month

of these retrievals across an 8×8 sensor pixel box (~ 6 km at nadir) is taken as representative and reported in the Level 2 (L2) data product at this resolution, known as a *cell* or *retrieval pixel*. Some cells contain retrievals of both types (full and backup), and these mixed cells are considered to also be backup retrievals in this study. Statistics of the individual retrievals within each cell are used to assign each cell a quality assurance (QA) flag value, as an indicator of whether quality issues with a given retrieval are expected. QA flags take values of 0 (no retrieval), 1 (potential poor quality), or 3 (good quality, no issues suspected); this 0–3 rating follows prior DB/SOAR convention, although note that there is no intermediate QA = 2 subset of the data. Again, specifics of all these thresholds and algorithm decisions are provided by Sayer et al. (2018). Generally, data with QA = 1 are not intended for use, and only QA = 3 data are used in the present analysis.

S-NPP is in a Sun-synchronous polar orbit with a daytime node equatorial crossing local solar time of 1:30 p.m., similar to the A-Train satellite constellation, although on a separate orbital track. L2 products consist of 6 min of scan data in the along-track direction; the across-track scan width is 3,040 km meaning that consecutive orbits overlap, even at the equator. Level 3 (L3) products are also generated as statistical summaries on a 1° horizontal grid. L3 daily (D3) files are created by aggregating L2 data (using only data passing QA thresholds), while L3 monthly (M3) files are aggregates of D3 products. These files are provided in Network Common Data Format version 4 format and the scientific data sets (SDS) within conform to Climate and Forecast data conventions version 1.6. The aerosol-related SDS from these files used within this study are listed in Table 1. Note that these SDS names refer to aerosol optical thickness (AOT) rather than AOD, which was a project requirement for compliance with heritage data sets. AOT and AOD are synonymous, and both are defined as the vertically integrated extinction due to aerosol scattering and absorption. Data products are available covering March 2012 onward.

3. Validation Against AERONET

Most satellite AOD validation exercises use some variation on the basic technique of Ichoku et al. (2002). In brief, this method uses spatial averaging of satellite data and temporal averaging of AERONET data to account for the fact that AERONET provides point measurements of AOD with a sampling frequency of 5–15 min (dependent on instrument configuration) in cloud-free conditions, while satellites provide an instantaneous swath measurement with individual retrieval footprints covering several to tens of kilometers. This fundamental difference introduces some uncertainty beyond that of AERONET alone, dependent on the level of heterogeneity in the underlying aerosol field (Virtanen et al., 2018). Since the publication of prior SOAR validation papers (Sayer, Hsu, et al., 2012; Sayer, Hsu, Bettenhausen, et al., 2017; Sayer, Hsu, Lee, et al., 2017), updates have been made to both the AERONET data version and the specifics of the validation methodology, and as a result these are described in full below.

In Sayer et al. (2018), a preliminary validation of the SOAR VIIRS data set was provided by comparing with ship-based AOD observations from the Maritime Aerosol Network (MAN, Smirnov et al., 2009). MAN is a valuable resource for open ocean validation but is somewhat limited in spatial and temporal coverage, and Sayer

et al. (2018) obtained 836 matches from approximately 5 years of VIIRS data. An advantage of AERONET over MAN is a larger data volume, but a disadvantage is the restricted sampling to coastal and island regions. Thus, the two are complementary, and results from the MAN comparison are discussed in the text to provide additional context.

3.1. AERONET Data Description and Matchup Protocol

The Sun photometers used in AERONET infer spectral AOD with an uncertainty of ~ 0.01 in the midvisible (Eck et al., 1999) by direct Sun observations of solar irradiance, which is attenuated by Rayleigh scattering, trace gas absorption, and aerosols as it passed through the atmosphere. AERONET provides data from several hundred sites, processed in a consistent way (Holben et al., 1998); this consistency is important when evaluating data from multiple sites.

This study uses the AERONET direct Sun level 2 (cloud-screened, post-deployment calibrated, and quality-assured; Smirnov et al., 2000) data products, from the newly released version 3. Version 3 includes improvements to sensor characterization and cloud aerosol discrimination over version 2 (T. Eck, D. Giles, A. Smirnov, personal communication, July 2018), particularly in the detection of stable optically thin cirrus cloud layers and rapidly evolving fine mode aerosol plumes. The geolocation of some sites has also been updated with more accurate site latitudes, longitudes, and elevations. Here VIIRS and AERONET data from all sites between March 2012 and December 2017 are considered; AERONET level 2 data have a latency of several months to years due to the requirement for postdeployment calibration, and so data from more recent months are at present largely unavailable.

All instruments provide data at a standard set of wavelengths (440, 675, 870, and 1,020 nm for AOD), and some include additional wavelengths. In this analysis, AERONET AOD are interpolated spectrally to 550 nm as well VIIRS band central wavelengths. Prior SOAR validation studies interpolated using the closest available AERONET wavelength and the AE, where AE (denoted α) is defined

$$\alpha = -\frac{d \log(\tau(\lambda))}{d \log(\lambda)} \approx -\frac{\log \frac{\tau_{\lambda_1}}{\tau_{\lambda_2}}}{\log \frac{\lambda_1}{\lambda_2}}, \quad (1)$$

for AOD (τ) around some wavelength λ . In this study, AERONET AOD spectral interpolation is performed with a least squares fit of all available AERONET AODs within the 440- to 870-nm wavelength range (typically 4, more for some configurations) to a quadratic polynomial, as follows (for coefficients a_0 , a_1 , and a_2 calculated on a point-by-point basis)

$$\log(\tau_\lambda) = a_0 + a_1 \log(\lambda) + a_2 \log(\lambda)^2. \quad (2)$$

This quadratic formulation is more robust to calibration problems in individual channels and accounts for the fact that in fine mode dominant aerosol conditions the relationship between $\log(\tau)$ and $\log(\lambda)$ is not linear but curved, dependent on fine mode particle size (Eck et al., 1999; Schuster et al., 2006). For low-AOD cases the difference is negligible, so has little effect on the conclusions of prior studies using the simpler (equation (1)) AE. In smoke- or haze-dominated cases, it results in a better representation of the true AOD at 550 nm.

The AE is also evaluated in this study. For VIIRS the AE provided is defined between 550 and 870 nm and is compared with AERONET AE computed over the closest wavelength range. Typically, this is 500–870 nm, but 440–870 nm is used for a few cases where the AERONET 500 nm channel is absent.

As in prior studies, the spatioemporal aggregation method of Ichoku et al. (2002) is followed and VIIRS and AERONET data are compared by averaging satellite data within 25 km of the AERONET site and AERONET data within ± 30 min of the satellite overpass. To further mitigate inherent differences in the sampling approaches of the two measurement types, two modifications are made to the strategy from prior analyses:

1. The median, rather than mean, values are reported. This is helpful when, for example, a satellite averaging area includes aerosol plumes which pass near to but not over the AERONET site.
2. Only aerosol retrievals with a surface elevation within 100 m of the AERONET site are considered, to decrease any systematic sampling biases from the fact that coastal/island Sun photometers are often not at sea level. For example, for a typical background AOD ~ 0.15 with the bulk of the aerosol within a well-mixed boundary layer of ~ 1.5 km, a Sun photometer situated at 100 m above the larger-scale surface level would miss ~ 0.01 of the AOD, which would lead to an apparent but untrue positive bias in the satellite data. Thus,

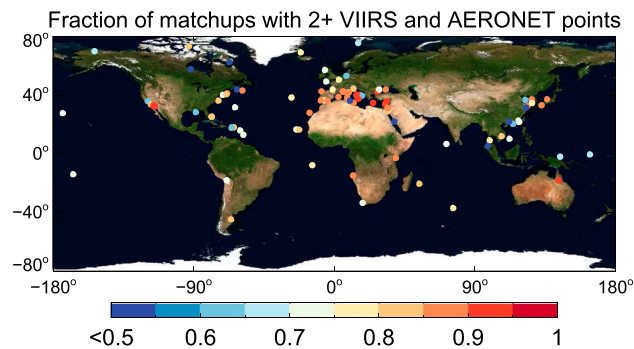


Figure 2. Fraction of AERONET/VIIRS matchups with at least two contributing points from each data set, for the full SOAR retrieval. Data shown for those 98 sites with at least 50 valid matchups with the full retrieval algorithm.

this threshold should ensure to a first order that biases resulting from elevation mismatches are typically similar to or smaller than the AERONET AOD uncertainty, which is dominated by radiometric calibration (Eck et al., 1999).

A matchup is valid if there is at least one VIIRS retrieval in the spatial window and at least one AERONET observation in the temporal window. This results in 146 sites with matchups with the full retrieval processing path and 213 with the backup turbid/shallow path (as many AERONET sites are in complex coastal environments). Of these sites, 98 and 126 have at least 50 matchups with the full and backup algorithms, respectively.

Many other satellite validation analyses which follow the recipe of Ichoku et al. (2002) put a threshold on the minimum number of satellite retrievals or AERONET observations in the averaging window required for a matchup to be considered valid. The purpose of this is to decrease the aforementioned inherent differences from the satellite versus ground-based observation types by removing heterogeneous scenes. However, this may introduce additional sampling-related biases into the aggregate statistics, as more heterogeneous (and often difficult to retrieve) scenes may be excluded, particularly near coasts where by nature there are fewer over-water satellite retrievals available. A danger is that this skews the sample to locations and times of year where features which limit satellite/AERONET data availability (e.g., coastline shape, presence and structure of cloud fields) are minimized and may therefore provide an unrepresentative assessment of satellite retrieval performance.

For example, requiring at least two AERONET and two VIIRS data points within the spatiotemporal window removes 19.5% of full retrieval matchups from consideration. This may paint a different picture of retrieval error characteristics than a more comprehensive comparison would, due to the exclusion of more complex cases. Figure 2 shows that the lost matchups are not spread evenly. While there are exceptions, the large cluster of sites around the Mediterranean tend to lose fewer. Island sites tend to lose more, as do those in complex terrain where there is often not enough over-water area for multiple VIIRS retrievals to fit. The six sites losing more than half their matchups (Anmyon, Churchill, Ieodo Station, Iqualuit, KAUST, and USM Penang) all fit into this latter category. In this analysis the use of medians and elevation thresholds as above is intended as an alternative approach to filter data while not screening out these more difficult-to-retrieve scenes, providing a more representative picture. There is scope to further refine the matchup methodology; the spatiotemporal averaging thresholds suggested by Ichoku et al. (2002) and adopted here are somewhat arbitrary but serve as reasonable global-average values to balance sampling coverage with representativeness. Future work should investigate region-specific or site-specific thresholds, such that data volume can be increased while avoiding adding significant sampling-related uncertainty.

3.2. Evaluation Metrics

The main metrics used to evaluate the SOAR data are as follows:

1. The correlation coefficient, as a measure of how well the satellite data track the variability of the AERONET data. Spearman's rank correlation coefficient is used rather than the more common Pearson's linear correlation coefficient. The reasons for this include the facts that the relationship between AERONET and satellite AOD may not be linear and also that Spearman's correlation is less sensitive to extreme outliers which may be unrepresentative of the behavior of the data set. These outliers may arise due to, for example, a clean background site having infrequent high-AOD transport events, where these rare events drive

Table 2
Global and Regional Statistics for the VIIRS-AERONET 550-nm AOD Comparison, for Full (Backup Turbid) Retrieval Matchups

Region name	Number of matchups	Correlation	Bias	f_{EE}^a	f_G	RMSE
Global	38,600 (37,946)	0.87 (0.88)	0.021 (0.025)	0.66 (0.66)	0.52 (0.49)	0.068 (0.084)
REM/CLN	4,019 (2,703)	0.77 (0.76)	0.018 (0.008)	0.65 (0.76)	0.55 (0.70)	0.064 (0.056)
EUR/NAM	12,190 (19,712)	0.84 (0.84)	0.016 (0.023)	0.71 (0.63)	0.59 (0.53)	0.053 (0.061)
SAF	405 (659)	0.88 (0.85)	0.020 (0.036)	0.69 (0.55)	0.54 (0.41)	0.065 (0.076)
NAME	5,721 (3,919)	0.91 (0.95)	0.034 (0.033)	0.58 (0.62)	0.43 (0.44)	0.070 (0.077)
MED	10,904 (4,532)	0.86 (0.86)	0.022 (0.027)	0.67 (0.59)	0.52 (0.47)	0.057 (0.067)
SAM	663 (1,121)	0.84 (0.76)	0.019 (0.028)	0.68 (0.55)	0.57 (0.48)	0.054 (0.10)
NEA	3,277 (2,623)	0.90 (0.90)	0.017 (0.036)	0.60 (0.47)	0.41 (0.33)	0.12 (0.17)
SEA/AUS	1,421 (2,677)	0.83 (0.84)	0.025 (0.050)	0.57 (0.39)	0.44 (0.28)	0.098 (0.15)

^aNote that EE is $\pm(0.03+0.1\tau_A)$ for the full and $\pm(0.03+0.15\tau_A)$ for the backup algorithm.

- up Pearson's correlation coefficient, or else sampling mismatches (e.g., high-AOD plumes or cloud contamination in one data set but not the other), which would tend to drive down Pearson's correlation coefficient. Where the true relationship is linear and departures are Gaussian, the two metrics are equivalent.
2. The median bias between the data sets, defined VIIRS-AERONET, as a measure of the general offset. Again, medians are more robust to outliers which can skew the means.
 3. The root-mean-square error (RMSE), which is a commonly-reported metric, although is dependent upon the typical level of AOD at individual sites as well as the presence of outliers.
 4. The fraction (f_{EE}) of points matching AERONET within the retrieval's expected level of error (EE). The EE is indicated to be a one standard deviation confidence interval such that one standard deviation (~68%) of matchups agree with AERONET within this bound. Based on theoretical grounds and confirmed with MAN data, Sayer et al. (2018) estimated the EE for the full SOAR VIIRS algorithm to be $\pm(0.03+0.1\tau_A)$, where subscripted A refers to AERONET AOD (i.e., a diagnostic, not prognostic, uncertainty estimate). The EE for the backup 4-band algorithm was not quantified, but herein is taken to be $\pm(0.03+0.15\tau_A)$, on the basis that the aerosol contribution to the TOA signal is smaller in the near infrared/swIR bands used in that algorithm, and uncertainty is therefore likely to be larger.
 5. The fraction (f_G) of points matching AERONET within the Global Climate Observing System (GCOS) group goal AOD uncertainty for a climate data record, which is the larger of 0.03 or 10% (GCOS, 2011). An advantage of this metric is that it is not sensor-specific and so can be used as a benchmark across multiple different data sets, e.g., Popp et al. (2016).

3.3. Global and Regional Direct Sun Comparison

Global and regional summary statistics are shown in Table 2, for the regions shown in Figure 3 (and see region name acronym definitions in that figure). Regional assignments are somewhat subjective but represent an

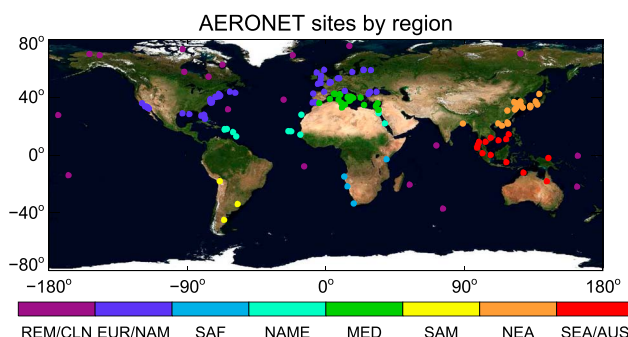


Figure 3. Regional assignment for the AERONET sites used in this study, for sites with at least 50 matchups in total between the full and backup retrieval algorithms. REM/CLN indicates remote/clean sites, EUR/NAM Europe/North America, SAF Southern Africa, NAME North Africa/Middle East, MED Mediterranean basin, SAM South America, NEA North-eastern Asia, and SEA/AUS South-eastern Asia/Australasia.

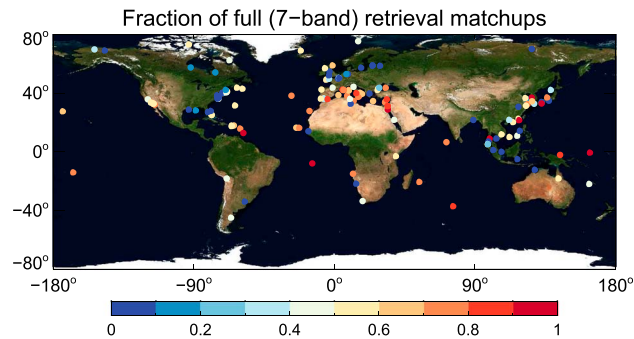


Figure 4. Fraction of VIIRS-AERONET matchups from the full seven-band retrieval algorithm. Data shown for sites with at least 50 matchups in total between the full and backup retrieval algorithms.

attempt to cluster sites subject to similar aerosol/surface conditions together, with the understanding that sites may sample different aerosol mass types through the course of seasons or years. For example, several sites in the Caribbean are included in the North Africa/Middle East region as they are predominantly low-AOD maritime sites but, particularly in the summer, can sample long-range transported African mineral dust (Prospero et al., 2014). The remote/clean (REM/CLN) category includes those typically low-AOD sites without strong local aerosol sources, predominantly on islands or isolated coastlines, although again some of these sites may seasonally sample transported continental air masses and it is impossible to fully exclude continental influence. The Mediterranean basin is further separated from the EUR/NAM and NAME regions both because sites often sample air masses from both regions and also because the relationship between water constituents and resulting ocean color there is known to be somewhat distinct from that of other large water bodies (e.g., Volpe et al., 2007). As such it is plausible that the surface reflectance model applied in SOAR (Sayer et al., 2017, 2018) may lead to larger uncertainties in this region; indeed, Table 2 suggests that errors in MED are slightly larger on average than EUR/NAM.

Further, Figure 4 shows, for the same sites, the number of full seven-band retrieval matchups as a fraction of the total number (seven-band plus backup four-band) of matchups. This shows strong heterogeneity: retrievals at island sites are most commonly with the full retrieval algorithm, while in complex coastal sites the backup algorithm is applied more often. This provides further motivation to analyze the retrievals from these two approaches separately. Note that a single AERONET measurement may contribute to the statistics for both the full and backup VIIRS algorithms, if retrievals of both types are within the spatial averaging window. Globally, a similar number of total matchups are obtained from the full and backup retrievals. Although the overwhelming majority of VIIRS retrievals are for the full seven-band algorithm (Figure 1), this is reflective of the nonrandom placement of AERONET sites available and the limitations of looking at global average results without consideration of their spatial distributions. Global statistics are very strongly weighted toward the EUR/NAM and MED regions, each accounting for about a quarter of matchups, due to the uneven distribution of sites. Only 10% of matchups with the full algorithm come from sites within the remote ocean category, which is that expected to correspond to conditions over the bulk of the world's oceans.

From Table 2, overall 66% of matchups fall within the EE of $\pm(0.03 + 0.1\tau_A)$ for the full and $\pm(0.03 + 0.15\tau_A)$ for the backup algorithms in both cases. This is slightly short of the 71% found by Sayer et al. (2018) for 836 matchups with ship-based MAN data. NAME and SEA/AUS have slightly poorer f_{EE} than average, while EUR/NAM has slightly higher. In all regions, the bias is slightly more positive than the 0.013 from MAN comparisons reported by Sayer et al. (2018). Despite this, the overall picture from these statistics is similar to Sayer et al. (2018). There is some difference in regional patterns of statistics between the full and backup algorithms, which may reflect the different sampling of those algorithms with individual regions (Figure 4), although correlation coefficients tend to show less variation. This AERONET comparison complements the prior MAN analysis and shows that there is some regional variability in quality (benchmarked primarily by bias and f_{EE}), but overall performance is similar to expectations from Sayer et al. (2018).

The GCOS compliance fraction (52% for full retrieval and 49% for backup retrieval) is within the range of results from three Advanced Along-Track Scanning Radiometer (AATSR) AOD retrieval algorithms reported by Popp et al. (2016) of 31%, 58%, and 66%. The sample for the AATSR analysis was drawn from only 1 month of data and consisted of 102 or fewer matchups, however, and so those numbers might not be representative of

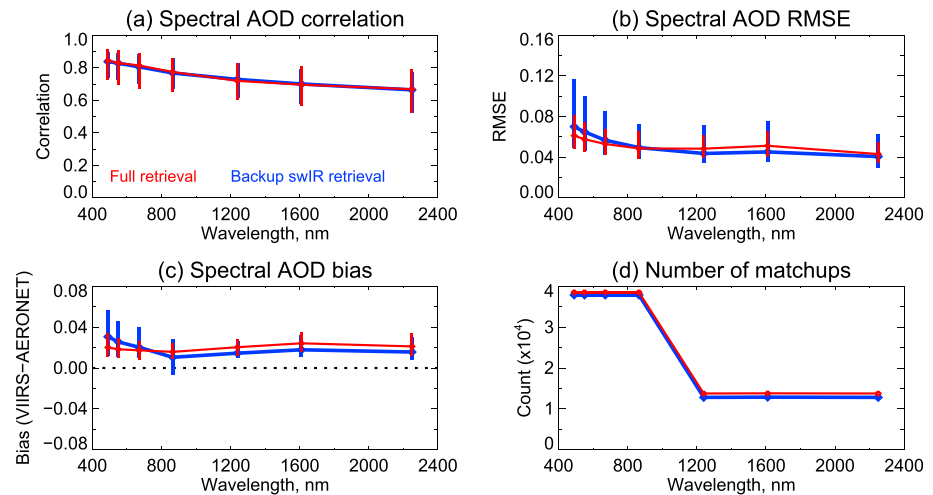


Figure 5. (a–d) Validation metrics for spectral AOD for the full (red) and backup (blue) retrieval algorithms between different sites. Symbols indicate median values across sites and vertical bars the central 68% of site-based statistics. Only sites with at least 50 matchups are considered.

broader-scale performance. Popp et al. (2016) did include results for a larger set of matches for one of the algorithms and found that f_G did not change much for that case. Note that this GCOS compliance fraction has not yet been widely adopted in satellite AOD validation exercises which makes comparative assessment difficult at the present time, so these results provide one benchmark for future studies with other data sets.

Expanding the analysis beyond 550 nm, Figure 5 shows the variability of site-by-site validation statistics for the bands at which SOAR provides retrieved AOD (490, 550, 670, 870, 1,240, 1,610, and 2,250 nm). Note that AERONET AOD is only computed for the three longest wavelengths here if the Sun photometer includes a filter at 1,600 nm, which is limited to about a third of the points collected. These statistics are presented as median and variability (central one standard deviation) of the statistics at individual sites, such that the larger vertical bars indicate larger variability between sites, and vice versa. As shown in Figure 4, many sites have matchups with both the full and backup retrieval algorithms.

Spectrally, the typical VIIRS-AERONET AOD correlation shows a decline from about 0.85 at 490 nm to 0.65 at 2,250 nm, which is expected because the dynamic range of AOD in the swIR is fairly small, since AOD tends to decrease with increasing wavelength. Variability between sites is around ± 0.1 at shorter wavelengths, and interestingly, results for full and backup retrievals are close to identical. Bias and root-mean-square error are closest to zero on average for the full retrieval up to about 870 nm and slightly closer to zero for the backup retrieval at longer wavelengths. This may reflect the fact that the backup retrieval is only fitting these bands, while the full retrieval is using all seven. For both these statistics, the full retrieval shows less site-to-site and less spectral variability, which could be explained by both a lower aerosol signal for the bands in the backup retrieval and also the different mechanisms leading to the backup retrieval being employed (e.g., shallow but clear water, different types of turbidity from continental outflow, and potential land contamination) affecting the error characteristics in more variable ways.

3.4. Dependence of Errors on AOD and Aerosol Type

Another relevant question is how the retrieval errors identified above depend on aerosol loading and type. This is addressed in Figure 6, which stratifies the available data into three broad categories based on the AERONET AOD and AE: *background* conditions, defined as $AOD \leq 0.2$; *dust-dominated* conditions, defined as $AOD > 0.2$, $AE \leq 1$; or *fine dominated* conditions, defined as $AOD > 0.2$, $AE > 1$. These groupings provide 29,667 (29,319), 4,031 (3,355), and 4,902 (5,272) matchups for the full (backup) retrieval algorithms, respectively. Since about 75% of the data belong to the background category, no separate *mixed high-AOD* group of intermediate AE is split out. Note that these categories are similar to the regions of state space for the SOAR aerosol optical models, although here the data are classified only using the AERONET data. Figure 6 shows median and variability of retrieval errors as a function of AOD, splitting the points within each category into 15 bins.

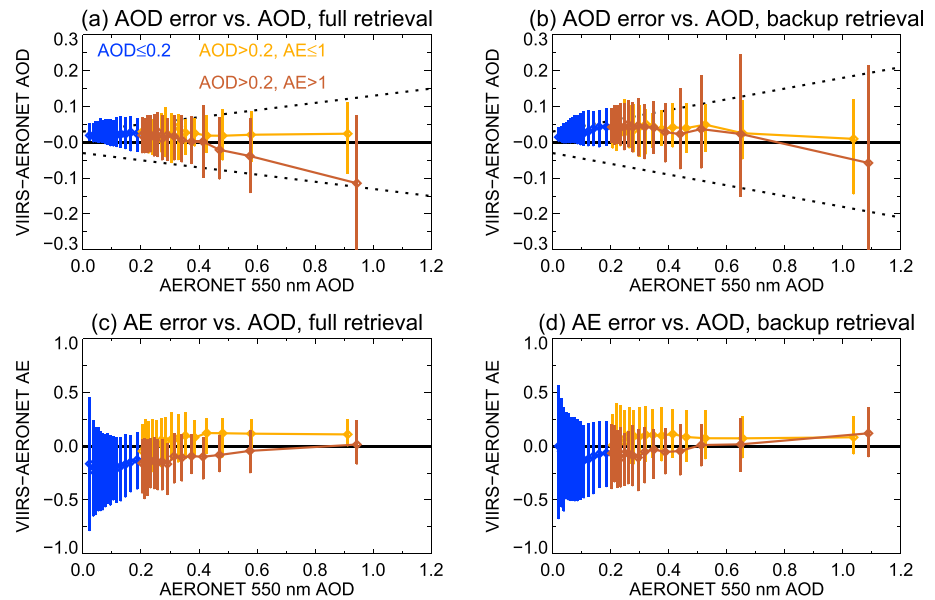


Figure 6. Median (points) and central 68% (vertical bars) of retrieval errors for 550-nm AOD (a, b) and AE (c, d), for the full (left column) and backup turbid (right column) retrieval algorithms. Data are split into background (blue), dust-dominated (orange), and fine mode dominated (brown) populations and binned (see text). For the top panels, dashed envelopes indicate the AOD EE: $\pm(0.03 + 0.1\tau_A)$ for the full and $\pm(0.03 + 0.15\tau_A)$ for the backup algorithms.

For the background grouping and full retrieval algorithm, the AOD tends to be overestimated by ~ 0.02 and the AE underestimated by ~ 0.2 . When AOD is low, the AERONET AE can be quite uncertain (Wagner & Silva, 2008), as it is essentially a gradient between two small numbers; this fact alone explains part of the decreasing width of the VIIRS AE error distribution as AOD increases from near zero to 0.2, although the fact that the negative offset appears fairly systematic suggests some systematic bias in the data. The AOD bias is more positive and AE bias less negative in the corresponding category for the backup algorithm. This difference in AOD/AE bias (as opposed to just a change in noise) between full and backup algorithms suggests some spectral dependence in the calibration or forward model bias at shorter versus longer wavelengths.

When AOD is low, the main factors influencing AOD/AE retrieval error are sensor calibration and the surface reflectance model. Aerosol optical model also contributes, but generally to a lesser extent, because the largest contributions to the TOA signal in these cases tend to be from surface reflectance and Rayleigh scattering. Separating out the contributions of each factor is not trivial. As previously noted and described by Sayer, Hsu, Bettenhausen, et al. (2017), the VIIRS TOA reflectances used in the SOAR data set are cross calibrated against MODIS Aqua. That analysis found that applying such a cross calibration decreased AOD retrieval errors. The MODIS absolute calibration uncertainty for solar bands is expected to be of order 2% (Toller et al., 2013). Uncertainties in the correction for absorption of trace gases, which act in a similar way to radiometric calibration biases, could contribute an uncertainty of a similar amount; SOAR uses the coefficients detailed by Patadia et al. (2018). Numerical experiments with this type of retrieval algorithm (e.g., Sayer, Hsu, et al., 2012; Sayer et al., 2018) suggest that calibration uncertainties of this magnitude can be responsible for an AOD bias of ~ 0.01 , dependent on the spectral correlation of the calibration bias. The AE is particularly sensitive to this spectral correlation of bias. It is therefore plausible that a significant proportion of the bias in the data set is linked to the absolute calibration and trace gas absorption corrections of the sensor, and future work should continue to assess the absolute and time dependence of the calibration (for both MODIS and VIIRS), as well as make use of updated gas spectroscopic absorption data bases and ancillary reanalysis data sets as they become available.

The SeaWiFS and AVHRR applications of SOAR showed similar positive AOD offsets in low-AOD conditions (Sayer, Hsu, et al., 2012; Sayer, Hsu, Lee, et al., 2017), although the wavelengths available to each sensor do not exactly match. This suggests some contribution from the radiative transfer model (surface/aerosol assumptions), as this has been basically the same for the various applications of SOAR. The aerosol optical model used in low-AOD cases, described in Sayer et al. (2012), is based on AERONET almucantar scan inversions (Dubovik

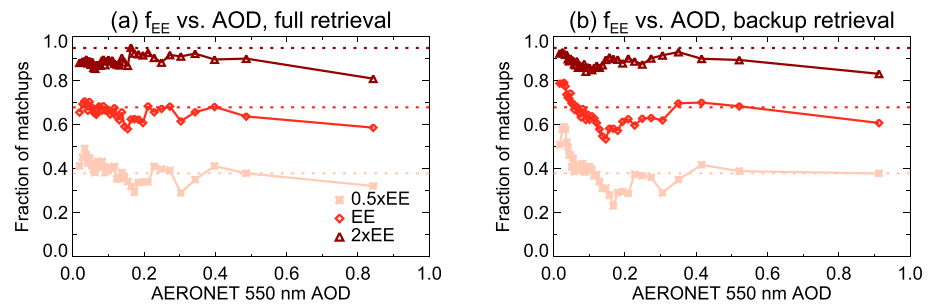


Figure 7. Fraction of matchups in agreement with AERONET within 0.5 (pink), 1 (red), and 2 (dark red) times the EE as a function of AOD, for the full (a) and backup turbid (b) retrieval algorithms. Note that the EE is $\pm(0.03 + 0.1\tau_A)$ for the full and $\pm(0.03 + 0.15\tau_A)$ for the backup algorithms. Points and solid lines indicate binned results while dashed lines indicate the theoretical expected behavior.

& King, 2000) from AERONET version 2. For future applications of SOAR, assessing how different version 3 inversions are from version 2, and the effect of these differences on the TOA signal, would guide refinement of this optical model and possibly resolve some of these residual biases. The surface reflectance model used (Sayer et al., 2017, 2018) is similar to those of many other algorithms, including contributions from Sun glint, whitecaps (foam), and in-water scattering and absorption from plankton pigments. Many of the coefficients within these models are inherently empirical in nature and may contribute to systematic offsets at some wavelengths. Ultimately, without an on-orbit absolute calibration source and decreases to the uncertainties of some of the assumptions going into the radiative transfer forward models, it is unlikely that biases can be entirely removed across all wavelengths. Thus, it is likely that empirical bias corrections of retrieval output, such as Zhang and Reid (2006), will continue to be necessary to further reduce errors.

When AOD is higher, differing behavior is seen in the retrieval errors dependent on whether the column is optically dominated by fine (e.g., smoke and haze) or coarse (e.g., dust and ash) particles. For coarse particles the AOD bias is close to zero, and overall uncertainties appear smaller than the EE, for both full and backup retrieval algorithms. The AE in these cases has a small positive bias (0–0.15) and the typical retrieval error quickly shrinks to around ± 0.2 , becoming closer to ± 0.1 for the full algorithm when AERONET AOD is higher than about 0.4. The dust optical model used in SOAR is described by Lee et al. (2017) and treats particle shape using the same ellipsoid shape distribution as in AERONET inversions (Dubovik et al., 2006), which significantly decreases spectral and angular retrieval artifacts compared to the spherical particle assumption. One further point to note is that the majority of points within this coarse dominated category are obtained from AERONET sites sampling dust from North Africa or the Arabian Peninsula, which is the main region used by Lee et al. (2017) to develop the constraints for the dust optical model. It is plausible, but currently hard to assess due to a dearth of AERONET data, that dust from other source regions might exhibit different error characteristics due to different mineralogical compositions, size, and shape distribution.

Table 3
Prognostic AOD Expected Error Coefficients $a + b\tau_v$ for the VIIRS SOAR Data Set and Coefficient of Determination of the Fit of the Expression

Aerosol model	Intercept a	Gradient b	R^2
Full retrieval			
Maritime	0.00	0.36	0.99
Dust	0.07	0.031	0.21
Fine dominated	0.01	0.21	0.98
Mixed	0.03	0.15	0.92
Backup retrieval			
Maritime	0.00	0.42	0.99
Dust	0.09	0.018	0.07
Fine dominated	0.01	0.27	0.98
Mixed	0.04	0.21	0.91

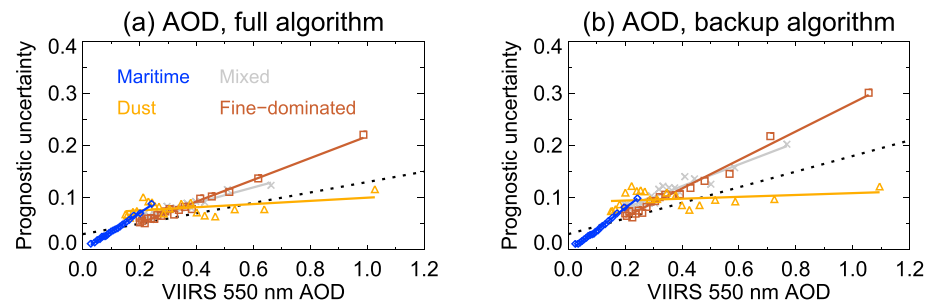


Figure 8. Prognostic expected errors for 550 nm AOD retrieval for the (a) full and (b) backup retrieval algorithms. Points and lines show binned and the linear fit of prognostic uncertainties for retrievals from the maritime (blue), dust (orange), fine mode dominated (brown), and mixed (gray) aerosol optical models. The dashed line shows the diagnostic expected error, $\pm(0.03 + 0.1\tau_A)$ for the full and $\pm(0.03 + 0.15\tau_A)$ for the backup algorithms but plotted as though it were relative to VIIRS.

For the fine mode dominated subset, AE is retrieved with negligible bias for both full and turbid algorithms, and the magnitude of the error is around ± 0.2 for moderate and high AOD. The larger AE errors compared to the dust-dominated subset might reflect the fact that AE is sensitive not only to the fine/coarse aerosol partition but also to fine mode particle size (Schuster et al., 2006). For smoke aerosols, it is well known that type of fuel and moisture variations can cause large variation in smoke properties at the point of emission (e.g., reviews by Reid, Eck, et al., 2005; Reid, Koppmann, et al., 2005); Sayer et al. (2014) used AERONET inversions to define optical models for smoke from different source regions, which sample more aged or transported smoke, and found these differences in particle size and spectral shape of AOD persisted. As AOD increases within this subset, SOAR underestimates by on average 10% for the full and 5% for the backup algorithm, and the variability of the error is also larger than expected. These results together suggest that future versions of SOAR should include more than one fine mode dominated aerosol optical model, with different particle sizes and single scattering albedos (SSA), to better account for the diversity of optical properties of fine mode dominated aerosol columns. The SSA of the current fine dominated model varies between 0.95 and 0.98 in the midvisible, dependent on FMF (Sayer et al., 2018), which is likely too large for transported absorbing particles.

The overall AOD dependence of the fraction of matchups within EE (f_{EE}) is shown in Figure 7, splitting the data into 40 bins (~ 950 matchups per bin). If the EE is a good representation of actual retrieval error (i.e., the discrepancy against AERONET), then by Gaussian statistics $\sim 38\%$ of matchups should agree within half the EE, $\sim 68\%$ of matchups should agree within the EE, and $\sim 95\%$ of matchups should agree within twice the EE. Figure 7 shows that these patterns hold reasonably well across the data set, although the fraction within twice EE is around 5% lower than expected; some of these may be residual sampling mismatches. For both the full and turbid algorithms there appears to be a slight decline in performance for high-AOD cases. This is consistent with the tendency for negative offset in high-AOD fine mode dominated aerosol columns seen in Figure 6. As the EE is different for the two algorithms, comparing Figures 7a and 7b does not compare the absolute uncertainty in the retrievals but rather their relative compliance with their expected levels of uncertainty.

Retrieval error has also been examined as a function of near-surface wind speed, total column water vapor, climatological chlorophyll *a* concentration, and time since mission launch. These results are omitted here as it was found that changes in the bias or width of retrieval error distribution varied negligibly (generally by 0.01 or less) across the range of the equivalent parameter observed. These findings are consistent with the driving uncertainties in the retrieval algorithm being related to sensor calibration, aerosol optical models, and surface reflectance model.

3.5. Prognostic AOD Uncertainty Estimates

While diagnostic (i.e., relative to a reference truth) uncertainty estimates are useful for understanding the distributions of retrieval error and guiding future algorithm development, applications such as data assimilation (DA) typically require prognostic (predictive, i.e., relative to the retrieved state) uncertainty estimates. That is because they require that each retrieval has an associated uncertainty to be able to weight it within the DA system (Benedetti et al., 2018). Prognostic estimates may be obtained using error propagation techniques such as Optimal Estimation (OE, Rodgers, 2000) or more empirical methods (e.g., Sayer et al., 2013; Shi et al.,

2011; Zhang & Reid, 2006, for some aerosol examples). Generally, they are expressed as the magnitude of a confidence interval around the retrieved state.

To facilitate AOD DA applications of the SOAR data set, Table 3 shows coefficients for the prognostic expected uncertainty (EE_p) of 550 nm AOD, assuming a linear model

$$EE_p = a + b\tau_V, \quad (3)$$

where subscripted V indicates it is defined relative to the satellite (VIIRS) AOD. This methodology for determining these coefficients is a similar method to that in Sayer et al. (2013). The VIIRS AOD retrievals matched with AERONET are split into categories based on which aerosol optical model was chosen as best fit by the retrieval (since this is a prognostic estimate rather than diagnostic, it can only use information provided within the retrieval output). Then the retrievals are sorted by VIIRS AOD and divided into 20 equally populated bins per optical model (providing between 110 and 1,300 matchups per bin, dependent on optical model). Within each bin, the 68th percentile (i.e., one standard deviation) of absolute retrieval error, that is, $|\tau_V - \tau_A|$, is fit to the bin average τ_V to determine the coefficients a and b in equation (3). These fits are shown graphically in Figure 8. The resulting coefficients can be used directly to assign pixel-level uncertainties for DA or other applications; note that they only apply to QA = 3 data (as that is what is included in the AERONET comparison), and as indicated earlier, QA = 1 data should not generally be used.

For both algorithms and for all optical model bar dust, the coefficient of determination (R^2) on these fits is high, ranging from 0.91 to 0.99, confirming that the simple linear model appears to be valid over this range of data. The exception is the dust model, for which the increase in retrieval error with AOD is similar to the noise on the fit (i.e., the AOD dependence of errors is fairly small in this case). The type dependence of these fits is again in part reflective of biases in aerosol optical model assumptions. Note that the upper AOD bound of the maritime model allowed in the SOAR algorithm is 0.25, and the lower bounds of the dust, fine dominated, and mixed models are 0.15, 0.2, and 0.2, respectively. Additionally, there are few matchups available with AOD above one. This explains the limits of the lines shown in Figure 8, and it is possible that in cases of very high AOD the linearity found by these fits will break down.

An empirical approach like this has advantages and downsides. An advantage is that, because it is based on real validation results, it does not require having an accurate quantitative model of each component contributing to retrieval error like OE does. However, a limitation is that it is necessarily a more simplistic error

Table 4
As Table 3 Except for Data Split Into the Regions Shown in Figure 3

Aerosol model	Intercept a	Gradient b	R^2
Full retrieval			
REM/CLN	0.02	0.20	0.66
EUR/NAM	0.01	0.23	0.93
SAF	0.02	0.14	0.76
NAME	0.04	0.09	0.52
MED	0.02	0.19	0.92
SAM	0.01	0.24	0.81
NEA	0.02	0.20	0.98
SEA/AUS	0.02	0.21	0.88
Backup retrieval			
REM/CLN	0.01	0.29	0.88
EUR/NAM	0.00	0.33	0.96
SAF	0.01	0.28	0.93
NAME	0.04	0.08	0.86
MED	0.01	0.30	0.99
SAM	-0.01	0.59	0.89
NEA	0.01	0.29	0.99
SEA/AUS	0.01	0.39	0.98

Table 5
Global and Regional Statistics for the VIIRS-AERONET 550 nm FMF Comparison, for Full (Backup Turbid) Retrieval Matchups

Region name	Number of matchups	Correlation	Bias	RMS difference
Global	32,985 (33,132)	0.73 (0.68)	-0.036 (-0.004)	0.19 (0.19)
REM	3,730 (2,510)	0.79 (0.67)	-0.059 (-0.056)	0.19 (0.19)
EUR/NAM	10,571 (17,228)	0.62 (0.61)	-0.034 (0.0)	0.21 (0.20)
SAF	243 (515)	0.71 (0.70)	-0.018 (0.007)	0.17 (0.17)
NAME	4,744 (3,366)	0.58 (0.57)	-0.014 (0.0)	0.16 (0.16)
MED	8,614 (3,493)	0.72 (0.68)	-0.024 (0.001)	0.17 (0.18)
SAM	627 (1,062)	0.72 (0.39)	-0.072 (0.041)	0.18 (0.26)
NEA	3,086 (2,397)	0.65 (0.63)	-0.081 (-0.027)	0.18 (0.16)
SEA/AUS	1,370 (2,561)	0.67 (0.60)	-0.031 (0.043)	0.19 (0.21)

model, reliant on those times and locations where validation data are available, and there may be situations where the predicted retrieval uncertainty does not reflect reality. In particular, for example, as noted earlier the bulk of the dust aerosol model matchups is for sites sampling Saharan outflow so error characteristics for other dust-laden regions may differ (i.e., these metrics may not capture regional variations within a given type). A second limitation is that it folds in some of the uncertainties related to the matchup method and AERONET data rather than the retrieval itself (e.g., Virtanen et al., 2018). Practically speaking, it would also be reasonable to assume a minimum possible uncertainty of ~ 0.01 on the basis of the radiometric uncertainty in the TOA reflectance measurements.

To take a second view at prognostic uncertainty estimates, Table 4 provides equivalent coefficients to Table 3, except for data split into the regions shown in Figure 3 instead of by type. Note that the data are not further broken down by type, and are only split into 10 bins per region, due to the limited data volume in some regions. These may be more useful for regional analyses, with the caveat that the data volume in some regions is low, and may be dominated by a small number of sites. Coefficients of determination are in general a bit lower than in Table 3 but still high (0.52–0.98), likely in part because now the type dependence of the retrieval is not part of the error model. For those regions dominated by a single aerosol type (e.g., most matchups from REM/CLN are for the maritime model), numbers are quantitatively similar. For NAME, which is a mixture of dust and marine aerosols, the expressions are different (as Table 4 is including those lower-AOD points absent from the dust optical model state space) but results are numerically fairly similar.

For an empirical approach, the results in Tables 3 and 4 are a starting point for understanding in a prognostic sense the retrieval uncertainties in the SOAR VIIRS data. For future versions of the data set, a sensible goal would be to add in a second (nonempirical) technique such as OE. Through complementary use of the theoretical and empirical approaches, it may be that the strengths and limitations of each could be better understood. Aside from Popp et al. (2016) and work by DA teams (whose requirements and efforts are

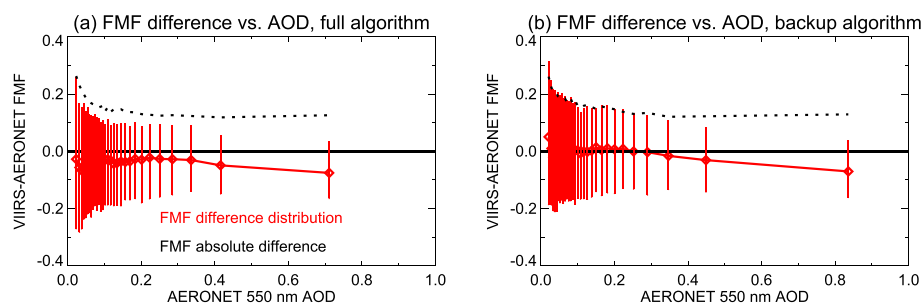


Figure 9. Median (points) and central 68% (vertical bars) of VIIRS-AERONET SDA retrieval differences for FMF for the full (a) and backup turbid (b) retrieval algorithms. Data split 30 equally populated bins. Dashed black lines indicate the 68th percentile of absolute difference within each bin.

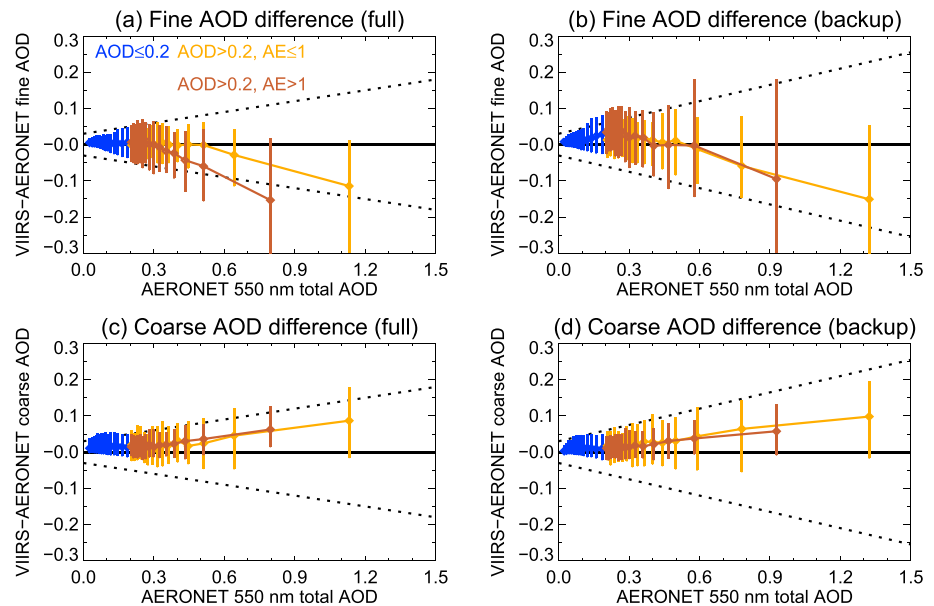


Figure 10. As in Figure 6 except for (a, b) fine mode and (c, d) coarse mode AOD from the comparison between VIIRS and AERONET spectral deconvolution algorithm data.

discussed by Benedetti et al., 2018), pixel-level uncertainty estimates provided within in remotely sensed AOD have in general not yet been robustly evaluated.

3.6. AERONET Spectral Deconvolution Algorithm

Besides the direct Sun AOD measurements and almucantar scan inversions, AERONET also applies a spectral deconvolution algorithm (SDA) to the direct Sun AOD to partition the fine and coarse mode contributions to the total AOD, under some assumptions of the spectral shape of fine and coarse particle extinction (O'Neill et al., 2001, 2003, 2006). This allows an evaluation of the FMF and fine/coarse mode AOD retrieved by SOAR. Note that this cannot be considered a validation to the same extent as the direct Sun AOD comparison in the previous section, because the uncertainty on SDA FMF, which is dependent on AOD and aerosol properties but is of order ± 0.1 (O'Neill et al., 2001), is not negligible.

For this purpose the same matchup methodology and basic analysis applied to the direct Sun AERONET data is used here. The SDA requires the presence of certain spectral bands and has some additional data quality checks to remove cases where its assumptions may not hold, so the data volume is somewhat lower than for the direct Sun comparison. The SDA output is provided at 500 nm, but for a consistent comparison with SOAR it is converted to 550 nm using the fine mode and total AOD and AE within the MAN SDA product, which adds negligible additional uncertainty. This wavelength adjustment makes the two FMF products more directly comparable.

Resulting global and regional statistics of the comparison are shown in Table 5. This shows little difference, in most cases, between regional and global statistics: a tendency for a small underestimation (of order -0.036) in FMF, with a RMS difference of a little under 0.2, and little difference overall between full and backup algorithms. Correlation is lower than for AOD, and for the full algorithm varies between 0.58 and 0.79, dependent on region. When considering the correlation and RMS difference in particular, it is important to bear in mind that the uncertainty on SDA FMF alone of around ± 0.1 (O'Neill et al., 2001) is not small compared to the possible range of values for this parameter (i.e., 0–1). These results are also consistent with the low offset in VIIRS AE in low-AOD conditions seen in Figure 6 and with the similar small negative offset (-0.026) and RMS difference of 0.18 reported by Sayer et al. (2018) for open ocean comparisons with the shipborne MAN SDA data set.

Unlike AOD, validation of FMF or similar quantities from other satellite data sets is uncommon, and there is no GCOS target guideline, and so it is difficult to assess this performance in the context of other published results. Kahn and Gaitley (2015) performed an evaluation of FMF from Multiangle Imaging Spectroradiometer (MISR) retrievals against AERONET and included two examples from ocean sites (Tahiti in the remote Pacific Ocean and Capo Verde in the Saharan dust Atlantic outflow region). Due to sampling limitations of MISR,

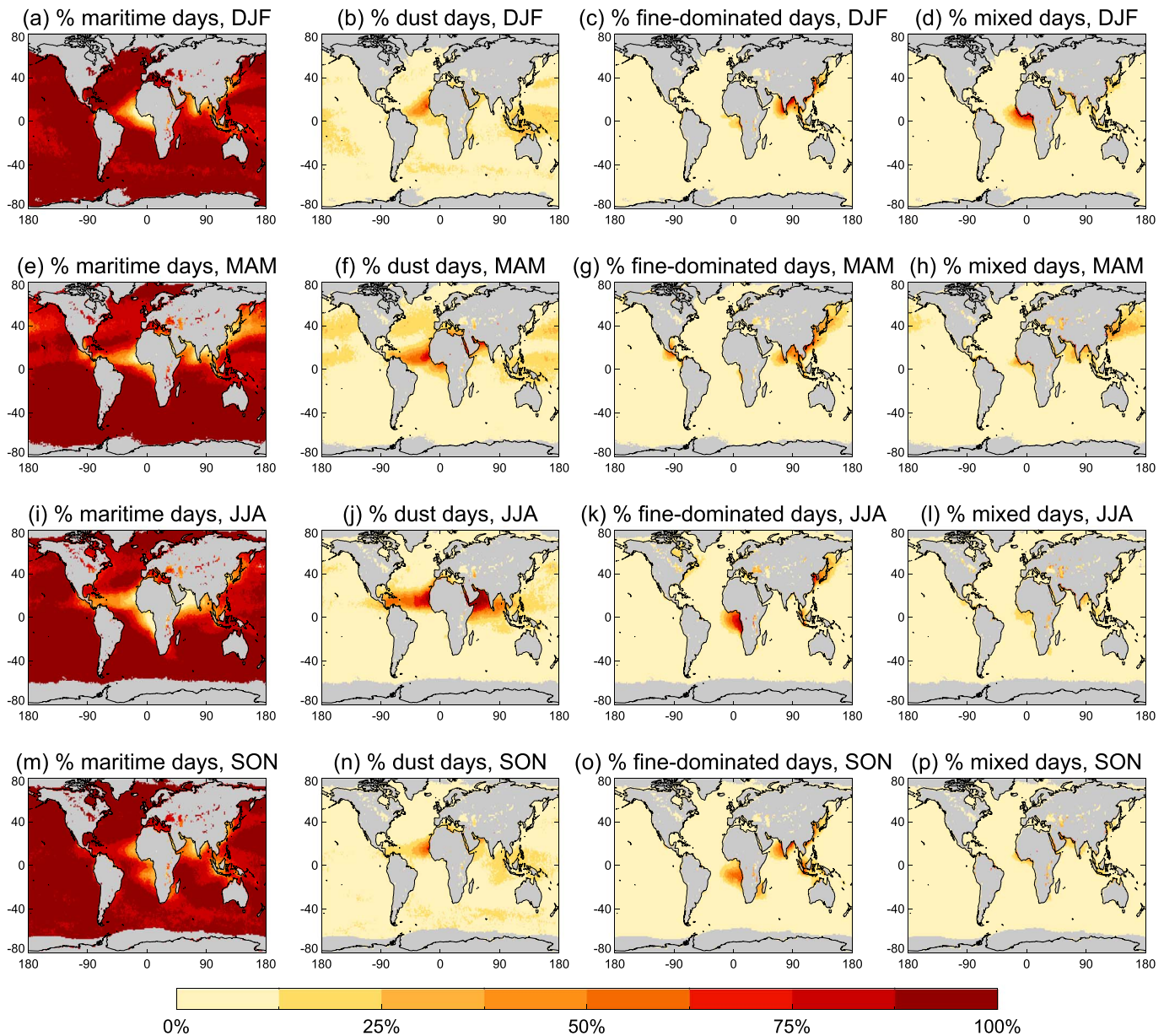


Figure 11. (a–p) Percentage of days in each season (top to bottom) where the VIIRS daily modal aerosol optical model choice was (left to right) maritime, dust, fine dominated, or mixed. Grid cells in gray indicate fewer than 30 available days. DJF = December, January, February; MAM = March, April, May; JJA = June, July, August; SON = September, October, November.

this was performed in terms of monthly climatological values for different regimes of AOD, rather than as instantaneous matchups, so is not directly comparable to these results. For these two sites MISR was found to broadly reproduce the intraannual variation of FMF, with offsets in mean values of 0.1 or smaller. Standard deviations of MISR FMF tended to be slightly larger than those of AERONET within a given month. Kleidman et al. (2005) provided a comparison of FMF from an older version of the MODIS Dark Target (DT) over-water aerosol retrieval algorithm, although with more limited sampling, and it is not clear whether those results still hold for the latest MODIS data product version. They found that MODIS had a lower dynamic range of FMF compared to SDA results (i.e., insufficiently frequent values close to 0 or 1). Kleidman et al. (2005) did not report a RMS difference but provide a correlation of 0.73 for points where AOD was 0.1 or higher. This is similar to the global average correlation in Table 5 without an AOD filter; if the SDA/VIIRS comparison is filtered to also require $AOD \geq 0.1$, correlation increases from 0.73 (0.68) to 0.81 (0.77), and RMS difference decreases from

Table 6
Confusion Matrix Showing Percentage Classification of Aerosol Type From AERONET and VIIRS for the Direct Sun Matchups With the Full (Turbid) Algorithms

VIIRS-retrieved aerosol type	AERONET-defined aerosol type			
	Dust	Fine dominated	Maritime	Mixed
Dust	5.4 (4.2)	0.1 (0.2)	4.5 (3.9)	0.8 (0.7)
Fine dominated	0.0 (0.1)	8.1 (10.0)	3.3 (6.0)	1.1 (1.3)
Maritime	0.9 (0.7)	1.2 (0.7)	67.3 (65.0)	1.4 (1.0)
Mixed	1.0 (1.0)	0.9 (1.0)	1.7 (2.3)	2.1 (2.0)

0.19 (0.19) to 0.16 (0.17) for the full (backup) algorithms. This suggests that the SOAR FMF retrieval is similar to, or possibly slightly better than, MODIS DT.

Figure 9 is similar to Figure 6, binning the SDA matchups in 30 bins (over 1,000 matchups per bin) and showing the VIIRS-AERONET difference in FMF as a function of AERONET AOD. The small negative offset in FMF (0 to -0.05) is present throughout the range of AOD. Also shown is the 68th percentile of absolute difference; this decreases with increasing AOD and quickly plateaus for $AOD \sim 0.2$ or higher at $0.13-0.14$ for both algorithms. If the uncertainty on the SDA FMF is about 0.1, then by adding in quadrature (under the assumption that errors in the SOAR and SDA algorithms are independent), this suggests that the uncertainty in the SOAR FMF may also be only around 0.1, as $(0.1^2 + 0.1^2)^{1/2} \approx 0.14$.

As another analog to Figure 6, Figure 10 shows the difference between VIIRS and AERONET fine mode and coarse mode AOD as a function of AERONET total AOD. By definition fine mode AOD is computed simply as the product of total AOD and FMF, and coarse mode AOD is the remaining difference. The EE envelope for total AOD is also shown; note that it is not trivial to define or evaluate an EE specific to fine mode or coarse mode AOD, as the SDA uncertainty on each component is dependent both on the total direct Sun AOD uncertainty (nominal ± 0.01) and the FMF uncertainty itself (nominal ± 0.1), and by definition an error in an individual AERONET FMF causes equal but opposite errors in estimated fine and coarse mode AOD. The results match expectations from Figures 6 and 9; the slight low offset in VIIRS FMF means that, when total AOD is high, total coarse mode AOD is slightly higher on average than AERONET, and total fine mode AOD slightly lower. For fine dominated aerosol loadings the negative offset outweighs the positive, such that total AOD has a low bias as seen in Figure 6. These results are again consistent with the MAN SDA comparison in Sayer et al. (2018), who found that the majority of the bias in total AOD could be attributed to a small overestimate in the coarse mode contribution.

4. Evaluation of Aerosol Type

Unlike AOD or AE, aerosol *type* is an inherently subjective concept which is hard to define quantitatively, particularly since in many real-world conditions an air mass may contain nonnegligible amounts of aerosols of different geographical and/or chemical origins, in potentially several different vertical layers. Nevertheless, due to the underdetermined nature of the aerosol remote sensing problem, most retrieval algorithms include constraints on aerosol optical properties each designed to reflect air masses dominated by aerosol from some specific region. These are often referred to as types as convenient shorthand by algorithm developers interested in comparing assumptions and by data users who wish to identify separately, for example, areas affected by dust storms or wildfire smoke. They tend to be based, dependent on the technique in question, of clustering aerosol optical and/or microphysical properties by some method. Russell et al. (2014) and references therein provide extensive discussion and examples on this topic.

In SOAR, the four optical models are designed to reflect background maritime, dust-dominated, fine mode (e.g., smoke and industrial pollution) dominated, or mixed (e.g., dust plus smoke) aerosol columns (Sayer et al., 2018). The retrieval algorithm reports the best fitting aerosol model for each pixel, with no geographical constraints applied to model selection. The purpose of this section is to examine their occurrence and attempt an evaluation of how reasonable the results of this best fit model selection are.

The VIIRS D3 products contain information on the number of retrievals choosing each optical model in each grid cell, as well as the modal model choice per cell; the M3 products contain a histogram of the daily modal

choice within each month, as well as the mode of daily modes within each month. Figure 11 shows, for each season, the percentage of days dominated by each of the four optical models, calculated from the histograms from the M3 products for the VIIRS mission to date. Qualitatively, this reveals largely the expected patterns, based on model simulations or satellite analyses (e.g., Colarco et al., 2010; Jeong & Li, 2005; Kahn & Gattley, 2015; Penning de Vries et al., 2015). Over the bulk of the low-AOD ocean the maritime optical model is picked most of the time; areas downwind of dust and fine dominated (smoke and haze) source regions tend to choose those models, and the mixed-type model is most common in between the dust and fine dominated regions. One exception is dust being the most common aerosol type 12.5–25% of the time over the tropical Pacific, when there is no obvious upwind source. This region often has widespread thin cirrus cloud cover (Eleftheratos et al., 2011), which is often subvisual (McFarquhar et al., 2000); it may therefore be that it is not always detected by the SOAR cloud mask. Cirrus clouds exhibit a roughly spectrally neutral signal across the visible/swIR at TOA, which is similar to the behavior of dust, so it may be that a small residual cirrus contribution to the signal manifests in this way.

AERONET direct Sun data do not directly provide an equivalent metric to evaluate the SOAR aerosol-type selections. However, the AERONET data can be split by AOD and AE into different regimes which overlap with the parts of state space sampled by the SOAR optical models. This provides an AERONET-based aerosol-type classification to evaluate the SOAR data. For this purpose, the AERONET matchups were taken and classified as *maritime* if $AOD < 0.2$, *dust* if $AOD > 0.2$, $AE \leq 0.6$; fine dominated if $AOD > 0.2$, $AE > 1.2$, and mixed if $AOD > 0.2$, $0.6 < AE \leq 1.2$. These AOD/AE bounds fall in the boundaries in AOD/AE space permitted by the SOAR retrieval (Sayer et al., 2018), and while subjective, classification of AERONET days does not change significantly if thresholds are adjusted within ± 0.05 for AOD and ± 0.1 for AE. The AERONET almucantar scan inversion product (Dubovik & King, 2000) provides other aerosol properties, including size distribution and SSA, which are useful for aerosol-type classification (e.g., Russell et al., 2014). However, that is not used here because the SSA information it provides is only quantitatively reliable when the AOD in the blue band is above 0.4 and the solar zenith angle is large (Dubovik et al., 2006). These constraints remove the overwhelming majority of potential matchups because the oceanic sites tend to be at lower AOD, and for the tropical sites where AOD is often higher the solar zenith angle is often insufficient. Thus, the AERONET direct Sun matchups seem most appropriate for use for a categorical classification for the present purpose.

Applying these type classifications to the AERONET data from the direct Sun matchups gives the confusion matrix shown in Table 6. Each row indicates, for matchups classified as that aerosol type by VIIRS, how the matchups were classified by AERONET. The trace of this matrix therefore gives the percentage of matchups where the classifications by VIIRS and AERONET are in agreement. This corresponds to 82.9% and 81.1% of cases for the full and backup algorithms, respectively, a very similar (and high) proportion. The bulk of matchups (around two thirds in each case) are classified as maritime by both data sets. The most common misclassification (around 4% of matchups) corresponds to data classed as maritime by AERONET but dust by VIIRS; this is consistent with the small positive and negative biases in VIIRS AOD and AE, respectively, in, for example, Figure 6. The next largest misclassification is with the mixed aerosol type, which as noted is slightly sensitive to the AOD/AE boundaries imposed on this classification, as there is some overlap in this space for the SOAR optical models. Confusion between dust and fine dominated aerosols is rare (under 0.3% of cases in total), indicating that the data are reliable for distinguishing between fine mode and dust-dominated aerosol columns.

Finally, Figure 12 shows the spatial distribution of the fractional agreement for those sites with at least 50 matchups. Spatial variation is reasonably low, indicating (so far as can be diagnosed from the available AERONET matchups) no region with lack of skill. The poorest performance is in tropical and eastern Asia with the backup algorithm, although even here sites typically make a classification consistent with AERONET 50–80% of the time. The lower performance in this region is likely due to a combination of aforementioned issues of complex aerosol mixtures and complex surface types, whose diversity may not be fully captured by the SOAR algorithm, and (for the tropical sites in particular) subvisual cirrus cloud contamination.

5. Comparison to Other Over-water Satellite Data Sets

This section compares the global and regional over-water patterns of VIIRS SOAR AOD/AE retrievals to those of other satellite data sets whose records overlap with VIIRS. This provides context to how VIIRS fits in with the broader satellite record, with the goal being to examine the time series of AOD in a similar way to how a data

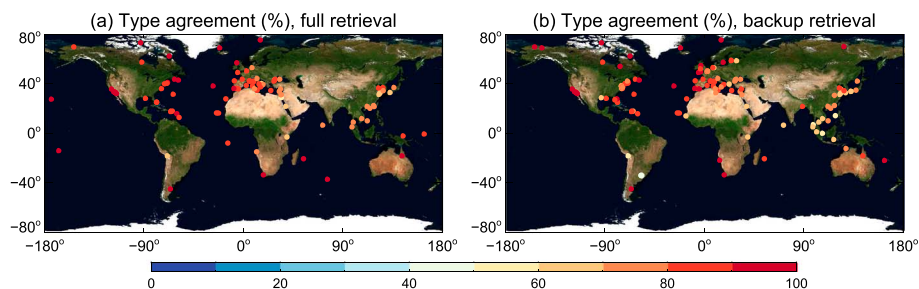


Figure 12. Percentage of matchups where AERONET and VIIRS define the same aerosol type (see text), for matchups with the full (a) and backup (b) SOAR algorithms. Data shown for sites with at least 50 matchups with the given algorithm.

user interested in spatiotemporal patterns would, and examines whether they are consistent in magnitude and seasonality. In all cases here, monthly mean data products (calculated as means of daily means), with any data selection or filtering as provided within those standard data products, are used. Although cosampling data sets using daily rather than monthly averages may reduce sampling-related differences, that is again not the purpose: this analysis looks at how consistent the available data products are for general applications, not how consistent they can be made through modification.

Each data product has slightly different characteristics, due to the nature of the sensor in question and the historical development of the data set. In all cases, the AOD nearest 550 nm is used, as is the AE closest to the wavelength range 550–870 nm. Some data sets provide monthly AE as the mean of daily AE, and some provide monthly spectral AOD from which a monthly AE can be calculated. As AE is an intensive variable (i.e., it depends on aerosol type but not amount), these two calculation methods can in principle provide different results, if AOD and AE are not homogeneous within a region for a given month. For those data sets which enable both calculations, testing the effect (not shown) reveals the difference, depends on region, tends to be fairly small (systematically ± 0.1 or so for a given region), and does not generally affect seasonality. So the default for each data set in question is used here, again, to mimic the way a data user would approach from standard L3 monthly products.

Unless otherwise stated, the source data products are provided at 1° horizontal resolution. All global and regional results in this section are area weighted (as equal-angle grid cells are otherwise larger at the equator than at high latitudes). Comparisons are performed for all available months overlapping with VIIRS.

5.1. Data Sets Used

5.1.1. MODIS Dark Target

The MODIS DT ocean data sets are the closest conceptually to SOAR; both MODIS and VIIRS are broad-swath single-view passive imaging radiometers, and both algorithms are multispectral fits of TOA reflectance to determine the total AOD and partition between fine and coarse modes for a given set of candidate aerosol optical models. Despite this, there are numerous differences in the specifics of the radiative transfer and retrieval procedure (Levy et al., 2013; Tarré et al., 1997). The MODIS swath (2,330 km) is narrower than that of VIIRS but still provides near-daily global views of a given location on the Earth's surface. A VIIRS DT data product is in development (Levy et al., 2015) but not yet publicly available. Note that the U.S. National Oceanic and Atmospheric Administration (NOAA) also provides a VIIRS over-water product based on an implementation of the DT algorithm (Jackson et al., 2013). The NOAA product was previously compared with SOAR in Sayer et al. (2018) and so is not included in the present comparison.

Here the monthly products (MOD08_M3 for Terra and MYD08_M3 for Aqua) from the latest Collection 6.1 (C6.1) are used. C6.1 has not yet been validated, but L2 validation results from Collection 6 and earlier versions (Kleidman et al., 2005; Levy et al., 2013; Sayer et al., 2012) revealed a typical AOD retrieval uncertainty of order $\pm(0.04+10\%)$, with a tendency of high bias ~ 0.02 (from a limited sample of low-AOD cases; higher for Terra than Aqua) and an AE retrieval with some skill but a lower dynamic range than AERONET. Recently, Levy et al. (2018) reported that these Terra/Aqua AOD offsets persist in C6.1; they also tested cross calibrating the two MODIS sensors against each other (which is not done in the operational C6.1 products) and found this reduced but did not remove the differences.

5.1.2. MISR

A key feature of the MISR sensor is that it obtains near-simultaneous views of a given point on the Earth's surface from nine cameras; this multiview capability provides additional information about surface and atmospheric scattering and absorption (Martonchik et al., 1998). A disadvantage is the narrow (~330 km) swath, meaning that repeat views are obtained only about once per week at equatorial latitudes and once every few days at high latitudes. MISR flies on the Terra platform.

This study uses the monthly product (MIL3MAEN), which is provided at 0.5° horizontal resolution, from the latest version 23. The main changes relevant between this and the prior version 22, which is the version which has been used in the majority of prior MISR-based studies, are an increase in the L2 retrieval resolution (pixel size 4.4 km compared to previous 17.6 km) and updates to the surface reflectance model, sensor calibration, and retrieval convergence and QA procedures (Witek et al., 2018). The MISR AOD retrieval uncertainty is context dependent but stated in general terms as the greater of 0.03 or 20% (Kahn et al., 2010); beginning in version 23, pixel-level uncertainty estimates for each retrieval are also provided, using a technique outlined in Witek et al. (2018). Large-scale validation or intercomparisons using MISR version 23 has not yet been published, although team expectations (M. Garay, personal communication, July 2018) are that a previously reported (e.g., Kahn et al., 2010) systematic positive bias in AOD in low-AOD conditions should be mitigated with this new version. Performing the analysis in this section with version 22 (omitted for brevity) reveals that version 23 AOD is indeed lower by ~0.03–0.05, dependent on region, than version 22 data.

5.1.3. CALIOP

The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) differs from the other instruments here in that it is an active sensor. CALIOP measures vertical profiles of atmospheric backscatter and depolarization at 532 and 1,064 nm (Winker et al., 2009); aerosol/cloud features in these profiles are identified and classified by type, and extinction is retrieved, using a series of algorithms (Omar et al., 2009; Winker et al., 2009; Young & Vaughan, 2009). The fact that CALIOP is a lidar offers many important advantages such as a profiling capability, with a 100 m footprint; the ability to see in gaps between clouds; an insensitivity to 3-D radiative transfer effects, which can confound imager-based retrievals; the ability to provide data above snow/ice-covered surfaces; and the ability to provide aerosol data both during day and night. However, it also has some disadvantages: first, as a backscatter lidar, it (similarly to other sensors) has limited aerosol-related information content and must make estimate the extinction-to-backscatter ratio (lidar ratio) to estimate extinction and AOD. Second, as its measurements are nadir only, the revisit cycle for a given location is once per 16 days, and as the orbit is controlled tightly, sampling is sparse and many parts of the world are never observed.

The CALIOP L3 aerosol profile monthly product creation is described by Tackett et al. (2018) and is provided at 2° latitudinal by 5° longitudinal resolution. The current standard L3 products use version 3 L2 inputs; version 4 L2 products have been developed (Kim et al., 2018; Young et al., 2018), based on improved calibration (Getzewich et al., 2018; Kar et al., 2018) and addressing algorithmic issues which (in part) contributed to a low bias in version 3 data. However, version 4 inputs have not yet been incorporated into the standard L3 products. To take advantage of these version 4 updates, this analysis applies the same filtering and aggregation as described in Tackett et al. (2018) but with version 4 rather than version 3 L2 inputs. Using the standard version 3 CALIOP products (not shown) results in generally similar behavior and CALIOP AOD being lower by generally ~0.02 (albeit with some spatiotemporal variability). The resulting data set is otherwise conceptually equivalent to the standard L3 product composited from cloud-free profiles (CAL_LID_L3_APro_CloudFree), separately for daytime (D) and nighttime (N) orbital nodes. At present it is available through the end of 2016, and February 2016 is missing when data were unable to be acquired due to a temporary satellite anomaly.

These CALIOP L3 products only include aerosol features in the troposphere; the stratospheric AOD is not presently included in this data set. The resulting bias is systematically negative and varies with latitude but is small (<0.01; Friberg et al., 2018). The wavelength difference between CALIOP and VIIRS AOD is expected to cause a small systematic positive offset in CALIOP of 5% or less (dependent on AE), therefore \ll 0.01 in most cases, that is, fairly negligible.

5.1.4. POLDER GRASP

The Polarization and Directionality of the Earth's Reflectance (POLDER) instrument series has flown on three satellites; two were short lived, although the final operated from 2004 to 2013 within the A-Train constellation, providing nearly 2 years of overlap with VIIRS. POLDER's unique contribution is that it is not only multispectral and multiangular but also measures the polarization state of reflected light, which further increases the information content of the measurements for retrieval of surface and atmospheric properties (Hasekamp &

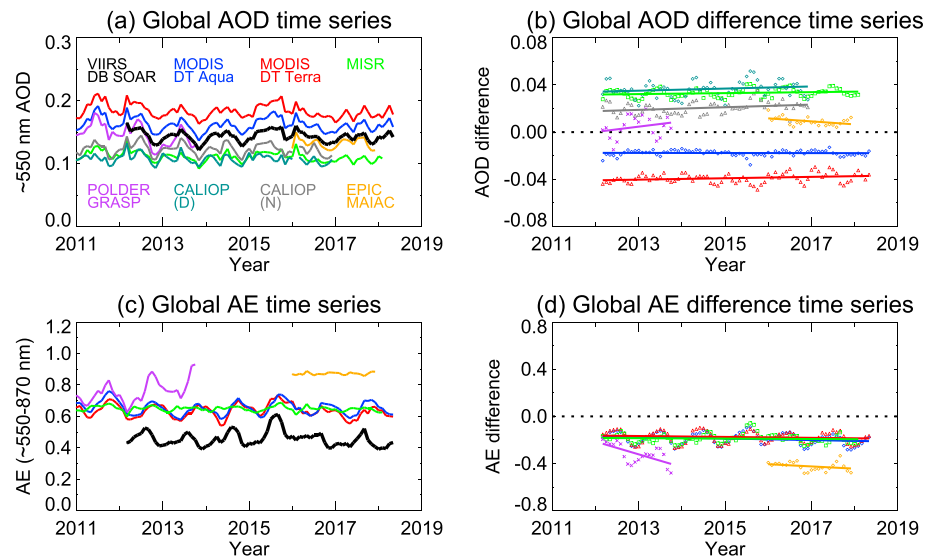


Figure 13. Time series (left column) of (a, b) AOD near 550 nm and (c, d) AE and VIIRS-other AOD and AE differences (right column), for over-water data from the period 2011–2018. The straight lines in the right panels are the linear fits to the monthly difference time series.

Landgraf, 2005). This high information content is exploited by the Generalized Retrieval of Aerosol and Surface properties (GRASP) algorithm, described by Dubovik et al. (2011, 2014). GRASP is built from the heritage of the AERONET inversion algorithm, retrieves (as the name suggests) aerosol and surface properties simultaneously, and makes no location-specific assumptions other than assignment of a pixel as land or water. Despite the short overlap with VIIRS, POLDER is of particular interest because the forthcoming multiviewing, multichannel, multipolarization (3MI) instrument (Marbach et al., 2015) builds upon the POLDER design. 3MI is planned to fly on multiple platforms from 2020 to 2040, and given the launch of additional VIIRS sensors in the future, the two are expected to eventually have considerable overlap.

This analysis uses the monthly product from the GRASP POLDER version 1.1 processing; note that POLDER AOD is defined at 565 nm rather than 550. Data are only provided at latitudes equatorward of 60°. The effect of this restriction on the global time series is expected to be small, for the twin reasons that area weighting reduces the importance of high-latitude regions in the global mean, and high-latitude over-water data are frequently unavailable due to cloud, snow, polar night, and continental land masses.

5.1.5. DSCOVER EPIC MAIAC

The Earth Polychromatic Imaging Camera (EPIC) aboard the Deep Space Climate Observatory (DSCOVER) platform is unique in that DSCOVER views the Earth from a Lagrange point, meaning that every point on the Sunlit portion of the globe is imaged from six (winter) to 11 (summer) times per day, at view angles close to backscatter (Marshak et al., 2018). This is different from both Sun-synchronous and geostationary orbits. EPIC measures at 10 wavelengths in the ultraviolet and visible, with a horizontal pixel size around 8 km at the subsatellite point and larger toward the edges of the Earth's disk. Except for the blue band, the actual resolution is coarser due to 2 × 2 pixel onboard aggregation. This is somewhat coarser than most other imagers, which has consequences for cloud masking and resulting data coverage. DSCOVER arrived at its orbital location in June 2015, and at present L2 and L3 products for 2016–2017 are available.

Here data from the Multiangle Implementation of Atmospheric Correction (MAIAC) algorithm applied to EPIC measurements are used. MAIAC employs time series analysis and spatial processing to improve cloud detection, aerosol retrieval, and atmospheric correction (Lyapustin et al., 2011). MAIAC was developed initially for MODIS; the EPIC implementation is built on the same principles, and the data are reprojected onto a 10-km sinusoidal grid before L2 processing. Note that MODIS MAIAC data are not used here because at present in MODIS water retrievals are limited to land-containing 1,200 km tiles, such that approximately half the global ocean is not covered. As the longest available EPIC wavelength is at 780 nm, the AE for this data set is defined over the wavelength range 550–780 nm. Retrieval from all available time slots are aggregated into the daily average, which is then averaged to provide monthly mean AOD and AE, at 1° horizontal resolution.

5.2. Global Analysis

Figure 13 shows time series of the global (area-weighted) mean over-water AOD and AE for each data set, as well as the difference between the VIIRS SOAR results and others. Global mean data are not necessarily useful for scientific analysis of the aerosol system but do provide a big picture overview of relative offsets between the data sets and whether or not they are stable in time.

For AOD, the temporal patterns of the time series track each other very closely. However, there is a range of about 0.08 in global mean AOD between the lowest (MISR and CALIOP D) and highest (MODIS Terra DT) in the time series; seasonality in the offsets is much smaller than the spread between the data sets. Small wavelength differences in AOD (from 532 to 565 nm) are likewise secondary. The SOAR results lie close to the middle (and generally in the range 0.13–0.16), with small (less than 0.01) offsets from the POLDER and EPIC retrievals. Realistically, the uncertainty on the true global mean AOD is likely to be smaller than this range of 0.08 implies. For example, CALIOP detection limits mean that ~ 0.02 – 0.04 of the background AOD may be undetected (Kacenenbogen et al., 2011; Toth et al., 2018). The offset between CALIOP D and N data is likely due to lower detection sensitivities caused by solar background noise in the daytime, although diurnal variability and day/night biases in aerosol typing could also contribute. On the other side, the SOAR and DT retrievals have a positive offset versus AERONET, with MODIS Terra more positive than Aqua (Levy et al., 2013; Sayer, Smirnov, Hsu, Munchak, & Holben, 2012). The VIIRS TOA reflectances used in SOAR are radiometrically cross calibrated against MODIS Aqua TOA reflectances, which decreases VIIRS-retrieved AOD by ~ 0.015 (Sayer, Hsu, Bettenhausen, et al., 2017). The VIIRS SOAR-MODIS Aqua DT offset here suggests that algorithmic contributions to differences between the data from these similar sensors are therefore of a similar magnitude. The POLDER, MISR, and EPIC data products require further evaluation in this regard. One caution to note with this interpretation is that due to sampling and aggregation, L2 validation biases may not directly correspond into L3 biases.

The AE comparison reveals a somewhat different picture. Interannual variability tracks less closely between the data sets, which is as expected given the more uncertain retrieval of AE compared to AOD. Note that an AE analog is not computed for CALIOP because of the limited validation of CALIOP AOD at 1,064 nm. For the MODIS and MISR products, AE varies between 0.6 and 0.7, compared to 0.4–0.6 for SOAR. Some of this offset is expected due to the negative offset of SOAR AE around -0.15 reported in section 3.3. However, the MODIS and MISR algorithms both provide an AOD/AE retrieval as the average of that from multiple candidate aerosol optical models (up to 20 for MODIS and up to 74 for MISR). When the AOD is low, a large proportion of the models are able to fit the TOA reflectances within assumed uncertainties; while this does not affect the retrieved AOD so strongly, this does mean that the reported AE will be the average of many possible solutions, which will diminish the variability of this parameter and weight toward more central values. In contrast, the VIIRS, POLDER, and EPIC data sets report a single retrieval solution. POLDER and EPIC AE are positively offset compared to MODIS/MISR. As with AOD, additional validation of these data products is necessary. As noted previously, the order of computation of AE (from monthly average spectral AOD vs. average of daily AE) does not change the big picture of these offsets (less than ± 0.1).

These mean offsets, and decadal slopes (i.e., trends) of fits to the time series of offsets, are summarized in Table 7. It is important to note that these trends in the offset indicate the relative trending between the data sets and not trends in the absolute AOD itself. The fidelity of a data set for trend calculation depends on many factors, including the magnitude of the trend relative to factors such as data set stability, interannual variability, and errors relating to retrieval problems or sampling incompleteness. AOD trends are not assessed in the present work. Uncertainties on the decadal AOD offset trends in Table 7 are one standard deviation (1σ) estimates, assuming (due to typical lifetimes of aerosol events/systems of order days to weeks) and accounting for a lag-1 month autocorrelation following Weatherhead et al. (1998). Few of the trends are significantly different from zero within 1σ and none within 2σ ; for the time series with most overlap with VIIRS (i.e., MODIS and MISR) the 2σ on decadal relative stability is of order 0.01 and 0.2 for AOD and AE, respectively. All four of these sensors have on-orbit calibration. This is only slightly higher for CALIOP, but it is somewhat larger for POLDER and EPIC, which is expected due to the short (~ 2 year) overlap in each case. Note that POLDER and EPIC also both lack onboard calibration, so it is possible that drifts in these sensors may also be larger. POLDER was calibration vicariously by observations of clouds, Sun glint, and stable desert sites (Hagolle et al., 1999), while EPIC visible bands are calibrated against MODIS Terra and Aqua (Marshak et al., 2018).

Table 7
Offsets and Decadal Trend ($\pm 1\sigma$) Between VIIRS and Other Data Sets in Global Area-Weighted Over-Water AOD and AE

Sensor/algorithm	AOD		AE	
	Offset	Trend	Offset	Trend
MODIS DT Aqua	-0.018	-0.0004 (± 0.003)	-0.19	-0.067 (± 0.11)
MODIS DT Terra	-0.039	0.006 (± 0.005)	-0.17	-0.039 (± 0.10)
MISR	0.033	0.004 (± 0.006)	-0.19	-0.017 (± 0.078)
POLDER GRASP	0.005	0.046 (± 0.10)	-0.32	-1.1 (± 0.99)
CALIOP (D)	0.037	0.009 (± 0.013)	—	—
CALIOP (N)	0.021	0.011 (± 0.009)	—	—
EPIC MAIAC	0.009	-0.026 (± 0.019)	-0.42	-0.19 (± 0.27)

Recently, Weatherhead et al. (2017) provided an assessment for how long satellites need to overlap to determine offsets or drifts to within a desired confidence. Although that work did not explicitly deal with aerosols, the same principles apply. One key point is that the length of overlap required is inversely proportional to the desired precision of the estimate. Based on this multisensor lack of divergence relative to MODIS and MISR, this suggests that the 2σ stability of the VIIRS data is ~ 0.01 in AOD and ~ 0.2 in AE or better, which is also broadly consistent with stability relative to AERONET (not shown). Halving these uncertainties on the estimates of the stability would thus be expected to require an approximate quadrupling of the overlapping data record lengths (unless significant decreases in noise and autocorrelation of the difference time series can be achieved). Given that the Terra and Aqua platforms are not expected to last another decade, it is unlikely that much improvement in these stability estimates can be made.

Global maps of AOD and AE, together with the MODIS Aqua daytime cloud fraction, are shown in Figure 14 for data from months of September during 2011–2017. MODIS Terra and CALIOP (N) data are omitted, as spatial patterns and values are very close to those of MODIS Aqua and CALIOP (D), respectively. September is shown as an illustrative month due to several interesting aerosol features (e.g., biomass burning in the tropics and dust outflow from North Africa and the Arabian Peninsula), plus the fact that as a month with an equinox it has somewhat balanced coverage of both northern and southern hemispheres. The general tendencies are similar for other months (not shown). The average of all available years in the period is shown because there is otherwise no overlap between POLDER and EPIC. This does mean that interannual variations could influence the composites for these two sensors. However, recently, Lee et al. (2018) used MODIS and MISR data to evaluate how many years were required to define a climatology of AOD within a sampling uncertainty (relative to full mission) of less than 0.01. They found that over the bulk of the open ocean, 2 years appeared sufficient, while in coastal outflow regions (e.g., dust/smoke belts) the time period was more often 4–12 years, dependent on region and distance from aerosol source. Monthly results were not shown (only annual), although these results suggest that, at least over the open ocean, interannual variability is not contributing significantly to differences for POLDER/EPIC AOD.

All data sets have the same general picture of AOD. Spatial variability reveals that differences in the background over-water AOD, which corresponds to the bulk of the globe, are likely to dominate the offsets seen in Figure 13. This indicates the need for better understanding, and ideally reducing, factors contributing to this open ocean difference. There are also differences in the intensity of high-AOD regions (e.g., African dust and smoke outflow). The differences do not appear to be largest in areas with the strongest cloud fraction (Figure 14l), although (particularly for POLDER and EPIC, which have larger pixel sizes so must screen clouds more aggressively to avoid contamination) data gaps are larger in cloudier regions.

For AE, there is greater diversity in spatial patterns. All data sets capture lower AE associated with outflow of Saharan dust into the tropical Atlantic. However, only VIIRS and POLDER show higher AE associated with African smoke outflow into the southern Atlantic, even though the enhanced AOD is reproduced broadly. The reasons for this are unclear. Except for this region, the spatial patterns of VIIRS, MODIS, and POLDER data are fairly similar, albeit with different baseline values and contrast. MISR and EPIC show diminished spatial

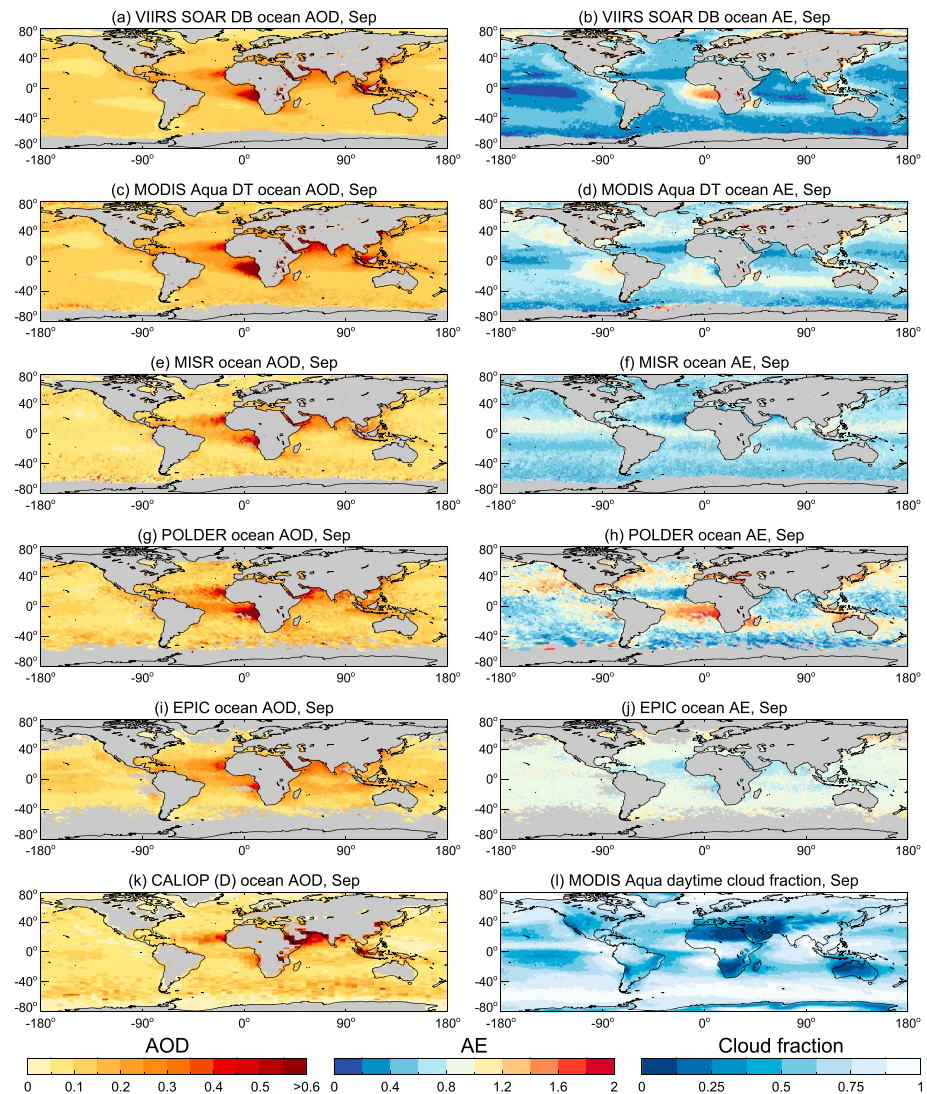


Figure 14. Multiannual (all available data from 2011 to 2017) average over-water AOD (a, c, e, g, i, k) and AE (b, d, f, h, j, l) for the month of September for assorted satellite AOD data sets. Also shown (bottom right) is the MODIS Aqua multiannual mean September daytime cloud fraction. Grid cells without data are shown in gray.

variability of AE compared to other data sets; for MISR, this is again consistent with the fact that the AE being reported may be the average AE from nearly all possible mixtures (and so algorithm decisions about averaging aimed at optimizing AOD retrieval may be harming the AE retrieval). Decreasing the considerable diversity in retrieved AE may be difficult because the bulk of the ocean consists of low-AOD regions where AE retrieval is subject to higher uncertainty. Note that the higher information content of multiangle and polarimetric measurements means that, in theory, MISR and POLDER should be more reliable for AE retrieval in low-AOD conditions than the other sensors. Continual improvement of sensor calibration, and realism of aerosol optical models, may produce larger changes in the diversity of AE than it does in AOD.

5.3. Regional Analysis

The borders of the regions chosen for multisensor time series comparison are shown in Figure 15; these encompass a range of different aerosol loadings and types, as well as different types of challenge to retrieval algorithms. Although all borders are somewhat subjective, manual testing finds that changing them by 5° typically changes AOD by 0.02 or less and AE by 0.1 or less and has a small influence on seasonality and the differences between the data sets. The resulting regional time series of AOD and AE are shown in Figure 16.

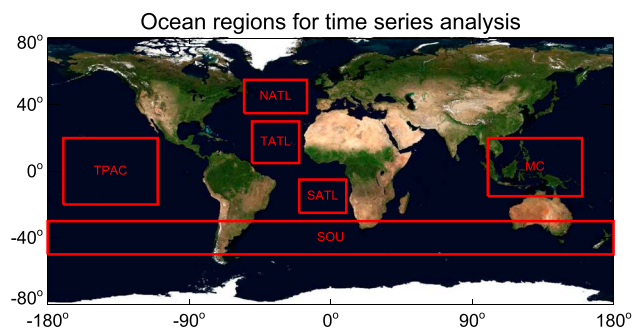


Figure 15. Regions used for multisatellite over-water aerosol optical depth and Ångström exponent time series comparison. Region names and bounds are the Tropical Pacific (TPAC; 20°S to 20°N, 170°W to 110°W), Tropical Atlantic (TATL; 5°N to 30°N, 50°W to 20°W), North Atlantic (NATL; 35°N to 55°N, 55°W to 15°W), South Atlantic (SATL; 25°S to 5°S, 20°W to 10°E), Southern Ocean (SOU; 50°S to 30°S, 180°W to 180°E), and Maritime Continent (MC; 15°S to 20°N, 100°E to 160°E).

The Tropical Pacific region is far from major aerosol sources so acts as a proxy for the behavior of the data sets in open ocean conditions. While the magnitude of AOD is slightly lower, the differences between the data sets, and the total diversity between them, are very similar to the global average results from Figure 13 and Table 7. Since AOD optical model assumptions tend to be less important to AOD retrievals when the AOD is low (as here), this implies that a large proportion of the variability is driven by fundamental factors such as calibration, surface reflectance models, and cloud screening/retrieval suitability choices. For example, Sayer, Hsu, Bettenhausen, et al. (2017) found that cross-calibrating VIIRS against MODIS Aqua decreased VIIRS-derived AOD by about 0.015 on average, Sayer, Thomas, and Grainger (2010) found that assumptions about glint and whitecaps could lead to systematic AOD retrieval changes around 0.01 following from small changes in assumed wind speed; Kaufman et al. (2005) estimated an overall thin cirrus contamination of ~ 0.01 in MODIS AOD; Zhao et al. (2013) found that increasing the strictness of cloud screening changes could decrease monthly mean AOD by 0.04. Indeed, empirical corrections developed for earlier versions of MODIS over-water aerosol

data by Zhang and Reid (2006) recognized this and included these factors in their parametrizations. Achieving consistency between AOD retrievals, even over the open ocean, traditionally seen as the most simple case, therefore requires additional efforts in these directions.

The Tropical Atlantic region samples maritime aerosol with, particularly in the spring and summer, large-scale transport of Saharan mineral dust. The peaks in AOD are well matched in both timing and relative intensities by all algorithms; the total diversity between data sets remains approximately 0.1. Unlike in the global and open ocean regions, CALIOP data are not on the low end of data sets but higher. The other positive outlier in these is generally MODIS DT; this is expected because the MODIS algorithm (unlike the others) assumes spherical dust aerosols, which is known to lead to geometry-dependent biases in AOD, which in this scattering angle range are generally positive and do not cancel out when averaging to monthly time scales (Lee et al., 2017). Sayer et al. (2018) found a similar positive offset in the NOAA VIIRS aerosol product (Jackson et al., 2013), based on the DT algorithm, in dust outflow regions. AE shows limited temporal variability in most data sets, which is as expected as both dust and sea spray aerosols are dominated by coarse (low AE) particles.

Both the North Atlantic and Southern Ocean are highly cloudy, with cloud fractions (calculated from MODIS Aqua daytime data) of ~ 0.8 – 0.9 and 0.7 – 0.8 , respectively (dependent on season). This limits the spatial coverage of the data sets and means that decisions about pixel suitability for retrieval via cloud masking or cloud adjacency may be more significant. In these regions the data diverge into two groups of differing AOD magnitude and seasonality: one is MODIS, VIIRS, POLDER, and EPIC (similar in seasonality but less so in magnitude), while the other is MISR and CALIOP. Note that expanding or contracting the latitudinal ranges of these regions does not alter these groupings, suggesting that they are not dominated by subregional variability. The former group tends to show larger AOD and AE in local summer months; for the latter group, MISR (as in other regions) shows very limited seasonality, while CALIOP does not provide AE. For North Atlantic it is conceivable that the summertime peak is due to aerosols transported from Europe or North America, which also exhibit a maximum of AOD in the summer (e.g., Remer et al., 2008). Transported fine mode aerosols would also explain the AE seasonality in this region. Although, if this is the explanation, it is curious why it is not seen by CALIOP or MISR AOD. For SOU, wind-driven sea spray provides one conceivable source and is observed in many global aerosol models and reanalyses (e.g., Bellouin et al., 2013; Colarco et al., 2010). However, Toth et al. (2013) used CALIOP and ship-based data to conclude that the bulk of this AOD enhancement is likely due to a combination of cloud contamination/adjacency effects and surface model inadequacies in this high-wind region for the imager-based retrievals, and many models are (directly or indirectly) tuned to these satellite retrievals.

The South Atlantic predominantly samples maritime aerosols, although from July to October it is also influenced by strongly absorbing smoke transported from biomass burning in central Africa (Edwards et al., 2006; Das et al., 2017), which is one of the main global biomass burning source regions (van der Werf et al., 2010). High cloud cover during the biomass burning season ($\sim 80\%$ from MODIS Aqua daytime data) also acts as a limiting factor for coverage, making this a complicated region for AOD retrieval. All sensors reproduce this expected seasonality in AOD, to a greater or lesser extent, although only POLDER and VIIRS SOAR see a strongly

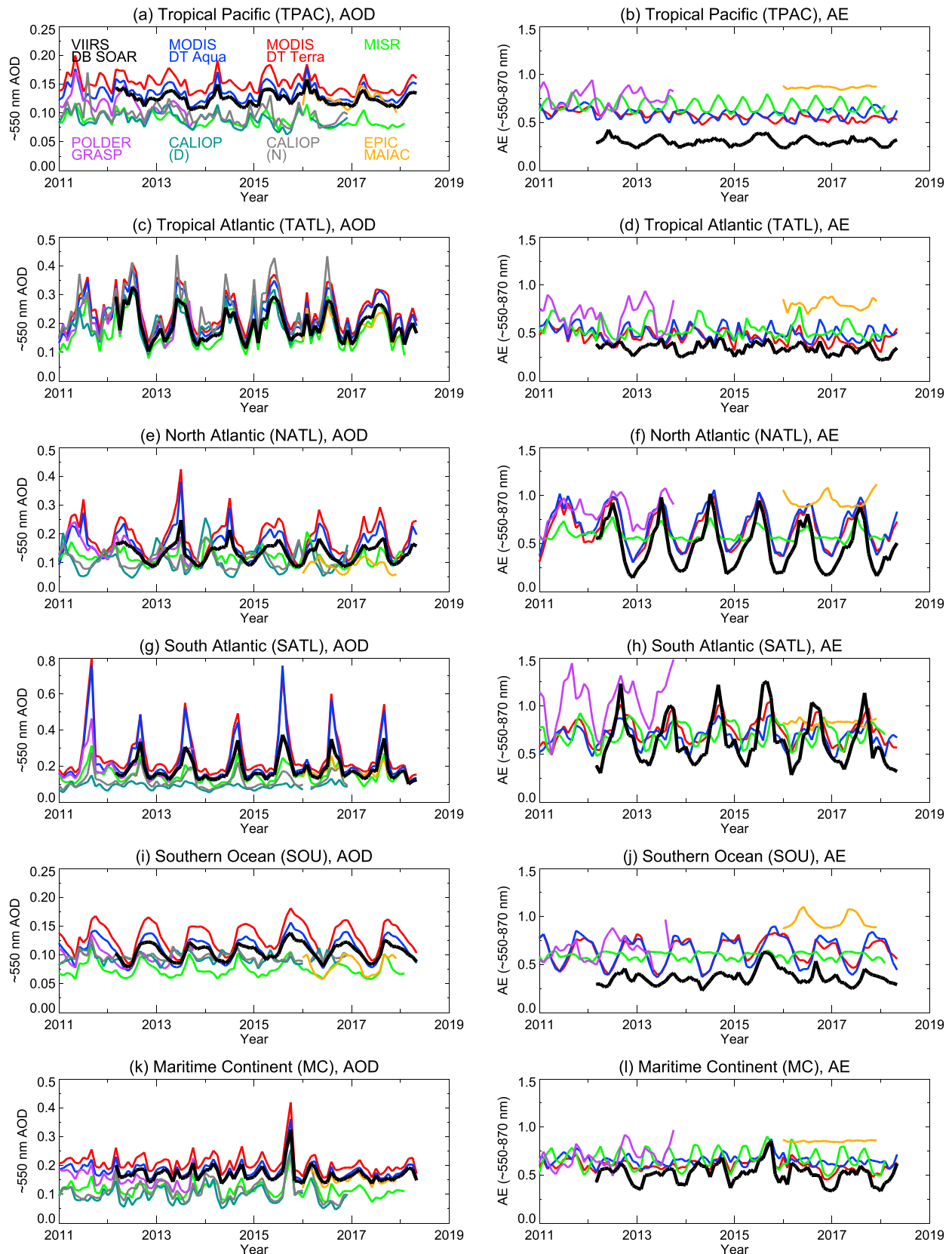


Figure 16. Time series of (a, c, e, g, i, k) AOD near 550 nm and (b, d, f, h, j, l) AE for the assorted satellite data sets, for the regions defined in Figure 15.

increased AE associated with this transported smoke during this period (MODIS seasonality is similar but diminished). POLDER and MISR see a second maximum in AE around November–December. Eck et al. (2013) reported a roughly linear change of SSA at 440 nm for smoke from this source region from ~ 0.83 in July to ~ 0.94 in November (similar pattern but lower SSA at longer wavelengths); aside from POLDER, all the retrievals have a fixed and limited set of optical models available, indicating the potential for changing biases during the biomass burning season. These SSA values are all more strongly absorbing than the set of optical models available to the retrieval algorithms (except POLDER); Sayer et al. (2014) validated older versions of the MODIS and MISR data sets using AERONET data at Ascension Island (8°S , 14°W , in the smoke outflow region) and found negative biases of magnitudes consistent with this SSA overestimation. For the same reason, and consistent with Section 3.4, the VIIRS data are expected to share this bias. It is therefore surprising that the MODIS biomass burning AOD peak is significantly stronger than seen in other data sets. Coupled with the weaker peak in AE and high cloud fraction, it is plausible that MODIS is providing retrievals closer to clouds, or is more affected by cloud masking, than the other data sets in this region. The weaker peaks in CALIOP are also surprising; the CALIOP lidar ratio adopted for smoke of 70 sr^{-1} at 532 nm (Liu et al., 2015; Omar et al., 2009) is close to that predicted from the optical model for smoke from this region developed in Sayer et al. (2014) of $\sim 75 \text{ sr}^{-1}$. If the non-CALIPSO retrievals are close to the true AOD, then either both lidar ratios are low by a factor of ~ 2 , which is very unlikely, or else factors such as failure to correctly detect the base of elevated smoke layers (Rajapakshe et al., 2017) or sampling limitations are causing large differences here.

The final region analyzed is the Maritime Continent, which is another region predominantly corresponding to clean maritime conditions but with seasonal contributions from transported smoke. From February to April a peak is observed associated with transport from biomass burning in southeastern Asia (e.g., Reid et al., 2013), while later in the year a peak corresponds to Indonesian biomass burning smoke. This latter was particularly severe in 2015, due to the strong El Niño causing widespread and intense fires (Field et al., 2016). For AOD this seasonality is captured by all data sets, albeit still with an offset around 0.1–0.15 between them; for AE, the signal is less clear, although elevated AE is seen for the 2015 Indonesian fire period. This is another region where high cloud cover (particularly of cirrus) limits observability (Reid et al., 2013); as the sensor most able to screen for this, it seems likely that CALIOP is giving a better representation of the true AOD during the clean periods than some other data sets. During intense smoke periods, imager-based retrievals are also known to overaggressively screen out thick smoke plumes and misclassify them as clouds, which also limits coverage (e.g., Shi et al., 2018, for MODIS).

6. Discussion and Conclusions

The analysis in this study has focused on two directions. The first is evaluating the performance of the VIIRS SOAR version 1 data set against AERONET from both the direct Sun and spectral deconvolution data sets. The second has been to place the VIIRS data set in context with other available satellite data sets and assess the level of consistency between them.

AERONET validation shows that the VIIRS data perform similarly to expectations, and similarly to the preliminary validation against shipborne observations presented alongside the algorithm by Sayer et al. (2018). Specifically, for the full retrieval algorithm the AOD uncertainty is about $\pm(0.03 + 10\%)$, increasing to $\pm(0.03 + 15\%)$ for retrievals performed in turbid/shallow waters with the backup algorithm. Globally, about 95% of retrievals are performed with the full algorithm. The overall bias in AOD is small and positive (~ 0.02), although in conditions of high-AOD loadings dominated by fine mode particles such as smoke, the data have a tendency for a low bias of $\sim 10\%$. These AOD error characteristics are comparable to those reported by other modern satellite AOD retrieval algorithms, establishing that the SOAR VIIRS data are suitable for the same types of scientific applications as them. The underlying radiometric uncertainty in spectral TOA reflectance measurements contributes to an expected floor in retrieval uncertainty over-water of AOD of ~ 0.01 , indicating that further refinement should be able to decrease this error. The consistency of the positive bias in low-AOD conditions with results from the application of SOAR to SeaWiFS and AVHRR (Sayer, Hsu, et al., 2012; Sayer, Hsu, Lee, et al., 2017) suggests that while sensor calibration plays a part, some of this error is likely due to algorithmic assumptions. The errors also remain slightly larger than the GCOS goals for an aerosol climate data record (the greater of 0.03 or 10%; GCOS, 2011), indicating that further refinement of all these algorithms, and ideally the availability of more advanced satellite sensors, is necessary. With the notable exception of Popp et al. (2016), GCOS compliance is not yet routinely reported in validation studies, and it would be helpful for these statistics to be included in the future.

Historically, FMF and AE have also not been validated as routinely as AOD. This is in part due to a lack of a robust ground truth, especially in the low-AOD conditions which predominate over the open ocean. Comparison with AERONET direct Sun and SDA data suggests that the VIIRS retrievals have a small negative offset (i.e., relatively too much AOD from the coarse mode) in low-AOD conditions. The AE offset largely disappears for smoke- or dust-dominated conditions, while the FMF difference does not, although in both cases the typical level of difference decreases with increasing AOD and plateaus around 0.2 (for AE) and 0.14 (for FMF), respectively. These are again similar to or better than the few published evaluations of these quantities from other data sets. Using an optical-based classification of aerosol type from AERONET and comparing against the optical model retrieved by SOAR shows agreement a little better than 80% of the time. Together, these suggest the potential of the data for more advanced applications, for example, data assimilation of not just total AOD but also the fine/coarse partition or masking by type to identify columns dominated by dust storms. This is an advantage of the SOAR algorithm for applications which require categorical classification of aerosol air masses or wish to perform a forward radiative transfer calculation consistent with retrieval assumptions. Some other approaches (e.g., MODIS DT and MISR) instead provide as standard an average of multiple plausible solutions rather than a single self-consistent optical model. Averaging solutions can mitigate some sources of AOD bias and also tends to artificially smooth derived products like AE, and an averaged solution cannot so directly be used for a forward calculation.

The second part of the analysis reveals that, despite the similar quality of AOD retrieved by individual data sets, there remains a large diversity of about 0.08 in regional and global overocean AOD and significantly larger diversity in AE. Much of this diversity is attributable to known issues with individual sensors and algorithms while considerable uncertainty is likely a result of differences in sampling and aggregation from L2 to L3. These sampling effects have not yet received much attention, with most studies using L3 data uncritically, although in recent years steps have been taken to begin quantifying them (e.g., Sayer, Thomas, Palmer, Grainger, 2010; Schutgens et al., 2016, 2017).

About half the uncertainty in indirect aerosol forcing is from lack of knowledge of the preindustrial natural aerosol background (Carslaw et al., 2013). This is in large part due to the interactions between aerosols and clouds (Meskhidze & Nenes, 2010), although the effect of an aerosol perturbation on the cloud system is dependent on the baseline level of aerosols. To decrease these uncertainties, reconciling these offsets between sensors is a key part. As well as continual assessment of sensor calibration, surface models, and optical models, often underappreciated factors which may introduce systematic offsets such as L3 aggregation or correction for trace gas absorption (Patadia et al., 2018) should be given more prominence. Regional AOD biases from thin cirrus contamination (or overzealous screening) have long been known to be an issue (e.g., Kaufman et al., 2005), although Lee et al. (2013) present a method for preretrieval correction for this, decreasing retrieval errors and increasing data set coverage. Pierce et al. (2010) also demonstrated the potential of MISR to detect and retrieve cirrus cloud optical depth simultaneously with aerosols and separate the contributions. It would be beneficial to further assess and implement corrective techniques into L2 processing algorithms.

Returning specifically to SOAR, the VIIRS data are roughly in the middle of the range of AOD from the various satellite data products. While AE has a low offset from other sensors, validation results indicate that AE and FMF both have quantitative skill. Together, the comparison with AERONET and other satellite products identifies no drift in the AOD, with the two standard deviation confidence level on a detectable drift likely to be around 0.01 per decade. SOAR has already been applied to the SeaWiFS and a subset of AVHRR sensors. Further refining these algorithms and continuing application of SOAR forward in time with VIIRS on the S-NPP platform and follow-on U.S. Joint Polar Satellite System platforms, and backward in time with MODIS, will eventually provide an ~40-year data record. The use of a consistent algorithmic approach (as far as possible given differing sensor characteristics) combined with careful attention to radiometric calibration, and taking advantage of the multiple years of overlap between satellite sensors, should mean that this record is able to overcome some difficulties with current trend studies (Li et al., 2009).

Acknowledgments

More information about the Deep Blue aerosol project, including further documentation and links, can be found at <https://deepblue.gsfc.nasa.gov>. This research was funded by NASA's radiation science programme, managed by Hal Maring. The Atmosphere SIPS at the University of Wisconsin are thanked for VIIRS data processing. Work at the Atmosphere SIPS is being performed for NASA under contract NNG15HZ38C. AERONET data are available from <https://aeronet.gsfc.nasa.gov>; the AERONET team and site Principal Investigators and managers are thanked for the creation and maintenance of the AERONET data record, which is central to the continued assessment of remotely sensed and modeled aerosol data. MODIS and VIIRS data are available from the NASA LAADS at <https://ladsweb.nascom.nasa.gov>. CALIOP and MISR data are available from the NASA Langley Research Center Atmospheric Science Data Center at <https://eosweb.larc.nasa.gov>; B. Getzewich, K. Beaumont, and T. Murray (NASA LaRC/SSAI) are thanked for assistance with the special CALIOP processing. The GRASP code and POLDER data are available from <https://www.grasp-open.com>. EPIC MAIAC data are available on request to A. Lyapustin (NASA GSFC). All data servers are acknowledged for the hosting of these data sets and supporting documentation. The authors thank M. J. Garay (JPL), R. A. Kahn (NASA GSFC), O. Kalashnikova (JPL), R. C. Levy (NASA GSFC), V. Sawyer (NASA GSFC/SSAI), Y. Shi (NASA GSFC/USRA), and attendees at NASA's AeroCenter seminar series for useful discussions about the different satellite AOD products and time series. We thank three anonymous reviewers for their comments and suggestions about this manuscript. The authors declare no conflicts of interest.

References

Bellouin, N., Quaas, J., Morcrette, J.-J., & Boucher, O. (2013). Estimates of aerosol radiative forcing from the MACC re-analysis. *Atmospheric Chemistry and Physics*, *13*, 2045–2062. <https://doi.org/10.5194/acp-13-2045-2013>

Benedetti, A., Reid, J. S., Knippertz, P., Marsham, J. H., Di Giuseppe, F., Rémy, S., et al. (2018). Status and future of numerical atmospheric aerosol prediction with a focus on data requirements. *Atmospheric Chemistry and Physics*, *18*, 10,615–10,643. <https://doi.org/10.5194/acp-18-10615-2018>

Carlaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., et al. (2013). Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, *503*, 67–71. <https://doi.org/10.1038/nature12674>

Colarco, P., da Silva, A., & Diehl, T. (2010). Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *Journal of Geophysical Research*, *115*, D14207. <https://doi.org/10.1029/2009JD012820>

Das, S., Harshvardhan, H., Bian, H., Chin, M., Curci, G., Protonotariou, A. P., et al. (2017). Biomass burning aerosol transport and vertical distribution over the South African-Atlantic region. *Journal of Geophysical Research: Atmospheres*, *122*, 6391–6415. <https://doi.org/10.1002/2016JD026421>

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

Dubovik, O., & King, M. D. (2000). A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. *Journal of Geophysical Research*, *105*(D16), 20,673–20,696. <https://doi.org/10.1029/2000JD900282>

Dubovik, O., Lapyonok, T., Litvinov, P., Herman, M., Fuertes, D., Ducos, F., et al. (2014). GRASP: A versatile algorithm for characterizing the atmosphere. Newsroom: SPIE. <https://doi.org/10.1117/2.1201408.005558>

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B., Mishchenko, M., Yang, P., et al. (2006). The application of spheroid models to account for aerosol particle non-sphericity in remote sensing of desert dust. *Journal of Geophysical Research*, *111*, D11208. <https://doi.org/10.1029/2005JD006619>

Dubovik, O., Smirnov, A., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F., & Slutsker, I. (2006). Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements. *Journal of Geophysical Research*, *105*, 9791–9806. <https://doi.org/10.1029/2000JD900040>

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*(D24), 31,333–31,349.

Eck, T. F., Holben, B. N., Reid, J. S., Mukelabai, M. M., Piketh, S. J., Torres, O., et al. (2013). A seasonal trend of single scattering albedo in southern African biomass-burning particles: Implications for satellite products and estimates of emissions for the world's largest biomass-burning source. *Journal of Geophysical Research: Atmospheres*, *118*, 6414–6432. <https://doi.org/10.1002/jgrd.50500>

Edwards, D. P., Emmons, L. K., Gille, J. C., Chu, A., Attié, J.-L., Wood, S. W., et al. (2006). Satellite-observed pollution from Southern Hemisphere biomass burning. *Journal of Geophysical Research*, *111*, D14312. <https://doi.org/10.1029/2005JD006655>

Eleftheratos, K., Zerefos, C. S., Varotsos, C., & Kapsomenakis, I. (2011). Interannual variability of cirrus clouds in the tropics in El Niño Southern Oscillation (ENSO) regions based on International Satellite Cloud Climatology Project (ISCCP) satellite data. *International Journal of Remote Sensing*, *32*(21), 6395–6405. <https://doi.org/10.1080/01431161.2010.510491>

Field, R. D., van der Werd, G. R., Fanin, T., Fetzer, E. J., Fuller, R., Jethva, H., et al. (2016). Indonesian fire activity and smoke pollution in 2015 show persistent nonlinear sensitivity to El Niño-induced drought. *PNAS*, *113*(33), 9204–9209. <https://doi.org/10.1073/pnas.1524888113>

Friberg, J., Martinsson, B. G., Andersson, S. M., & Sandvik, O. S. (2018). Volcanic impact on the climate—The stratospheric aerosol load in the period 2006–2015. *Atmospheric Chemistry and Physics*, *18*(15), 11,149–11,169. <https://doi.org/10.5194/acp-2017-1200>

GCOS (2011). Systematic observation requirements for satellite-based data products for climate, 2011 update (report GCOS-154): world Meteorological Organization (WMO) Global Climate Observing System (GCOS).

Getzewich, B. J., Kar, J., Vaughan, M. A., Lee, K.-P., Tackett, J. L., Avery, M. A., et al. (2018). CALIPSO lidar calibration at 532 nm: Version 4 nighttime algorithm. *Atmospheric Measurement Techniques Discussions*, *11*, 1459–1479. <https://doi.org/10.5194/amt-2018-206>

Hagolle, O., Goloub, P., Deschamps, P.-Y., Cosnefroy, H., Briottet, X., Baillet, T., et al. (1999). Results of POLDER in-flight calibration. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(3), 1550–1566. <https://doi.org/10.1109/36.763266>

Hasekamp, O. P., & Landgraf, J. (2005). Retrieval of aerosol properties over the ocean from multispectral single-viewing-angle measurements of intensity and polarization: Retrieval approach, information content, and sensitivity study. *Journal of Geophysical Research*, *110*, D20207. <https://doi.org/10.1029/2005JD006212>

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*, 1–16. [https://doi.org/10.1016/S0034-4257\(98\)00031-5](https://doi.org/10.1016/S0034-4257(98)00031-5)

Hsu, N. C., Jeong, M.-J., Bettenhausen, C., Sayer, A. M., Hansell, R., Seftor, C. S., et al. (2013). Enhanced Deep Blue aerosol retrieval algorithm: The second generation. *Journal of Geophysical Research: Atmospheres*, *118*, 9296–9315. <https://doi.org/10.1002/jgrd.50712>

Hsu, N. C., Lee, J., Sayer, A. M., Carletta, N., Chen, S.-H., Tucker, C. J., et al. (2017). Retrieving near-global aerosol loading over land and ocean from AVHRR. *Journal of Geophysical Research: Atmospheres*, *122*, 9968–9989. <https://doi.org/10.1002/2017JD026932>

Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., et al. (2002). A spatio-temporal approach for global validation and analysis of MODIS aerosol products. *Geophysical Research Letters*, *29*(12), 8006. <https://doi.org/10.1029/2001GL013206>

Jackson, J. M., Liu, H., Laszlo, I., Kondragunta, S., Remer, L. A., Huang, J., & Huang, H.-C. (2013). Suomi-NPP VIIRS aerosol algorithms and data products. *Journal of Geophysical Research: Atmospheres*, *118*, 12,673–12,689. <https://doi.org/10.1002/2013JD020449>

Jeong, M.-J., & Li, Z. (2005). Quality, compatibility, and synergy analyses of global aerosol products derived from the advanced very high resolution radiometer and total ozone mapping spectrometer. *Journal of Geophysical Research*, *110*, D10508. <https://doi.org/10.1029/2004JD004647>

Kacenenbogen, M., Vaughan, M. A., Redemann, J., Hoff, R. M., Rogers, R. R., Ferrare, R. A., et al. (2011). An accuracy assessment of the CALIOP/CALIPSO version 2/version 3 daytime aerosol extinction product based on a detailed multi-sensor, multi-platform case study. *Atmospheric Chemistry and Physics*, *11*, 3981–4000. <https://doi.org/10.5194/acp-11-3981-2011>

Kahn, R. A., & Gaitley, B. J. (2015). An analysis of global aerosol type as retrieved by MISR. *Journal of Geophysical Research: Atmospheres*, *120*, 4248–4281. <https://doi.org/10.1002/2015JD023322>

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., 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>

Kar, J., et al. (2018). CALIPSO lidar calibration at 532 nm: Version 4 nighttime algorithm. *Atmospheric Measurement Techniques*, *11*, 1459–1479. <https://doi.org/10.5194/amt-11-1459-2018>

- Kaufman, Y. J., Remer, L. A., Tanré, D., Li, R.-R., Kleidman, R., Mattoo, S., et al. (2005). A critical examination of the residual cloud contamination and diurnal sampling effects on modis estimates of aerosol over ocean. *IEEE Transactions on Geoscience and Remote Sensing*, 43(12), 2886–2897. <https://doi.org/10.1109/TGRS.2005.858430>
- Kim, M.-H., Omar, A. H., Tackett, J. L., Vaughan, M. A., Winker, D. M., Trepte, C. R., et al. (2018). The CALIPSO version 4 automated aerosol classification and lidar ratio selection algorithm. *Atmospheric Measurement Techniques*, 11, 6107–6135. <https://doi.org/10.5194/amt-11-6107-2018>
- Kleidman, R. G., O'Neill, N. T., Remer, L. A., Kaufman, Y. J., Eck, T. F., Tanré, D., et al. (2005). Comparison of Moderate Resolution Imaging Spectroradiometer (MODIS) and Aerosol Robotic Network (AERONET) remote-sensing retrievals of aerosol fine mode fraction over ocean. *Journal of Geophysical Research*, 110, D22205. <https://doi.org/10.1029/2005JD005760>
- Kokhanovsky, A. A., & de Leeuw, G. (Eds.) (2009). *Satellite aerosol remote sensing over land*. Berlin: Springer. <https://doi.org/10.1007/978-3-540-69397-0>
- Lee, H., Garay, M. J., Kalashnikova, O. V., Yu, Y., & Gibson, P. B. (2018). How long should the MISR record be when evaluating aerosol optical depth climatology in climate models? *Remote Sensing*, 10(9), 1326. <https://doi.org/10.3390/rs10091326>
- Lee, J., Hsu, N. C., Bettenhausen, C., & Sayer, A. M. (2013). Retrieval of aerosol optical depth under thin cirrus from MODIS: Application to an ocean algorithm. *Journal of Geophysical Research: Atmospheres*, 118, 10,111–10,124. <https://doi.org/10.1002/jgrd.50806>
- Lee, J., Hsu, N. C., Sayer, A. M., Bettenhausen, C., & Yang, P. (2017). AERONET-based nonspherical dust optical models and effects on the VIIRS Deep Blue/SOAR over-water aerosol product. *Journal of Geophysical Research: Atmospheres*, 122, 10,384–10,401. <https://doi.org/10.1002/2017JD027258>
- Lenoble, J., Remer, L. A., & Tanré, D. (Eds.) (2013). *Aerosol remote sensing Edited by Lenoble, J., Remer, L. A., & Tanré, D.* Berlin: Springer. <https://doi.org/10.1007/978-3-642-17725-5>
- 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., Mattoo, S., Sawyer, V., Shi, Y., Colarco, P. R., Lyapustin, A. I., et al. (2018). Exploring systematic offsets between aerosol products from the two MODIS sensors. *Atmospheric Measurement Techniques*, 11, 4073–4092. <https://doi.org/10.5194/amt-11-4073-2018>
- Levy, R. C., Munchak, L. A., Mattoo, S., Patadia, F., Remer, L. A., & Holz, R. E. (2015). Towards a long-term global aerosol optical depth record: Applying a consistent aerosol retrieval algorithm to MODIS and VIIRS-observed reflectance. *Atmospheric Measurement Techniques*, 8, 4083–4110. <https://doi.org/10.5194/amt-8-4083-2015>
- 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, 2755–2770. <https://doi.org/10.5194/angeo-27-2755-2009>
- Liu, Z., Winker, D., Omar, A., Vaughan, M., Kar, J., Trepte, C., et al. (2015). Evaluation of CALIOP 532 nm aerosol optical depth over opaque water clouds. *Atmospheric Chemistry and Physics*, 15, 1265–1288. <https://doi.org/10.5194/acp-15-1265-2015>
- Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korokin, S., Remer, L., et al. (2011). Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *Journal of Geophysical Research*, 116, D03211. <https://doi.org/10.1029/2010JD014986>
- Marbach, T., Riedi, J., Lacan, A., & Schlüssel, P. (2015). The 3MI mission: Multi-viewing-channel-polarisation imager of the EUMETSAT polar system: Second generation (EPS-SG) dedicated to aerosol and cloud monitoring. In *Proceedings of SPIE* (Vol. 9613, 8 pp.) San Diego, CA. <https://doi.org/10.1117/12.2186978>
- Marshak, A., Herman, J., Szabo, A., Blank, K., Cede, A., Carn, S., et al. (2018). *Earth observations from DISCOVER/EPIC instrument* (Vol. 99, pp. 1829–1850). <https://doi.org/10.1175/BAMS-D-17-0223.1>
- Martonchik, J. V., Diner, D. J., Kahn, R. A., Ackerman, T. P., Verstraete, M. M., Pinty, B., & Gordon, H. R. (1998). Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4), 1212–1227. <https://doi.org/10.1109/36.701027>
- McFarquhar, G. M., Heymsfield, A. J., Spinhirne, J., & Hart, B. (2000). Thin and subvisual tropopause tropical cirrus: Observations and radiative impacts. *Journal of the Atmospheric Sciences*, 57(12), 1841–1853. [https://doi.org/10.1175/1520-0469\(2000\)057<1841:TASTTC>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<1841:TASTTC>2.0.CO;2)
- Meskhidze, N., & Nenes, A. (2010). Effects of ocean ecosystem on marine aerosol-cloud interactions. *Advances in Meteorology*, 13, 1–13. <https://doi.org/10.1155/2010/239808>
- O'Neill, N. T., Dubovik, O., & Eck, T. F. (2001). Modified ångström coefficient for the characterization of submicrometer aerosols. *Applied Optics*, 40(15), 2368–2375. <https://doi.org/10.1364/AO.40.002368>
- O'Neill, N. T., Eck, T. F., Smirnov, A., Holben, B. N., & Thulasiraman, S. (2003). Spectral discrimination of coarse and fine mode optical depth. *Journal of Geophysical Research*, 108(D17), 4559–4573. <https://doi.org/10.1029/2002JD002975>
- O'Neill, N., Eck, T., Smirnov, A., Holben, B., & Thulasiraman, S. (2006). Spectral deconvolution algorithm technical memo (Tech. rep.). Greenbelt, MD: NASA Goddard Space Flight Center. revision April 26, 2006, version 4 available online from http://aeronet.gsfc.nasa.gov/new_web/PDF/tauf_tauc_technical_memo1.pdf [Accessed December 2015].
- Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. A., Ferrare, R. A., et al. (2009). The CALIPSO automated aerosol classification and lidar ratio selection algorithm. *Journal of Atmospheric and Oceanic Technology*, 26, 1994–2014. <https://doi.org/10.1175/2009JTECHA1231.1>
- Patadia, F., Levy, R. C., & Mattoo, S. (2018). Correcting for trace gas absorption when retrieving aerosol optical depth from satellite observations of reflected shortwave radiation. *Atmospheric Measurement Techniques*, 11, 3205–3219. <https://doi.org/10.5194/amt-11-3205-2018>
- Penning de Vries, M. J. M., Beirle, S., Hörmann, C., Kaiser, J. W., Stammes, P., Tilstra, L. G., et al. (2015). A global aerosol classification algorithm incorporating multiple satellite data sets of aerosol and trace gas abundances. *Atmospheric Chemistry and Physics*, 15, 10,597–10,618. <https://doi.org/10.5194/acp-15-10597-2015>
- Pierce, J. R., Kahn, R. A., Davis, M. R., & Comstock, J. M. (2010). Detecting thin cirrus in Multiangle Imaging Spectroradiometer aerosol retrievals. *Journal of Geophysical Research*, 115, D08201. <https://doi.org/10.1029/2009JD013019>
- Popp, T., de Leeuw, G., Bingen, C., Brühl, C., Capelle, V., Chedin, A., et al. (2016). Development, production and evaluation of aerosol climate data records from European satellite observations (Aerosol_cci). *Remote Sensing*, 8(5), 421. <https://doi.org/10.3390/rs8050421>
- Prospero, J. M., Collard, F.-X., Molinié, J., & Jeannot, A. (2014). Characterizing the annual cycle of African dust transport to the Caribbean Basin and South America and its impact on the environment and air quality. *Global Biogeochemical Cycles*, 28, 757–773. <https://doi.org/10.1002/2013GB004802>
- Rajapakshe, C., Zhang, Z., Yorks, J. E., Yu, H., Tan, Q., Meter, K., et al. (2017). Seasonally transported aerosol layers over southeast Atlantic are closer to underlying clouds than previously reported. *Geophysical Research Letters*, 44, 5818–5825. <https://doi.org/10.1002/2017GL073559>

- Reid, J. S., Eck, T. F., Christopher, S. A., Koppmann, R., Dubovik, O., Eleuterio, D. P., et al. (2005). A review of biomass burning emissions part III: Intensive optical properties of biomass burning particles. *Atmospheric Chemistry and Physics*, 5, 827–849. <https://doi.org/10.5194/acp-5-827-2005>
- Reid, J. S., Hyer, E. J., Johnson, R. S., Holben, B. N., Yokelson, R. J., Zhang, J., et al. (2013). Observing and understanding the Southeast Asian aerosols system by remote sensing: An initial review and analysis for the Seven Southeast Asian Studies (7 SEAS) program. *Atmospheric Research*, 122, 303–468. <https://doi.org/10.1016/j.atmosres.2012.06.005>
- Reid, J. S., Koppmann, R., Eck, T. F., & Eleuterio, D. P. (2005). A review of biomass burning emissions part II: Intensive physical properties of biomass burning particles. *Atmospheric Chemistry and Physics*, 5, 799–825. <https://doi.org/10.5194/acp-5-799-2005>
- Remer, L. A., Kleidman, R. G., Levy, R. C., Kaufman, Y. J., Tanré, D., Mattoo, S., et al. (2008). Global aerosol climatology from the MODIS satellite sensors. *Journal of Geophysical Research*, 113, D14S07. <https://doi.org/10.1029/2007JD009661>
- Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: Theory and Practice*, Series on Atmospheric, Oceanic and Planetary Physics-Vol. 2. London, UK: World Scientific.
- Russell, P. B., Kacenelenbogen, M., Livingston, J. M., Hasekamp, O. P., Burton, S. P., Schuster, G. L., et al. (2014). A multiparameter aerosol classification method and its application to retrievals from spaceborne polarimetry. *Journal of Geophysical Research: Atmospheres*, 119, 9838–9863. <https://doi.org/10.1002/2013JD021411>
- Sayer, A. M., Hsu, N. C., Bettenhausen, C., Ahmad, Z., Holben, B. N., Smirnov, A., et al. (2012). SeaWiFS Ocean Aerosol Retrieval (SOAR): Algorithm, validation, and comparison with other data sets. *Journal of Geophysical Research*, 117, D03206. <https://doi.org/10.1029/2011JD016599>
- Sayer, A. M., Hsu, N. C., Bettenhausen, C., Holz, R. E., Lee, J., Quinn, G., & Veglio, P. (2017). Cross-calibration of S-NPP VIIRS moderate-resolution reflective solar bands against MODIS Aqua over dark water scenes. *Atmospheric Measurement Techniques*, 10, 1425–1444. <https://doi.org/10.5194/amt-10-1425-2017>
- Sayer, A. M., Hsu, N. C., Bettenhausen, C., & Jeong, M.-J. (2013). Validation and uncertainty estimates for MODIS Collection 6 "Deep Blue" aerosol data. *Journal of Geophysical Research: Atmospheres*, 118, 7864–7872. <https://doi.org/10.1002/jgrd.50600>
- Sayer, A. M., Hsu, N. C., Eck, T. F., Smirnov, A., & Holben, B. N. (2014). AERONET-based models of smoke-dominated aerosol near source regions and transported over oceans, and implications for satellite retrievals of aerosol optical depth. *Atmospheric Chemistry and Physics*, 14, 11,493–11,523. <https://doi.org/10.5194/acp-14-11493-2014>
- Sayer, A. M., Hsu, N. C., Lee, J., Bettenhausen, C., Kim, W. V., & Smirnov, A. (2018). Satellite ocean aerosol retrieval (SOAR) algorithm extension to S-NPP VIIRS as part of the 'Deep Blue' aerosol project. *Journal of Geophysical Research: Atmospheres*, 123, 380–400. <https://doi.org/10.1002/2017JD027412>
- 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., Smirnov, A., Hsu, N. C., & Holben, B. N. (2012). A pure marine aerosol model, for use in remote sensing applications. *Journal of Geophysical Research*, 117, D05213. <https://doi.org/10.1029/2011JD016689>
- Sayer, A. M., Smirnov, A., Hsu, N. C., Munchak, L. A., & Holben, B. N. (2012). Estimating marine aerosol particle volume and number from Maritime Aerosol Network data. *Atmospheric Chemistry and Physics*, 12, 8889–8909. <https://doi.org/10.5194/acp-12-8889-2012>
- Sayer, A. M., Thomas, G. E., & Grainger, R. G. (2010). A sea surface reflectance model for (A)ATSR, and application to aerosol retrievals. *Atmospheric Measurement Techniques*, 3, 813–838. <https://doi.org/10.5194/amt-3-813-2010>
- Sayer, A. M., Thomas, G. E., Palmer, P. I., & Grainger, R. G. (2010). Some implications of sampling choices on comparisons between satellite and model aerosol optical depth fields. *Atmospheric Chemistry and Physics*, 10, 10,705–10,716. <https://doi.org/10.5194/acp-10-10705-2010>
- Schuster, G. L., Dubovik, O., & Holben, B. N. (2006). Angstrom exponent and bimodal aerosol size distributions. *Journal of Geophysical Research*, 111, D07207. <https://doi.org/10.1029/2005JD006328>
- Schutgens, N. A. J., Gryspeerdt, E., Weigum, N., Tsyro, S., Goto, D., Schulz, M., & Stier, P. (2016). Will a perfect model agree with perfect observations? The impact of spatial sampling. *Atmospheric Chemistry and Physics*, 16, 6335–6353. <https://doi.org/10.5194/acp-16-6335-2016>
- Schutgens, N., Tsyro, S., Gryspeerdt, E., Goto, D., Weigum, N., Schulz, M., & Stier, P. (2017). On the spatio-temporal representativeness of observations. *Atmospheric Chemistry and Physics*, 17, 9761–9780. <https://doi.org/10.5194/acp-17-9761-2017>
- Shi, Y. R., Levy, R. C., Eck, T. F., Fisher, B., Mattoo, S., Remer, L. A., et al. (2018). Characterizing the 2015 Indonesia fire event using modified MODIS aerosol retrievals. *Atmospheric Chemistry and Physics Discussions*, 1–26. <https://doi.org/10.5194/acp-2018-468>
- Shi, Y., Zhang, J., Reid, J. S., Holben, B. N., Hyer, E. J., & Curtis, C. (2011). An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation. *Atmospheric Chemistry and Physics*, 11, 557–565. <https://doi.org/10.5194/acp-11-557-2011>
- Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., & Slutsker, I. (2000). Cloud-screening and quality control algorithms for the AERONET database. *Remote Sensing of Environment*, 73(3), 337–349.
- Smirnov, A., Holben, B. N., Slutsker, I., Giles, D. M., McClain, C. R., Eck, T. F., et al. (2009). Maritime aerosol network as a component of aerosol robotic network. *Journal of Geophysical Research*, 112, D06204. <https://doi.org/10.1029/2008JD011257>
- Tackett, J. L., Winker, D. M., Getzewich, B. J., Vaughan, M. A., Young, S. A., & Kar, J. (2018). CALIPSO lidar level 3 aerosol profile product: Version 3 algorithm design. *Atmospheric Measurement Techniques*, 11, 4129–4152. <https://doi.org/10.5194/amt-11-4129-2018>
- Tanré, D., Kaufman, Y. J., Herman, M., & Mattoo, S. (1997). Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances. *Journal of Geophysical Research*, 102(D14), 16,971–16,988. <https://doi.org/10.1029/96JD03437>
- Toller, G., Xiong, X., Sun, J., Wenny, B. N., Geng, X., Kuypers, J., et al. (2013). Terra and Aqua Moderate-Resolution Imaging Spectroradiometer collection 6 level 1B algorithm. *Journal of Applied Remote Sensing*, 7(1), 73557–1–73557-17. <https://doi.org/10.1117/1.JRS.7.073557>
- Toth, T. D., Campbell, J. R., Reid, J. S., Tackett, J. L., Vaughan, M. A., Zhang, J., & Marquis, J. W. (2018). Minimum aerosol layer detection sensitivities and their subsequent impacts on aerosol optical thickness retrievals in CALIPSO level 2 data products. *Atmospheric Measurement Techniques*, 11, 499–514. <https://doi.org/10.5194/amt-11-499-2018>
- Toth, T. D., Zhang, J., Campbell, J. R., Reid, J. S., Shi, Y., Johnson, R. S., et al. (2013). Investigating enhanced Aqua MODIS aerosol optical depth retrievals over the mid-to-high latitude Southern Oceans through intercomparison with co-located CALIOP, MAN, and AERONET data sets. *Journal of Geophysical Research: Atmospheres*, 118, 4700–4714. <https://doi.org/10.1002/jgrd.50311>
- van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., et al. (2010). Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chemistry and Physics*, 10, 11,707–11,735. <https://doi.org/10.5194/acp-10-11707-2010>
- Virtanen, T. H., Kolmonen, P., Sogacheva, L., Rodríguez, E., Saponaro, G., & de Leeuw, G. (2018). Collocation mismatch uncertainties in satellite aerosol retrieval validation. *Atmospheric Measurement Techniques*, 11, 925–938. <https://doi.org/10.5194/amt-11-925-2018>

- Volpe, G., Santoleri, R., Vellucci, V., Ribera d'Alcalà, M., Marullo, S., & D'Ortenzio, F. (2007). The colour of the mediterranean sea: Global versus regional bio-optical algorithms evaluation and implication for satellite chlorophyll estimates. *Remote Sensing of Environment*, 104(4), 625–638. <https://doi.org/10.1016/j.rse.2006.10.017>
- Wagner, F., & Silva, A. M. (2008). Some considerations about Ångström exponent distributions. *Atmospheric Chemistry and Physics*, 8, 481–489. <https://doi.org/10.5194/acp-8-481-2008>
- Weatherhead, E. C., Harder, J., Araujo-Pradere, E. A., Bodeker, G., English, J. M., Flynn, L. E., et al. (2017). How long do satellites need to overlap? Evaluation of climate data stability from overlapping satellite records. *Atmospheric Chemistry and Physics*, 17, 15069–15093. <https://doi.org/10.5194/acp-17-15069-2017>
- Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X.-L., Choi, D., Cheang, W.-K., et al. (1998). Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *Journal of Geophysical Research*, 103(D14), 17,149–17,161. <https://doi.org/10.1029/98JD00995>
- Winker, D. M., Vaughan, M. A., Omar, A. H., Hu, Y., Powell, K. A., Liu, Z., et al. (2009). Overview of the CALIPSO mission and CALIOP data processing algorithms. *Journal of Atmospheric and Oceanic Technology*, 26, 2310–2323. <https://doi.org/10.1175/2009JTECHA1281.1>
- Witek, M. L., Garay, M. J., Diner, D. J., Bull, M. A., & Seidel, F. C. (2018). New approach to the retrieval of AOD and its uncertainty from MISR observations over dark water. *Atmospheric Measurement Techniques*, 11, 429–439. <https://doi.org/10.5194/amt-11-429-2018>
- Young, S. A., & Vaughan, M. A. (2009). The retrieval of profiles of particulate extinction from Cloud Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) data: Algorithm description. *Journal of Atmospheric and Oceanic Technology*, 26, 1105–1119. <https://doi.org/10.1175/2008JTECHA1221.1>
- Young, S. A., Vaughan, M. A., Garnier, A., Tackett, J. L., Lambeth, J. B., & Powell, K. A. (2018). Extinction and optical depth retrievals for CALIPSO's version 4 data release. *Atmospheric Measurement Techniques Discussions*, 11, 5701–5727. <https://doi.org/10.5194/amt-2018-182>
- Zhang, J., & Reid, J. S. (2006). MODIS aerosol product analysis for data assimilation: Assessment of over-ocean level 2 aerosol optical thickness retrievals. *Journal of Geophysical Research*, 111, D22207. <https://doi.org/10.1029/2005JD006898>
- Zhao, T. X.-P., Chan, P. K., & Heidinger, A. K. (2013). A global survey of the effect of cloud contamination on the aerosol optical thickness and its long-term trend derived from operational AVHRR satellite observations. *Journal of Geophysical Research: Atmospheres*, 118, 2849–2857. <https://doi.org/10.1002/jgrd.50278>