**Atmospheric
Chemistry
and Physics**

Open Access

Coherent uncertainty analysis of aerosol measurements from multiple satellite sensors

M. Petrenko[1,2] **and C. Ichoku**[2]

[1]Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA
[2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Correspondence to: M. Petrenko (maksym.petrenko@nasa.gov)

Abstract. Aerosol retrievals from multiple spaceborne sensors, including MODIS (on Terra and Aqua), MISR, OMI, POLDER, CALIOP, and SeaWiFS – altogether, a total of 11 different aerosol products – were comparatively analyzed using data collocated with ground-based aerosol observations from the Aerosol Robotic Network (AERONET) stations within the Multi-sensor Aerosol Products Sampling System (MAPSS, http://giovanni.gsfc.nasa.gov/mapss/ and http://giovanni.gsfc.nasa.gov/aerostat/). The analysis was performed by comparing quality-screened satellite aerosol optical depth or thickness (AOD or AOT) retrievals during 2006–2010 to available collocated AERONET measurements globally, regionally, and seasonally, and deriving a number of statistical measures of accuracy. We used a robust statistical approach to detect and remove possible outliers in the collocated data that can bias the results of the analysis. Overall, the proportion of outliers in each of the quality-screened AOD products was within 7 %. Squared correlation coefficient (R^2) values of the satellite AOD retrievals relative to AERONET exceeded 0.8 for many of the analyzed products, while root mean square error (RMSE) values for most of the AOD products were within 0.15 over land and 0.07 over ocean. We have been able to generate global maps showing regions where the different products present advantages over the others, as well as the relative performance of each product over different land cover types. It was observed that while MODIS, MISR, and SeaWiFS provide accurate retrievals over most of the land cover types, multi-angle capabilities make MISR the only sensor to retrieve reliable AOD over barren and snow/ice surfaces. Likewise, active sensing enables CALIOP to retrieve aerosol properties over bright-surface closed shrublands more accurately than the other sensors, while POLDER, which is the only one of the sensors capable of measuring polarized aerosols, outperforms other sensors in certain smoke-dominated regions, including broadleaf evergreens in Brazil and South-East Asia.

1 Introduction

Remote sensing of aerosols from space has been a subject of extensive research, with multiple sensors retrieving global aerosol properties on a daily or weekly basis. During the past decade, the retrievals of atmospheric aerosol parameters have been available from a multitude of spaceborne sensors (Lee et al., 2009; Yu et al., 2006). The diverse algorithms used for these retrievals operate on different types of the remotely sensed signals and rely on different assumptions about the underlying physical phenomena. Significant effort has been made by the various aerosol algorithm teams to refine progressively these assumptions, from algorithm version to version, in order to derive and provide the most accurate products possible. However, despite these efforts, measurements of identical aerosol parameters from different sensors, including the most common observable and widely used aerosol optical depth or thickness (AOD or AOT or τ_a) parameter, often disagree with each other due to a variety of reasons including differences in the underlying surface properties at different locations, intrinsic sensor observation characteristics and retrieval approaches (Li et al., 2009). Therefore, it has become necessary to analyze consistently the available aerosol products wherever possible in order to establish the geographical locations where and un-

der what circumstances each of these products provides the greatest accuracy.

The unique attributes of a particular sensor may be advantageous for aerosol retrievals, depending on the parameter(s) being retrieved, especially under favorable atmospheric conditions. However, aerosol retrieval accuracy can also be affected by numerous other factors, including the retrieval algorithm's assumptions and parameterizations, the instrument characteristics (intrinsic design, calibration, and time-dependent degradation), the measurement configurations (solar and view geometry), the atmospheric conditions (cloudiness, aerosol mixing, layer height, and humidity), the surface background (vegetated, bare, snow-covered, inundated, or simply just dark or bright land surface or ocean), and others (Kokhanovsky et al., 2007).

Since the accuracy of aerosol retrieval from a sensor may be affected positively or negatively by these factors and conditions in different ways and to varying degrees, a synergetic use of similar aerosol parameters across the sensors is non-trivial, and the data synergy research is instead focused on combining orthogonal (i.e., non-conflicting) aerosol measurements. For example, the aerosol layer height information from the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) has been used to enhance aerosol retrievals from other sensors (Oo and Holz, 2011; Torres et al., 2012; Zhang et al., 2011), while the geometry information from the Advanced Along Track Scanning Radiometer (AATSR) was used to initialize the Moderate Resolution Imaging Spectroradiometer (MODIS) bidirectional reflection distribution function (BRDF) in order to derive AATSR AOD (Guo et al., 2009).

To characterize better the differences and uncertainties that exist between the aerosol retrievals from different sensors, several studies compared a limited number of sensors. For example, AOD retrievals from MODIS were separately compared to retrievals from the MISR Multi-angle Imaging Spectroradiometer (Kahn et al., 2007, 2011; Mishchenko et al., 2010; Zhang and Reid, 2010), the POLDER – POLarization and Directionality of the Earth's Reflectances sensor (Gérard et al., 2005), and CALIOP(Kittaka et al., 2011; Redemann et al., 2012). A larger set of sensors was intercompared using a synthetic benchmark (Kokhanovsky et al., 2010), and also based on detailed analysis of data from limited geographical regions (Cheng et al., 2012; Yu et al., 2012). In addition, a set of 9 aerosol products was evaluated over ocean and coastal AERONET (Aerosol Robotic Network) sites during the period of 1997–2000, highlighting regions of high retrieval agreement and disagreement (Myhre et al., 2005). However, all the satellite data used in that study had already undergone post-retrieval spatiotemporal aggregation at 1×1 degree grid resolution on a monthly mean basis (so-called Level 3 products) before they were used in the comparisons.

Finally, a recent study compared AERONET retrievals with a set of 5 spaceborne aerosol products archived at the ICARE Data and Services Center, including POLDER,

MODIS-Aqua (Dark Target retrievals), MERIS, SEVIRI, and CALIOP (Bréon et al., 2011). Although that study was based on a similar collocation framework as that used in the current study, our study focuses on a different set of sensors that provides a more extensive set of over-land spaceborne aerosol products. Furthermore, the presented study is based on the analysis of the spatiotemporally averaged and outlier-screened data, whereas that of Bréon et al. (2011) is predominantly based on the analysis of individually collocated spaceborne and ground-based data points that are the closest in space and time that would correspond to the central values in our collocated data subsets (we report a similar analysis in the Supplement to this paper).

In this work, 11 retrieval-scale (Level 2) aerosol products from multiple spaceborne sensors are intercompared during the recent "golden" period of 2006–2010 (see Fig. 1), when as many as seven major sensors were in operation and measuring aerosols concurrently. Specifically, we focus on aerosol products retrieved over land and ocean from MODIS on Terra and Aqua, MISR on Terra, the Ozone Monitoring Instrument (OMI) on Aura, POLDER on PARASOL, CALIOP on CALIPSO, and the Sea-viewing Wide Field of view Sensor (SeaWiFS) aboard the SeaStar spacecraft. At the time of this study (January 2013), all of the studied sensors were still active, with the exception of SeaWiFS, whose operation ended in December 2010. The analysis is based on the collocation of the satellite data products using the Multi-sensor Aerosol Products Sampling System (MAPSS) framework (Petrenko et al., 2012), which samples these satellite products relatively uniformly over the global AERosol Robotic NETwork (AERONET) of sun photometers and other important ground-based stations both over land and ocean.

The details of the MAPSS sampling approach are explained in Sect. 2, while the relevant characteristics of the aerosol data products from the different sensors and the corresponding data quality screening techniques are described in Sect. 3 and Sect. 4. Section 5 describes a novel statistical approach for detecting and removing possible data outliers that can exist in the collocated data and, as a result, bias the statistical analysis of these data. Section 6 presents the detailed analysis of the compared aerosol products, while Sect. 7 examines the accuracy of these products based on land cover type. Conclusions are presented in Sect. 8.

2 Sampling method

The different aerosol-measuring sensors have different spatial resolutions, of which some have square-shaped footprints while others have rectangular pixel shapes. The nominal ground pixel sizes of the analyzed aerosol products at nadir are summarized in Table 1, and these sizes become progressively larger away from nadir. To ensure a uniform and fair sampling of the aerosol products for cross-evaluation with

Table 1. Ground-based and spaceborne atmospheric aerosol products analyzed in the study. In the product designation titles, "O" at the end of the title of a product signifies ocean retrievals, "L" – land retrievals, "DT" – land retrievals using the MODIS Dark Target algorithm, and "DB" – land retrievals using the MODIS Deep Blue algorithm. The AERONET AOD retrievals were interpolated to the studied wavelengths of the spaceborne sensors. The indicated local equatorial crossing times (LT) are based on the original orbital designs, and can change during the lifetimes of the satellites. SeaWiFS mission has ended in December 2010. * While 388 nm was the main observational wavelength used in the study for OMI, 500 nm was used where the collocated AERONET data were not available in the UV range.

| Sensor	Platform	Product	Designation in the study	Study Wave-length	Spatial Resolution (km × km)	Equator crossing time	Data available from
AERONET	N/A	AOT	AERONET	Varies	N/A	N/A	Varies
MODIS	Terra	MOD04	TMODIS DT TMODIS DB TMODIS O	550 nm	10 × 10	10:30 a.m. 1:30 p.m.	Mar 2000 Jul 2002
	Aqua	MYD04	AMODIS DT AMODIS DB AMODIS O				
MISR	Terra	MIL2ASAE	MISR	558 nm	17.6 × 17.6	10:30 a.m.	Feb 2000
OMI	Aura	OMAERUV	OMI	388 nm *500 nm	13.7 × 23.7	1:38 p.m.	Oct 2004
POLDER	PARASOL	P3L2TLGC P3L2TOGC	POLDER3 L POLDER3 O	865 nm 670 nm	19 × 19	1:30 p.m.	Mar 2005
CALIOP	CALIPSO	05kmALay	CALIOP	532 nm	5 × 0.1	1:32 p.m.	Jun 2006
SeaWiFS	SeaStar	SWDB	SeaWiFS L SeaWiFS O	550 nm	13.5 × 13.5	12:00 p.m.	Sep 1997
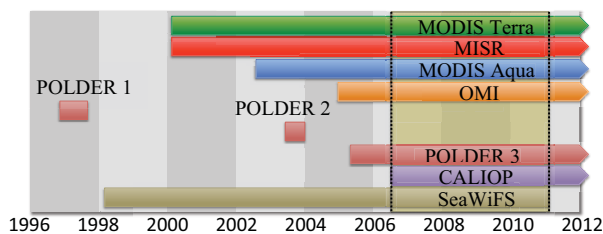
Fig. 1. Periods of operation of major past and current aerosol-measuring satellite sensors. The pair of dotted vertical lines marks the "golden" period (between the start of CALIOP in July 2006 and the end of SeaWiFS in December 2010) when as many as seven of these sensors were measuring aerosols concurrently. The golden period was used as the base for the studies reported in the rest of this paper.

AERONET and for comparison with one another, we used the framework of Multi-sensor Aerosol Products Sampling System (MAPSS) that was originally developed by Ichoku et al. (2002) for validation and analysis of MODIS aerosol products (Chu et al., 2002; Ichoku et al., 2003, 2005; Levy et al., 2010; Remer, 2002) and later expanded to support aerosol products retrieved by other spaceborne sensors (Petrenko et al., 2012). MAPSS subsets the aerosol products by extracting pixels covering approximately the same area on the ground centered over AERONET sun photometer measurement sites and over certain other point locations that are not addressed in this study.

Assuming an imaginary circle of 55 km diameter whose center coincides with each AERONET station, all space-borne aerosol product pixels falling within the circle are extracted. An aerosol pixel is regarded as being within the circle if the coordinates of the pixel center fall within 27.5 km from the coordinates of the circle center, where the distance between the coordinates of the two points is determined using the Haversine formula (Sinnott, 1984). Based on the nominal spatial resolution of the sensors in Table 1, the approximate maximum number of pixels within the 55 km diameter sample space at nadir for the different sensors is as follows: MODIS – 25, MISR – 9, OMI – 8, POLDER – 9, CALIOP – 11, and SeaWiFS – 16. The actual number of pixels within the sampling circle decreases for the aerosol retrievals away from the nadir of the satellite scene, and can be further reduced in the presence of clouds or other factors preventing retrieval of aerosol parameters. Based on the extracted sample, statistics of each aerosol parameter retrieved within the sampling areas are calculated and include mean, median, standard deviation, as well as the value of the central point in the sample, i.e., the pixel in the spaceborne subset that is the closest (i.e., whose center has the smallest distance) to the ground station, or the individual data point in the ground-based subset that is measured the closest in time to the overpass of the satellite.

In this paper, results are reported based on the analysis of the mean values. Although not reported in this paper because of the space considerations, a similar analysis was performed based on the central values and is reported in the

Supplement to this paper. It is appropriate to use the mean values in this paper, so as to maintain the uniform sampling criterion across the different sensors and their respective retrieval pixel sizes to facilitate a fair intercomparison. On the other hand, an analysis based on the central pixel values such as that reported in the Supplement can provide further details on the effect of difference in sampling aerosol products from individual sensors, as well as more accurately characterize the performance of the sensors in the presence of a strong point source of pollutant particles. Additionally, it should be noted that since the mean value of a sample can be computed even if its central value is missing, the reported analysis of the central values is based on a somewhat reduced volume of the collocated data points when compared to the reported analysis based on the mean values.

To collocate AERONET data in time and space with the satellite data, AERONET measurements acquired within ± 30 min of each satellite sensor overpass are also extracted and the corresponding statistics are derived. Additionally, for the convenience of aerosol data intercomparison and validation, AERONET AODs are interpolated to the wavelengths of spaceborne sensors in Table 1 based on the established wavelength dependence of AOD (Eck et al., 1999). It is pertinent to note that this interpolation process might introduce an additional source of uncertainty when intercomparing the aerosol products. Also, because of the wavelength dependence of AOD, the difference in the compared wavelengths of the spaceborne products should be considered when intercomparing the relative performance of the products. Furthermore, although many AERONET stations provide observations in the range of 340–1020 nm, certain stations report AOD in the range of 440–1200 nm. For such stations that have no measurements in the UV region, we have evaluated OMI AOD at 500 nm instead of AOD at 388 nm, in order to avoid additional extrapolation biases.

Each AERONET station has a different period of operation, and the quantity of available AOD data points is not uniform across all stations; while many stations are still active, certain stations were active in the past and only for a short period of time. The overall availability of the collocated data during the analysis period of 2006-06-07 to 2010-12-11 is shown in Fig. 2, where for the purposes of this study the stations are classified as land-only, ocean-only, or land-and-ocean. This classification is based on analyzing collocated data of separate aerosol retrievals over land and ocean from the MODIS, SeaWiFS, and POLDER sensors and identifying stations that have AOD data points from the land datasets, ocean datasets, or both; note that the MISR, OMI, and CALIOP sensors provide only joint land-and-ocean datasets.

3   Aerosol products

The key properties of the 11 analyzed aerosol products are summarized in Table 1, while the original science dataset (SDS) names of the spaceborne aerosol products are outlined in the first column of Table 2, except for the POLDER products that do not have an established SDS product naming convention. The sampled satellite data products are derived directly from the retrieval level aerosol products (Level 2) that represent the highest available spatial resolution for each product/sensor combination and are free of aggregation artifacts that can be present in data at Level 3 (Hyer et al., 2011; Levy et al., 2009; Zhang and Reid, 2010).

Of the 11 sampled products, 3 are combined land-and-ocean products, 6 land-only products, and 4 ocean-only products. Furthermore, 6 aerosol products are retrieved from the twin MODIS-Terra and MODIS-Aqua sensors using the same set of 3 algorithms: the ocean algorithm is used for the retrievals over oceans and other large bodies of water; the land Dark Target (DT) algorithm is used over vegetated regions and other dark surfaces (Remer et al., 2005); and the land Deep Blue (DB) algorithm is used for deserts and barren lands (Hsu et al., 2004). Although the results between the two MODIS sensors are expected to be very close, they might still differ due to the different times of scene observation during the day and other factors summarized in Ichoku et al. (2005) and Remer et al. (2008).

The remainder of this section provides a brief description of the analyzed products and highlights some of the unique aerosol properties reported in these products. A more detailed overview can be found in the theoretical and validation works of the respective science teams of the products as cited below, while a general comparative overview of multiple products and retrieval algorithms are in Kokhanovsky et al. (2007), Lee et al. (2009), Li et al. (2009), and Yu et al. (2006).

AERONET (http://aeronet.gsfc.nasa.gov) sun photometers measure aerosol properties using ground-based observations of solar direct and diffuse irradiances, where the nominal accuracy of AOD measurements is within the range of 0.01–0.02. In this work, the AERONET product used is the aerosol optical depth or thickness (AOD or AOT), which is retrieved from the AERONET direct measurements of solar irradiance. Since AERONET measurements are made from the ground looking up, they present a distinct advantage over spaceborne retrievals in that they are not affected by uncertainties associated with the effects of surface properties as much as satellite measurements are (Dubovik et al. 2002; Holben et al. 1998, 2001). Furthermore, the Level 2.0 AERONET data used in this work are carefully calibrated, cloud screened, and quality assured (Smirnov et al., 2000) and therefore are especially suitable for use as the reference standard against which the satellite aerosol remote-sensing data are evaluated.

Table 2. Studied aerosol datasets, the matching data quality (QA) datasets, and the corresponding QA data screening criteria. Where provided, numbers in parentheses in the middle column indicate the base-1 layer index, base-0 bit number, and number of bits extracted from this QA dataset. For MODIS, MISR, OMI, and SeaWiFS, the QA values are integer numbers between 0 and 3, whereas for MODIS and SeaWiFS larger numbers indicate a better retrieval quality, and for OMI and MISR the opposite is true. For POLDER, QA is a real number between 0 (worst) and 1 (best). For CALIOP, a column is accepted only if all layers found in this column meet all listed QA conditions. The listed extinction QC values indicate retrievals that are unconstrained, constrained, have a reduced lidar ratio, or detected an opaque aerosol layer. CAD score and layer type and subtype flags indicate retrievals that classified a layer with a high confidence as containing aerosol and were able to determine the aerosol type. IAB condition is set to prevent the retrieval anomaly of overcorrecting the attenuation of overlying layers (Kittaka et al., 2011).

AOD dataset	QA dataset (layer/starting bit/number of bits)	QA condition
MODIS		
Corrected_Optical_Depth_Land	Quality_Assurance_Land (1/1/3)	QA=3
Deep_Blue_Aerosol_Optical_Depth_550_Land	Quality_Assurance_Land (5/1/2)	QA=3
Effective_Optical_Depth_Average_Ocean	Quality_Assurance_Ocean (1/5/3)	QA=[1,2,3]
MISR		
RegBestEstimateSpectralOptDepth	RegBestEstimateQA	QA=[0,1]
OMI		
FinalAerosolOpticalDepth	FinalAlgorithmFlags	QA=0
SeaWiFS		
aerosol_optical_thickness_550_land	aerosol_optical_thickness _confidence_flag_land	QA=3
aerosol_optical_thickness_550_ocean	aerosol_optical_thickness _confidence_flag_ocean	QA=[2,3]
CALIOP		
Column_Optical_Depth_Aerosols_532	ExtinctionQC_[532 and 1064]	QA=[0,1,2,16,18]
	CAD_Score	$-100 \leq CAD < -20$
	Feature_Classification_Flags (1/0/3)	Layer type=3
	Feature_Classification_Flags (1/9/3)	Layer subtype>0
	Integrated_Attenuated _Backscatter_[532 and 1064]	IAB$\leq$ 0.01
POLDER		
Aerosol optical thickness at 865 nm corresponding to the polarized particles	Quality index for the land inversion	QA$\geq$0.5
Aerosol optical thickness at 670 nm	Quality index for the ocean inversion	QA$\geq$0.2

The MODIS (http://modis.gsfc.nasa.gov) aerosol product (MOD04 and MYD04) comprises the column aerosol optical thickness and other physical properties of aerosols retrieved globally over land and ocean (Chu et al., 2002; Hsu et al., 2004; Ichoku et al., 2005; Levy et al., 2010; Remer, 2002; Remer et al., 2005). MODIS has a swath of 2300 km.

The MISR (http://www-misr.jpl.nasa.gov) aerosol product (MIL2ASAE) features aerosol retrievals based on observations from 9 independent camera angles. Though limited to a swath of 563 km, its multiple viewing angles allow MISR to measure certain aerosol properties that are not available from the other instruments (e.g., aerosol particle size). Furthermore, MISR multiple cameras enable retrievals under conditions that are unfavorable to single-view (e.g., nadir) instruments, such as over bright surfaces or sun glint, where the other instruments are unable to make reliable retrievals in the visible wavelengths (Kahn, 2005; Kahn et al., 2010; Martonchik et al., 2009).

The OMI (http://www.knmi.nl/omi/research/instrument/index.php) aerosol product (OMAERUV) measures the near-UV (near ultraviolet) aerosol absorption and extinction optical depth, as well as single scattering albedo, among other aerosol properties (Torres et al., 1998, 2007). In addition to offering a generous swath of 2800 km, OMI is capable of retrieving absorption optical depth in partially cloudy conditions that usually pose a challenge to other aerosol instruments.

The POLDER onboard PARASOL (http://www.icare.univ-lille1.fr/parasol) aerosol land product (P3L2TLGC) and aerosol ocean product (P3L2TOGC) are derived from measuring spectral, directional, and polarized properties of reflected solar radiation. Also with a swath of 2800 km, one of the main features of the POLDER instrument is its utilization of polarization properties of the measured radiation for retrieving anthropogenic aerosol optical depth (Bréon et al., 2002; Buriez et al., 2002; Deuzé et al., 1999, 2001; Herman et al., 1997). It is important to note that the POLDER operational algorithm retrieves AOD at 2 wavelengths (670 and 865 nm) over ocean and only at 1 wavelength (865 nm) over land. Furthermore over land, the POLDER algorithm retrieves only AOD that corresponds to polarized particles, i.e., mainly fine-mode particles originating from anthropogenic activities. However, since there are no specific recommendations on the regions suitable for using the polarized AOD retrievals either as a proxy for total AOD or as pure fine-fraction AOD, it is beneficial to explore this dataset in relation to other available AOD products in order to identify geographical regions where the polarized AOD from POLDER can potentially be treated as total AOD.
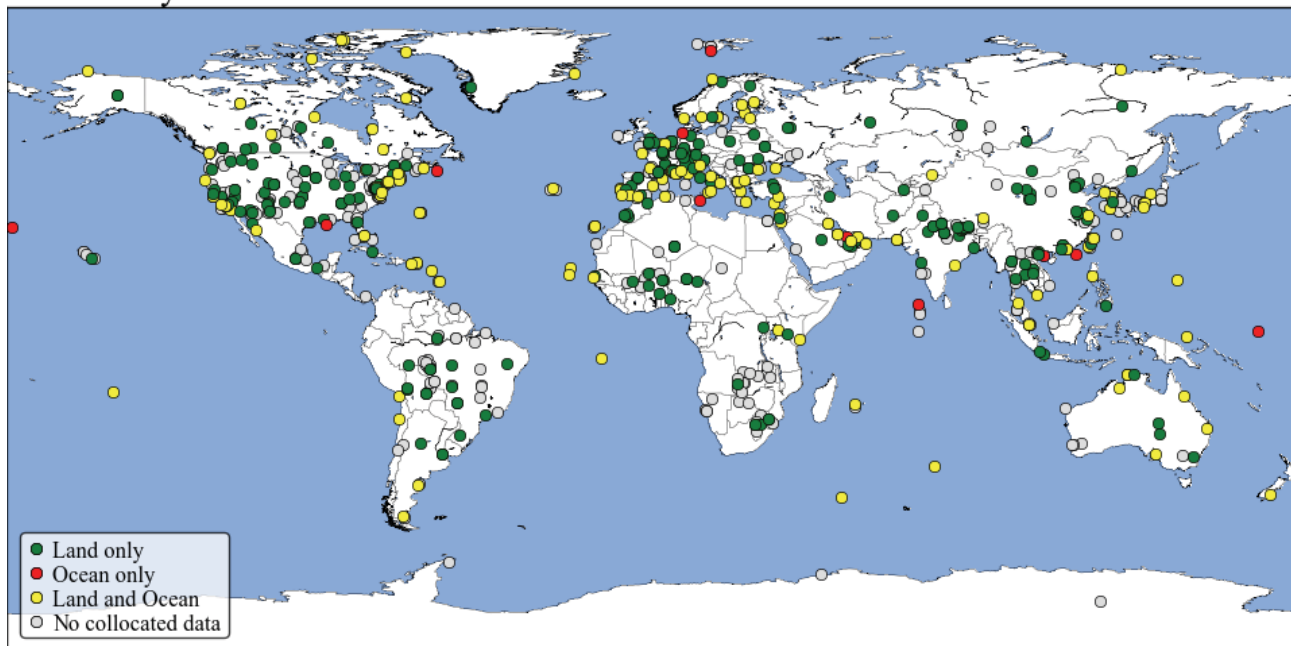
Fig. 2. Distribution of AERONET stations used in the study. Green, red, and yellow colors indicate stations that can be classified as land-only (233 sites), ocean-only (11 sites), or both land-and-ocean (149 sites), respectively. The classification was established based on data availability in separate over-land and over-ocean datasets in MODIS, SeaWiFS, and POLDER aerosol products. Gray color indicates stations that do not have any collocated data for the studied period of time.

The SeaWiFS (http://disc.sci.gsfc.nasa.gov/dust/) aerosol product (SWDB) uses the Deep Blue algorithm to derive aerosol optical thickness and Ångström exponent. Also based on an orbital ground-coverage swath of 2800 km, the key features of this product are the retrievals of aerosol properties over both bright desert and vegetated surfaces, avoidance of sun glint that improves aerosol retrievals over ocean, and a highly precise calibration of the SeaWiFS sensor (Hsu et al., 2004, 2012).

The CALIOP (http://www-calipso.larc.nasa.gov) aerosol product (05kmALay) represents atmospheric curtain slices portraying the vertical distribution of aerosols and clouds in the atmosphere, including the density and certain properties of individual aerosol layers (Omar et al., 2009; Winker et al., 2007). Since CALIOP is an active lidar sensor, it can provide both daytime and nighttime retrievals within a narrow swath of about 70 m. Although the lack of the daytime background solar illumination makes nighttime CALIOP retrievals more accurate, they are not used in this study because they cannot be intercompared with the AERONET retrievals, which are available only during the daytime.

Since each of the foregoing datasets has a few versions because of the periodic revisions and updates of their retrieval algorithms over time, the data versions that were current at the time of writing this paper (January 2013) were sampled, although the study has been designed in a highly flexible

way to enable rapid re-analysis as the new versions become available. The respective data versions used in this paper are AERONET AOD (Version 2), Terra and Aqua MODIS (Collection 051), MISR (Version 002), OMI (Version 003), POLDER (Versions L and K), SeaWiFS (Version 004), and CALIOP (Version 3-01). Therefore, all of the illustrations and analyses shown in this paper are based on these data versions for the respective aerosol sensors.

4   Data quality screening

While the AERONET Level-2 data are manually inspected to be free of retrieval defects and anomalies (Smirnov et al., 2000), such an approach is not feasible for the voluminous spaceborne data. Instead, all Level-2 aerosol products analyzed in this paper assign to AOD pixels one or more quality assurance (QA) flags that indicate a degree of "confidence" of the retrieval algorithms in their results. For MODIS and SeaWiFS, aerosol QA flags are integer numbers ranging from 0 to 3, with 3 representing the highest quality. For MISR and OMI data, the reverse is the case (i.e., 0 is the highest quality). Finally, for POLDER and CALIOP, QA data are a combination of one or more flags, most of which are real numbers ranging between 0 and 1, where 1 indicates the highest quality. By means of these QA flags, certain aerosol retrievals

are identified as "bad quality" and are considered to be not trustworthy enough for certain analyses. Therefore, users of these aerosol products have been advised to choose data corresponding to a range of QA values that is most appropriate for their specific needs.

To establish similar yet valid QA thresholds for the analyzed products, we consulted science teams of the analyzed products as well as data product validation results reported by these teams and other research groups. Based on this inquiry, we chose the acceptable QA values as described in Table 2. For the majority of the products, the thresholds are set based on selecting a limited subset of the possible QA values. An important exception is the POLDER aerosol products, where the QA flags are expressed as real numbers between 0 ("bad") and 1 ("excellent"). Since there are no formal recommendations on the acceptable range of these flag values, we have adopted thresholds suggested for the "quality of inversion" flag in Bréon et al. (2011), specifically 0.5 for land retrievals and 0.2 for ocean retrievals. It is also important to note that since the primary designation of this flag is to indicate the success of the retrieval algorithm, this flag does not always reflect the actual quality of the retrieved aerosol parameters, especially under certain less than favorable conditions (Fig. 3).

The original MAPSS framework was designed to facilitate data analysis experiments based on different values of QA flags. For this, MAPSS extracts QA flags over the sampling area and computes the statistical mode for integer QA flags and mean for real QA flags. These statistical modes of the integer QA flags and means of the real QA flags provide a single number for the quality assessment of each sample set, and can be used to screen the corresponding subset statistics while providing a convenient alternative compared to screening individual pixels (e.g., see Levy et al., 2010; Remer et al., 2008). However, it was observed that this approach has an unequal impact on the statistical properties of the different aerosol products (Petrenko et al., 2012).

As an example, consider Fig. 4, where the global collocated subset mean AOD values from OMI and Terra MODIS Deep Blue (TMODIS DB) are compared to the corresponding subset mean AOD values from AERONET. It can be observed that while filtering the mean TMODIS DB AOD values by the mode of QA flags improves the R^2 and RMSE statistics, when compared to computing the mean values based on individually screened TMODIS DB AOD pixels, this filtering significantly changes the distribution of the collocated data. Specifically, compared to screening individual pixels, QA mode filtering removes 50 % more of the collocated data points and degrades the slope of the fitted regression line as a result of removing certain high-biased points. The opposite behavior can be observed in the collocated OMI AOD and AERONET AOD datasets, where screening by QA mode degrades R^2 of the collocated data when compared to screening individual pixels, although RMSE is still improved and the slope of the fitted regression line remains the same,

since both screening approaches produce approximately the same number of the OMI subset data points.

This observation indicates a certain inhomogeneity in the uncertainties that are present in the aerosol products, as in some cases high biases in individual pixels might overwhelm the statistics derived from the sample set. Therefore, to avoid such biases and ensure a fair comparison between the analyzed aerosol products, the rest of this study is based on the QA "pre-filtering" approach, where individual pixels in a spatial sample are screened by their QA values before computing the statistics of this sample. This approach also closely models a typical use of the spaceborne aerosol data, where data users screen each pixel individually and do not consider QA values of its neighboring pixels. The data quantity impact of the described QA screening approach can be observed in Table 3, which provides the sizes of the analyzed datasets before and after the screening. It is noticeable that, depending on the product, the impact is quite different, with the two MODIS ocean AOD datasets and the MISR AOD dataset retaining almost all of their available datasets, whereas the two MODIS DT datasets retained only one-fourth of the complete collocated datasets.

It is important however to keep in mind that the QA values reported by the retrieval algorithms are to a large degree subjective to these algorithms and do not always reflect the actual quality of the retrievals. For example, in the absence of a proper aerosol or surface model, an algorithm can in certain cases use a wrong model to retrieve aerosol properties and mistakenly assign this retrieval a "good" QA flag (e.g., Kahn et al., 2010; Levy et al., 2010). Furthermore, a retrieval algorithm used might not have enough skill or even the possibility to recognize correctly certain conditions that are unfavorable for aerosol retrieval, e.g., sub-pixel cloud contamination in OMI retrievals (Torres et al., 1998). In yet another situation, an aerosol scene can be observed in only a portion of the available observation modes of a sensor, e.g., in only a few of the available observation directions in POLDER (Herman et al., 1997), which can lead to more confident yet less reliable results. The opposite case can also be true where an algorithm correctly retrieves aerosol properties but is not confident about the retrieval. As an example, consider Fig. 5, which explores how QA screening degrades the statistics of OMI AOD and Aqua MODIS Deep Blue AOD datasets when compared to AERONET AOD over Djougou, Benin, as a result of assigning a "bad" QA flag to sufficiently "good" retrievals.

5   Possible data outliers

Under rare circumstances, aerosol retrievals from spaceborne observations can produce aerosol properties that do not reflect the actual physical state of aerosol in the atmosphere. Some of the reasons for such retrievals were discussed in the previous section and might include, but are not limited

Table 3. Statistics of the studied aerosol datasets based on all AERONET stations during the period of 2006-06-07 and 2010-12-11. "N_{tot}" indicates the total number of the collocated spaceborne AOD – AERONET AOD data points, while "N_{filt}" indicates the number of data points after filtering (screening) the spaceborne data by QA as described in Sect. 4 and Table 2. "N_{out}" is the total number of the possible data outliers determined as explained in Sect. 5. The last 8 columns present the statistics on the collocated data based on regression fits also plotted in Fig. 6. Please see the Supplement for a breakdown of the listed statistics based on nominal ranges of AOD loading.

Dataset	N_{filt}	$N_{filt} \div$ N_{tot} (%)	N_{out}	$N_{filt} \div$ N_{out} (%)	Complete data				Outliers removed			
					R^2	RMSE	Slope	Intercept	R^2	RMSE	Slope	Intercept
All seasons												
TMODIS DT	76 781	69.5	1126	2.1	0.79	0.11	0.96	0.01	0.81	0.09	0.93	0.01
TMODIS DB	15 319	28.9	175	4.0	0.63	0.23	0.87	0.06	0.68	0.20	0.95	0.03
TMODIS O	16 399	99.6	561	3.4	0.82	0.08	1.02	0.02	0.87	0.06	0.97	0.03
AMODIS DT	75 648	64.9	1372	2.8	0.78	0.11	0.98	0.01	0.80	0.09	0.95	0.01
AMODIS DB	43 996	30.8	676	5.0	0.62	0.23	0.88	0.06	0.68	0.19	0.91	0.04
AMODIS O	17 903	99.5	638	3.6	0.80	0.08	0.94	0.03	0.85	0.06	0.90	0.03
MISR	13 673	99.8	642	4.7	0.73	0.13	0.60	0.07	0.83	0.07	0.76	0.04
OMI	56 072	94.5	1545	2.9	0.39	0.24	0.68	0.16	0.44	0.21	0.71	0.14
CALIOP	2056	91.3	130	6.9	0.35	0.22	0.51	0.06	0.55	0.15	0.63	0.03
POLDER3 L	17 288	95.5	898	5.4	0.24	0.20	0.16	0.02	0.30	0.16	0.20	0.01
POLDER3 O	9568	94.7	402	4.4	0.45	0.15	0.30	0.03	0.50	0.12	0.37	0.02
SeaWiFS L	46 546	39.6	628	3.4	0.76	0.12	0.81	0.03	0.81	0.10	0.87	0.02
SeaWiFS O	16 551	79.4	575	4.4	0.72	0.10	0.98	0.02	0.79	0.07	0.93	0.02
Fall												
TMODIS DT	22 986	69.1	292	1.8	0.83	0.11	1.04	−0.02	0.85	0.08	0.97	−0.01
TMODIS DB	5229	27.2	49	3.5	0.48	0.19	0.90	0.03	0.50	0.15	0.81	0.04
TMODIS O	5198	99.5	183	3.5	0.79	0.07	0.90	0.04	0.84	0.05	0.93	0.03
AMODIS DT	20 676	62.4	366	2.8	0.82	0.11	1.05	−0.01	0.85	0.08	0.98	0.00
AMODIS DB	13 238	29.7	148	3.8	0.59	0.20	0.97	0.02	0.63	0.18	0.94	0.02
AMODIS O	5191	99.5	180	3.5	0.74	0.07	0.79	0.04	0.81	0.06	0.81	0.04
MISR	3988	99.8	185	4.6	0.68	0.14	0.54	0.07	0.84	0.07	0.71	0.05
OMI	15 140	92.7	364	2.6	0.34	0.21	0.67	0.14	0.36	0.20	0.69	0.13
CALIOP	582	91.2	40	7.5	0.35	0.21	0.61	0.05	0.38	0.17	0.56	0.05
POLDER3 L	5603	95.4	280	5.2	0.42	0.14	0.37	0.00	0.50	0.11	0.43	−0.01
POLDER3 O	2606	93.3	117	4.8	0.54	0.12	0.36	0.01	0.62	0.09	0.46	0.00
SeaWiFS L	13 357	43.8	189	3.2	0.79	0.10	0.83	0.02	0.81	0.08	0.90	0.01
SeaWiFS O	4583	79.0	200	5.5	0.70	0.09	0.93	0.02	0.77	0.06	0.88	0.02
Winter												
TMODIS DT	10 324	61.1	146	2.3	0.79	0.11	0.76	0.03	0.81	0.10	0.78	0.03
TMODIS DB	2458	33.6	35	4.2	0.65	0.23	0.90	0.08	0.68	0.20	0.90	0.07
TMODIS O	3322	99.8	131	4.0	0.72	0.09	0.98	0.03	0.85	0.05	0.91	0.03
AMODIS DT	9081	54.6	150	3.0	0.78	0.13	0.77	0.03	0.79	0.12	0.78	0.03
AMODIS DB	8327	36.6	121	4.0	0.62	0.21	0.84	0.06	0.65	0.18	0.86	0.05
AMODIS O	3341	99.4	151	4.5	0.74	0.08	0.84	0.03	0.80	0.05	0.84	0.03
MISR	2373	99.8	121	5.1	0.78	0.10	0.59	0.06	0.83	0.08	0.65	0.05
OMI	9321	92.6	169	2.0	0.42	0.23	0.60	0.15	0.42	0.21	0.60	0.15
CALIOP	344	91.0	21	6.7	0.42	0.25	0.36	0.09	0.46	0.21	0.43	0.08
POLDER3 L	2106	95.6	120	6.0	0.29	0.24	0.12	0.02	0.31	0.20	0.13	0.02
POLDER3 O	1544	93.9	62	4.3	0.60	0.11	0.39	0.01	0.60	0.10	0.43	0.01
SeaWiFS L	7327	42.2	110	3.6	0.75	0.13	0.81	0.04	0.80	0.11	0.85	0.03
SeaWiFS O	2409	78.3	76	4.0	0.68	0.09	0.83	0.03	0.75	0.07	0.84	0.03

Table 3. Continued.

Dataset	N_{filt}	$N_{filt} \div N_{tot}$ (%)	N_{out}	$N_{filt} \div N_{out}$ (%)	Complete data				Outliers removed			
					R^2	RMSE	Slope	Intercept	R^2	RMSE	Slope	Intercept
Spring												
TMODIS DT	15 094	70.6	207	1.9	0.80	0.11	0.92	0.03	0.81	0.10	0.94	0.02
TMODIS DB	2465	30.2	26	3.5	0.65	0.30	0.78	0.12	0.66	0.27	0.88	0.08
TMODIS O	2882	99.8	95	3.3	0.86	0.10	1.07	0.03	0.88	0.08	1.02	0.03
AMODIS DT	16 978	67.3	279	2.4	0.78	0.12	0.96	0.03	0.80	0.11	0.98	0.02
AMODIS DB	9633	32.0	88	2.9	0.63	0.27	0.83	0.09	0.66	0.25	0.89	0.07
AMODIS O	3632	99.8	142	3.9	0.81	0.09	0.99	0.03	0.86	0.07	0.94	0.04
MISR	2916	99.9	153	5.3	0.76	0.14	0.63	0.08	0.83	0.09	0.74	0.05
OMI	12 858	95.6	367	3.0	0.41	0.27	0.66	0.18	0.46	0.25	0.69	0.16
CALIOP	474	92.6	28	6.4	0.37	0.25	0.56	0.05	0.43	0.23	0.62	0.03
POLDER3 L	3766	96.3	193	5.3	0.20	0.28	0.10	0.02	0.23	0.21	0.13	0.02
POLDER3 O	2422	95.7	99	4.3	0.41	0.19	0.26	0.03	0.42	0.16	0.29	0.03
SeaWiFS L	10 512	41.7	135	3.1	0.79	0.15	0.80	0.05	0.82	0.12	0.85	0.04
SeaWiFS O	3749	79.4	118	4.0	0.73	0.12	1.04	0.01	0.78	0.08	0.97	0.02
Summer												
TMODIS DT	28 377	72.3	373	1.8	0.77	0.10	1.01	0.00	0.78	0.09	1.01	0.00
TMODIS DB	5167	27.7	41	2.9	0.66	0.24	0.91	0.04	0.69	0.21	0.98	0.02
TMODIS O	4997	99.6	169	3.4	0.84	0.09	1.06	0.02	0.88	0.07	1.02	0.02
AMODIS DT	28 913	68.6	487	2.5	0.76	0.11	1.05	0.01	0.77	0.09	1.03	0.01
AMODIS DB	12 798	27.1	125	3.6	0.62	0.24	0.89	0.07	0.66	0.22	0.98	0.05
AMODIS O	5739	99.4	205	3.6	0.84	0.08	1.00	0.02	0.89	0.06	0.98	0.02
MISR	4396	99.9	219	5.0	0.72	0.13	0.62	0.08	0.82	0.08	0.79	0.04
OMI	18 753	96.1	492	2.7	0.35	0.24	0.72	0.16	0.39	0.22	0.75	0.14
CALIOP	656	90.7	30	5.0	0.34	0.18	0.54	0.05	0.49	0.14	0.64	0.02
POLDER3 L	5813	94.9	337	6.1	0.24	0.18	0.15	0.01	0.24	0.15	0.15	0.01
POLDER3 O	2996	95.5	115	4.0	0.39	0.16	0.27	0.04	0.40	0.14	0.30	0.03
SeaWiFS L	15 350	33.2	185	3.6	0.69	0.12	0.76	0.03	0.74	0.10	0.85	0.01
SeaWiFS O	5810	80.1	189	4.1	0.74	0.10	1.01	0.01	0.81	0.07	1.02	0.01

to, such factors as the lack of a proper aerosol model, incorrect assumptions about boundary conditions, cloud contamination, and several other factors. In Fig. 6, the possible abnormal retrievals can be visually identified by observing points that have a minimal data density and lie abnormally far from the fitted regression lines. Even though an actual fraction of such data points in a complete collocated dataset can be relatively minor, the extreme deviations of such points from the overall trend might significantly bias and misrepresent the overall statistics of the data. Therefore, when computing the overall statistics and inter-comparing the aerosol products, such data points should be treated as possible outliers and analyzed separately from the rest of the data.

In order to identify and separate the possible data outliers, we analyzed AOD residuals, i.e., the difference between spaceborne AOD and AERONET AOD observations, using the modified Z-score test (Iglewicz and Hoaglin, 1993; National Institute of Standards and Technology, 2012). This test is designed for testing data for multiple outliers in approximately normal datasets and works by finding data points that

differ from the mean value by more than 5 median absolute deviations. Unlike the standard deviation used in the traditional Z-score test, the median absolute deviation in the modified Z-score test is calculated based on the median of the data and is less sensitive to extreme values. When applied to the collocated AOD data, this test removes those spaceborne retrievals that grossly overestimate or underestimate ground-based observations as compared to the median retrieval error at a specific AERONET location. This is especially useful since many spaceborne retrieval algorithms tend either to under-estimate certain high-AOD events because of a pre-set maximum AOD threshold, or to over-estimate AOD in the presence of clouds as well as under very low-AOD conditions. However, it should be noted that the test may not remove such possible outliers that have a relatively small error.

It is pertinent to note that even though AOD data are known to follow the lognormal distribution (O'Neill et al., 2000), the AOD residuals of the analyzed products follow an approximately normal distribution as shown in Fig. 7 and Fig. 8, with the exception of the POLDER products that
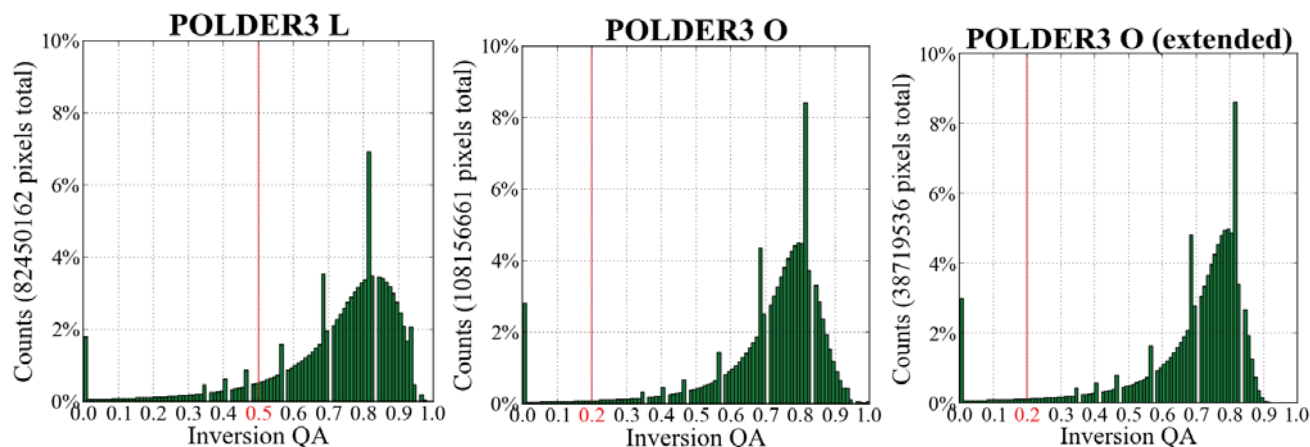
Fig. 3. The total quantities of POLDER3 land ("L") and ocean ("O") pixels based on different values of "quality of inversion" flag during the analyzed period of July 2006–December 2010. The "POLDER3 O (extended)" histogram is based on those pixels in ocean retrievals, where the retrieval algorithm considered the sensor viewing geometry conditions to be especially "favorable" and produced a set of additional aerosol parameters, such as spherical large-mode AOD, refractive index of fine mode, and others. The quality flag values are binned into 0.01 intervals, and the red lines indicate the 0.5 (land) and 0.2 (ocean) QA thresholds used in this study. Please note that even though certain retrievals can have a very high value of inversion QA (e.g., QA > 0.9 in ocean retrievals), if they were retrieved under less than favorable conditions, they may not necessarily be high-quality aerosol data, as there are almost no extended ocean retrievals with QA > 0.9.

mostly underestimate AOD, because their retrievals focus on anthropogenic fine-mode aerosols, and thus represent only the negative portion of the distribution. In the figures, it can be seen that the distributions have long tails, strongly indicating the presence of outliers. Furthermore, it can be observed that the slopes of the fitted lines are different from the slope of the 1 : 1 line. This indicates that the standard deviation of the analyzed residuals is different from 1, showing that these data do not follow the standard normal distribution, although this difference does not affect the test since the modified Z-score test normalizes residuals by the median absolute deviation of the data.

The overall effect of removing the possible outliers can be observed in the bottom-right sub-plots of Fig. 6, Fig. 7, and Fig. 8, as highlighted by the green frames, showing that 575 (4.4 %) outliers are removed from the SeaWiFS ocean AOD dataset. The total numbers of the removed outliers are provided in Table 3 and do not exceed 7 % of the total QA-screened data for any of the datasets when considering the all-season data. The global distribution of the possible data outliers is depicted in Fig. 9 and generally corresponds to the outlier locations reported by the science teams of the aerosol products, e.g., outliers around the coastal areas where the significant subpixel surface variations, shallow waters, sediments, and complex marine/inland aerosol mixtures complicate the retrievals, and also data outliers associated with uncertain retrievals by the MODIS and MISR algorithms in Amazon Basin and near the Sahara desert (Kahn et al., 2010; Levy et al., 2010), although a more detailed study is needed to determine the specific factors that lead to these outliers and their spatiotemporal distributions. Since the Z-score test

requires the use of a reference dataset (i.e., AERONET in our case), it cannot be directly applied to remove outliers in spaceborne data in an independent and systematic fashion. However, the results of this study could possibly be used to develop appropriate mitigation measures in the retrieval algorithms or to design specific data screening strategies for each of the products.

In the remainder of this paper, the reported results are based on the QA-screened data with the outliers removed.

6 Analysis

The overall data distribution for the analyzed spaceborne aerosol products is presented in Fig. 6, whereas the detailed linear regression fit statistics (Fox, 1997) for the products based on the treatment of the possible data outliers and the nominal delimiters of the four boreal seasons, namely, spring (March–May), summer (June–August), autumn (September–November), and winter (December to February), are listed in Table 3. The statistics are presented based on seasonal time frames rather than monthly or shorter time periods because there may not be sufficient coincident data for a scatterplot over such shorter time periods, due to the infrequency of satellite aerosol retrieval caused by cloud cover and other issues. Fortunately, many climatic events that are relevant to aerosol emission, transport, and distribution are often roughly aligned with these seasons.

The second column of Table 3 (N_{filt}) outlines the total volume of the collocated quality-filtered data available for each of the sensors depending on the boreal season. Although sensor swath width (Sect. 3) and data quality (Sect. 4) are
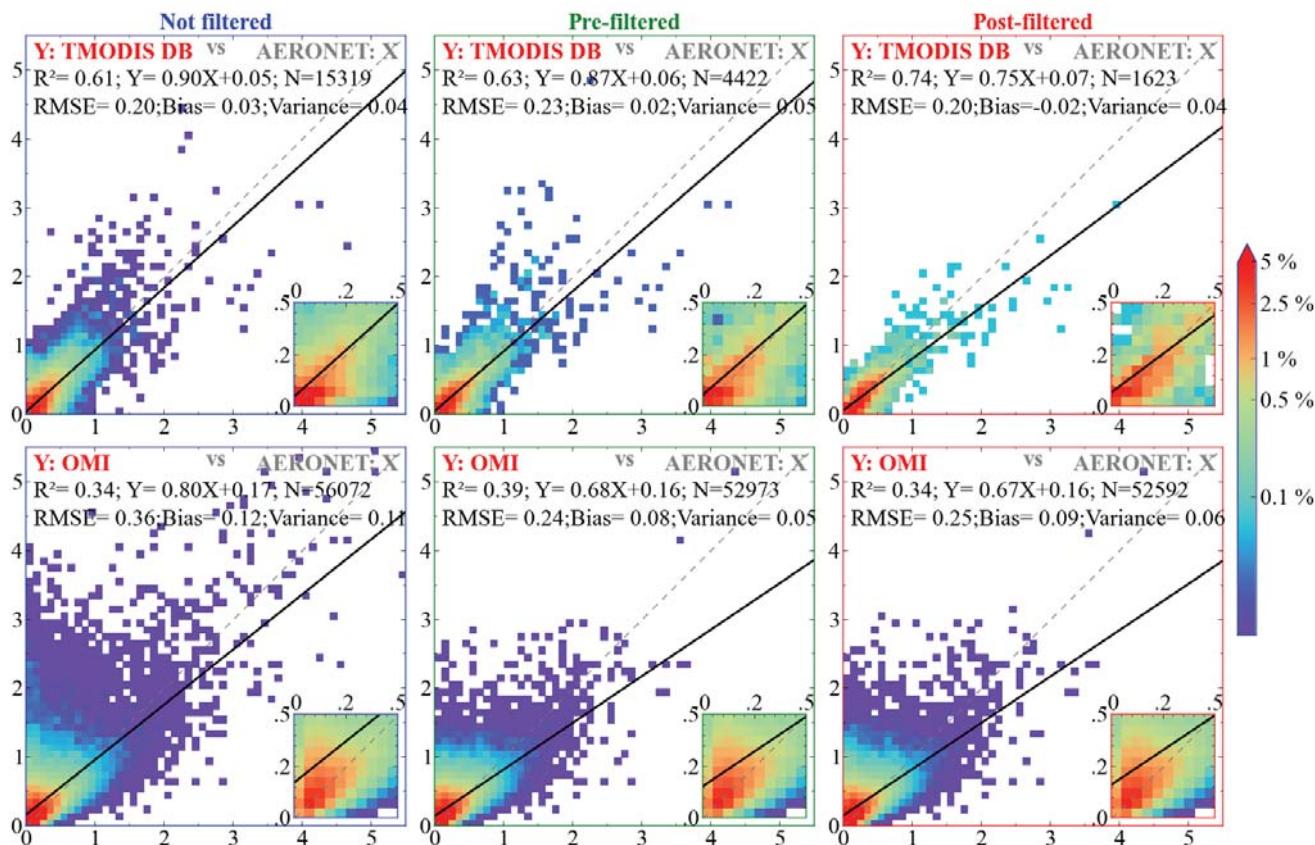
Fig. 4. Effects of two different data QA filtering schemes on the accuracy of the global collocated spaceborne AOD, as discussed in Sect. 4. AERONET AOD data are shown on the x axes, while AODs measured by spaceborne sensors are on the y axes. Density plots bin data into 0.1 AOD (0.05 AOD in magnified insets) intervals, where the color of each bin indicates the percentage of all data points that fall into this bin. Left column displays the original unfiltered data with all QA values. Middle column displays the data *pre-filtered* by QA, where individual pixels in each data sample were filtered based on their QA values *before* calculating the mean value of the sample. Right column shows the data *post-filtered* by QA, where the mean of each sample was calculated based on all pixels in the sample; *after* this, the whole sample was rejected if at least half of the pixels in the sample had QA values below the specified threshold. Note that OMI data have somewhat better properties when pre-filtered, while Terra MODIS–Deep Blue data are in a better agreement with AERONET AOD when post-filtered. The insets (0–0.5 AOD range magnified) are intended to enhance the visualization of the linear regression fits near the origin.

among the main factors that determine the available volume of the data (e.g., MODIS has approximately 4 times the swath width and 4 times the data volume of MISR), it can be seen that the seasonal changes in retrieval conditions also have a very considerable impact on the data. Thus, summer retrievals can have 2–4 times as much collocated data points as winter retrievals. The relative data volume differences between the studied data products should be carefully considered when interpreting the statistics discussed in the remainder of this paper.

In the presented statistics of Table 3, the slope value indicates by how much the satellite retrieval for the parameter under consideration is relatively underestimated or overestimated across different magnitudes, depending on whether the slope value is less than or greater than unity. The offset parameter indicates the extent to which the satellite retrieval

is biased. The squared linear correlation coefficient (R^2) indicates how consistent the parameter retrieval is across its magnitude range, that is how tightly the points are aligned close to the 1-to-1 line. Finally, the root mean square error (RMSE) indicates the accuracy of the retrievals measured as the average error in the spaceborne retrievals as compared to the ground-based AERONET retrievals.

In Table 3, it can be seen that all MODIS, MISR, and SeaWiFS aerosol products correlate with AERONET observations with $R^2 \geq 0.6$. Furthermore, MISR, SeaWiFS land, and MODIS Dark Target products have $R^2 \geq 0.7$, and MODIS ocean products have $R^2 \geq 0.8$. Also, once the possible outliers are removed, the SeaWiFS, MISR, and MODIS Dark Target products reach $R^2 \approx 0.8$, while the MODIS Deep Blue products have $R^2 \approx 0.7$. All the best-performing land products have RMSE values of about 0.15 (measured in the
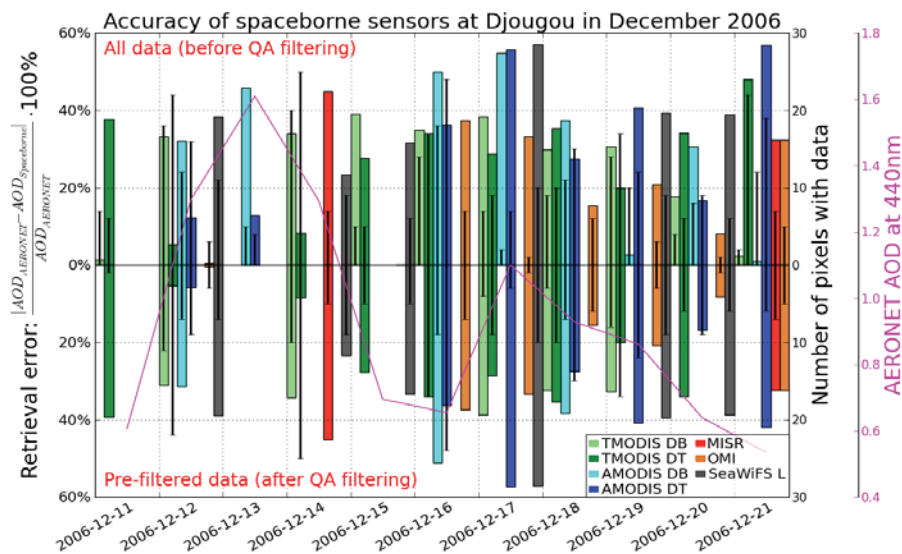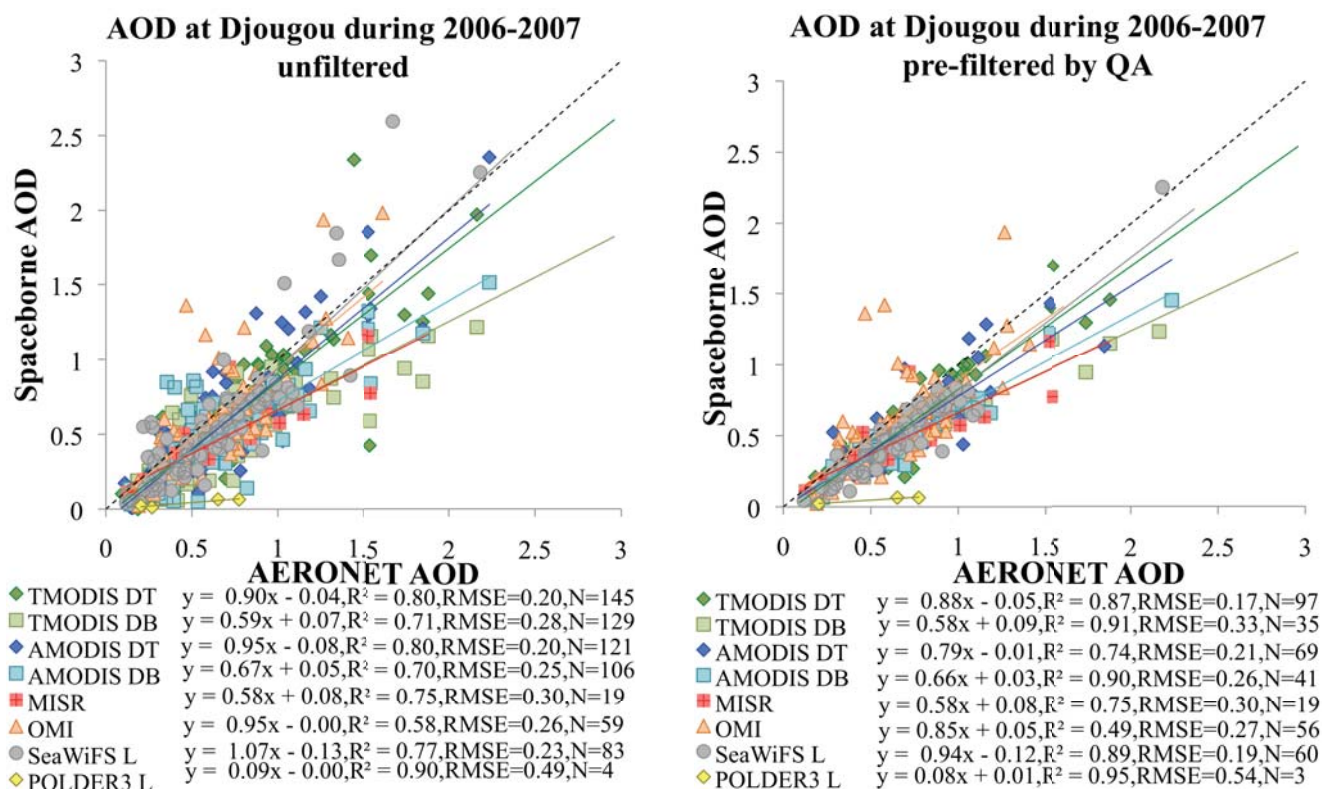
Fig. 5. Impact of QA screening on the statistical properties of AOD retrieved by the different sensors over Djougou, Benin. The top part of the figure shows scatterplots of 2 yr of data that are unfiltered (left) or pre-filtered (right) by QA. It can be observed that while filtering improved the properties of certain datasets, it degraded the properties of the others, particularly Aqua MODIS Deep Blue and OMI. This effect can be partially explained by observing that the retrieval algorithms can mistakenly assign bad QA to pixels with good retrievals, as demonstrated in a high-AOD event in the bottom part of the picture. In this figure, the magenta line indicates daily means of AERONET AOD at 440 nm, while bar heights reflect the number of all-QA (top half of the figure) and best-QA (bottom half) data pixels in each spaceborne sample, and error bars represent the mean relative accuracy of each sample computed based on its pixels. As an example, consider AMODIS DB (turquoise) retrieval on 13 December. Even though the mean AOD based on 22 pixels in this retrieval was within 10 % of the corresponding AERONET AOD, all these 22 pixels were marked as having a bad QA.
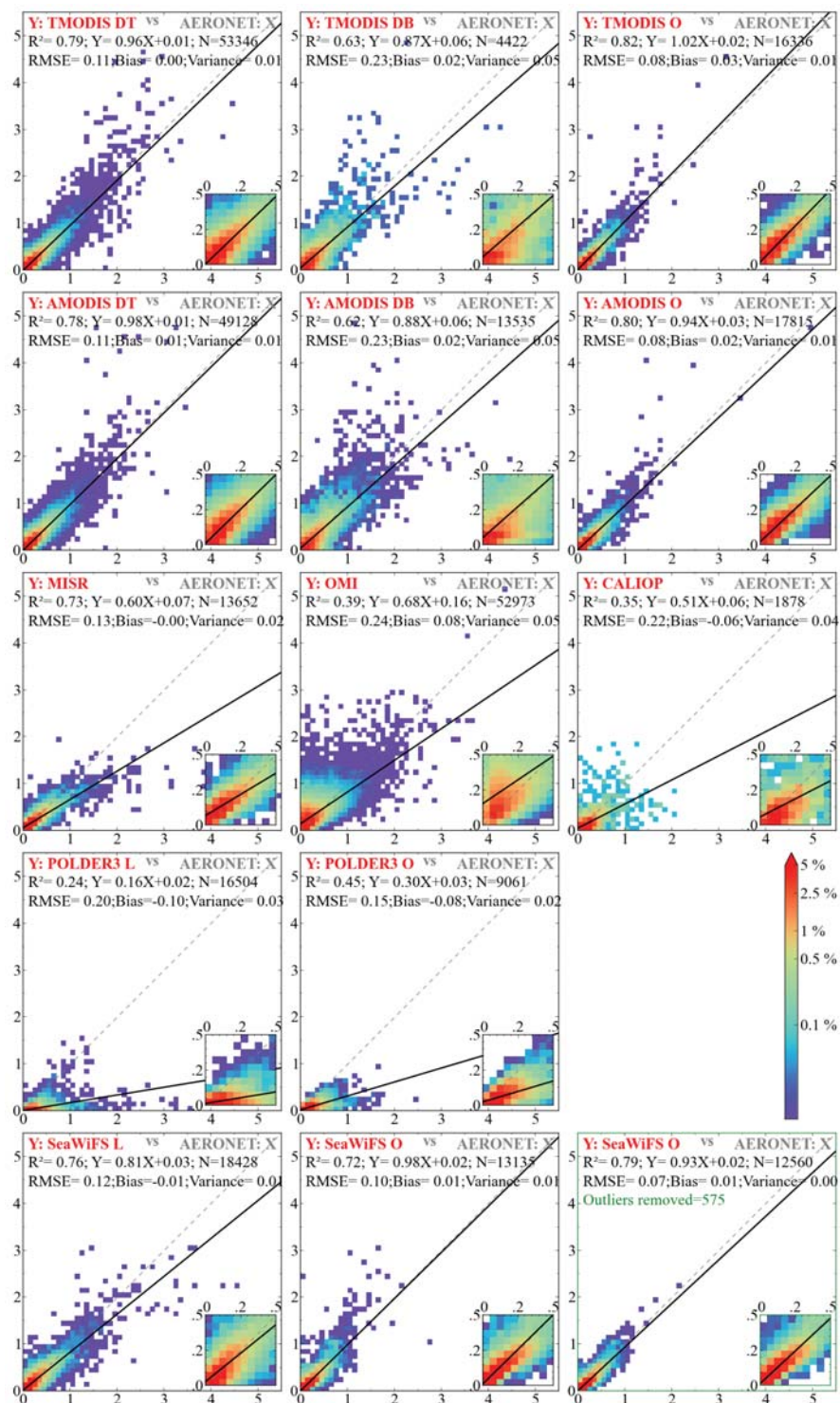
Fig. 6. Regression fits of AERONET AOD (*x* axes) to AOD measured by spaceborne sensors (*y* axes). Satellite data were pre-screened by QA as explained in Sect. 4. Density plots bin data into 0.1 AOD (0.05 AOD in magnified insets) intervals, where the color of each bin indicates the percentage of all data points that fall into this bin. Density plot in the bottom right corner (green frame) demonstrates the results of the possible data outlier detection and removal procedure described in Sect. 5, compared to the bottom middle plot. The insets (0–0.5 AOD range magnified) are intended to enhance the visualization of the linear regression fits near the origin.
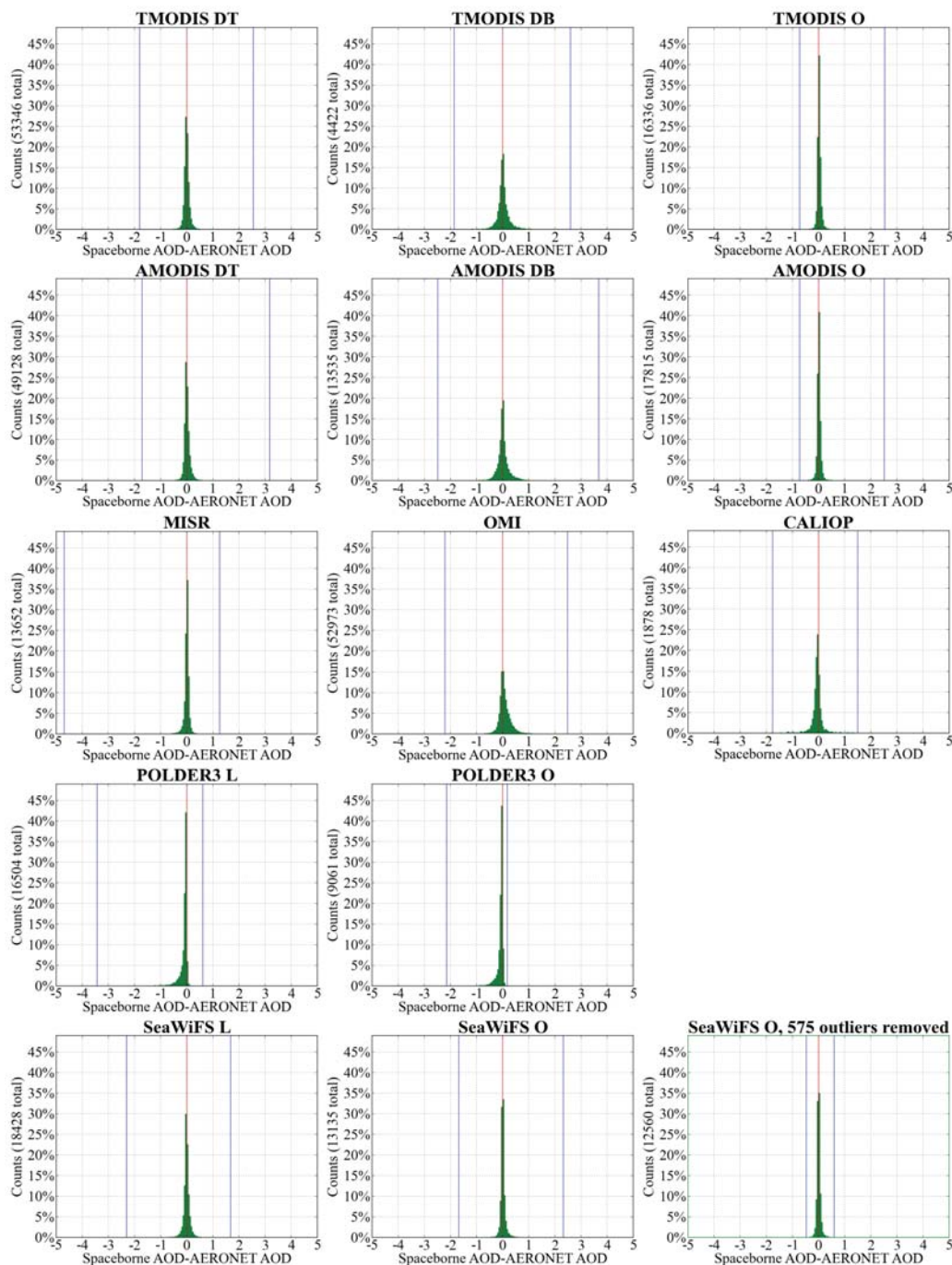
Fig. 7. Distribution of the difference (residuals) between spaceborne AOD and AERONET AOD. Satellite data were pre-screened by QA as explained in Sect. 4. In each histogram, the data are split into equal-length bins of 0.05 AOD. The red vertical line indicates the residual of 0 AOD, while the blue lines mark minimum and maximum residuals of each distribution. Histogram in the green frame demonstrates the results of the possible data outlier detection and removal procedure described in Sect. 5.

same units as AOD), with the exception of MODIS Deep Blue products that have RMSE of 0.23. Removing possible outliers improves (reduces) the RMSE of all products by 13–50 %. This indicates an opportunity for improvement of the aerosol data products by adjusting the retrieval algorithms in

the areas with the highest concentrations of the possible outliers.

For most of the sensors in Table 3 and Fig. 5, the slope of the fitted regression line is below 1.0 and the intercept is slightly above 0. This can be explained by the limitations of
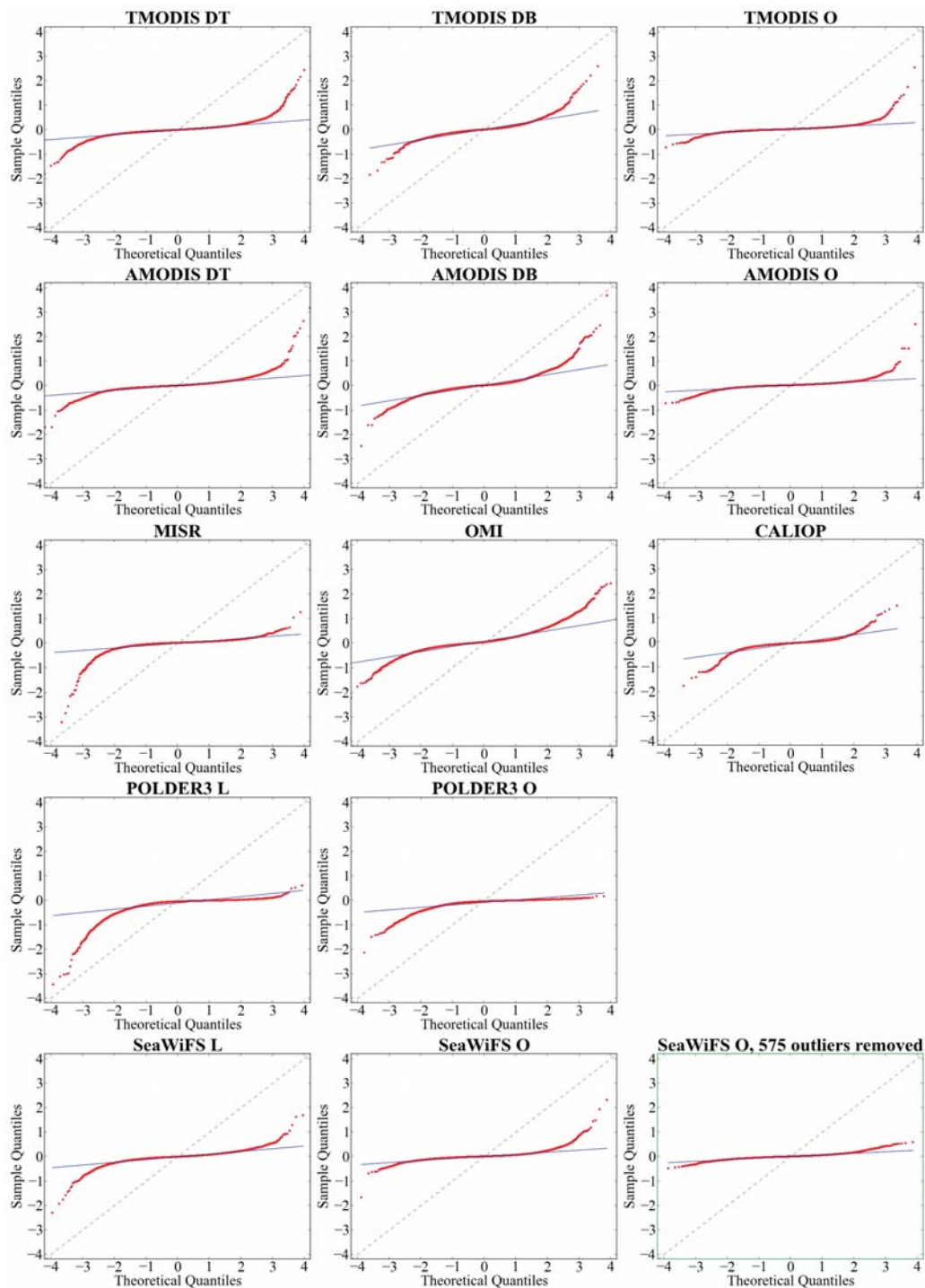
Fig. 8. Normality of the difference between spaceborne AOD and AERONET AOD. In each plot, points closely following the blue fitted line indicate the data that are approximately normally distributed. Curvatures around the center of the straight line represent the departure from the normality and indicate a presence of possible outliers, particularly at the tails of the distributions. The difference in the slope and offset of the fitted blue line from the gray 1 : 1 line indicates a deviation from the standard location (i.e., mean = 0) and scale (i.e., standard deviation = 1) of the normal distribution. Satellite data were pre-screened by QA as explained in Sect. 4. Plot in the green frame demonstrates the results of the possible data outlier detection and removal procedure described in Sect. 5.
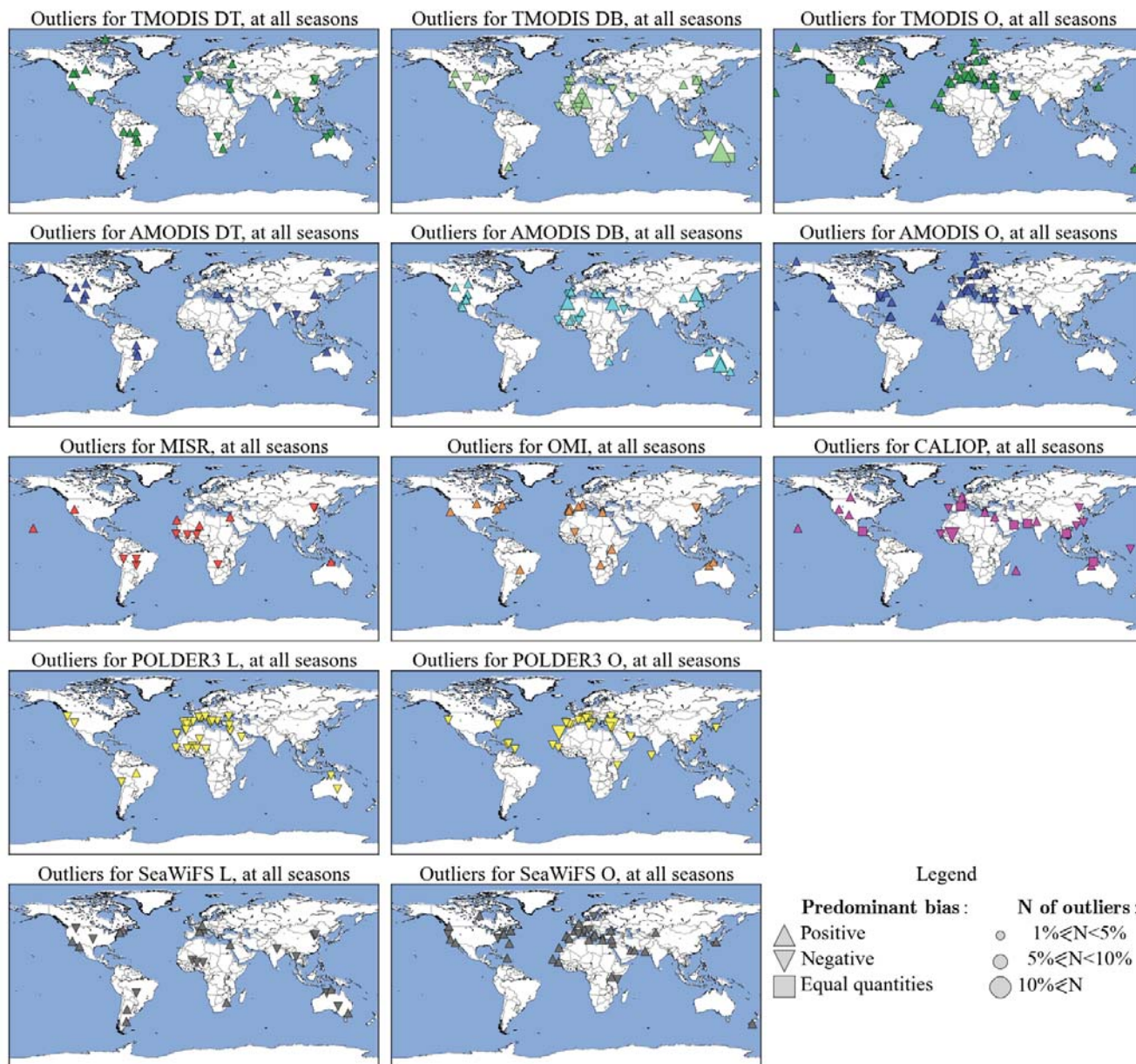
Fig. 9. Distribution of the possible data outliers for the studied spaceborne aerosol datasets. Displayed values are percentages from all outliers detected for each of the datasets as listed in the fourth column of Table 3. Stations with less than 1 % from the total number of outliers are not shown. The statistical technique for detection and removal of the possible data outliers is described in Sect. 5.

the spaceborne retrieval algorithms that tend to (1) overestimate low-AOD events when the AOD signal is very weak and almost indiscernible from the surface signal, resulting in a portion of the surface signal being mistaken for an AOD signal; and (2) underestimate high-AOD events because of the very weak surface signal, where a portion of the AOD signal might be mistaken for a surface signal. Furthermore, certain algorithms have pre-set limits on the highest possible retrieved value of AOD (e.g., 3.0 in MISR), which may further affect the reported statistics. Finally, certain censors

have peculiar features that impact the overall characteristics of their data. Among such features are sensitivity to sub-pixel cloud contamination in OMI retrievals that leads to an over-estimation of AOD (Torres et al., 1998), sensitivity to fine particles in POLDER land retrievals that leads to underestimation of AOD in coarse-mode-dominated regions (Herman et al., 1997), and also frequent under-estimation of AOD by daytime CALIOP retrievals (Kacenelenbogen et al., 2011). Since without a ground reference it is near impossible to recognize an underestimation of high-AOD events or over-
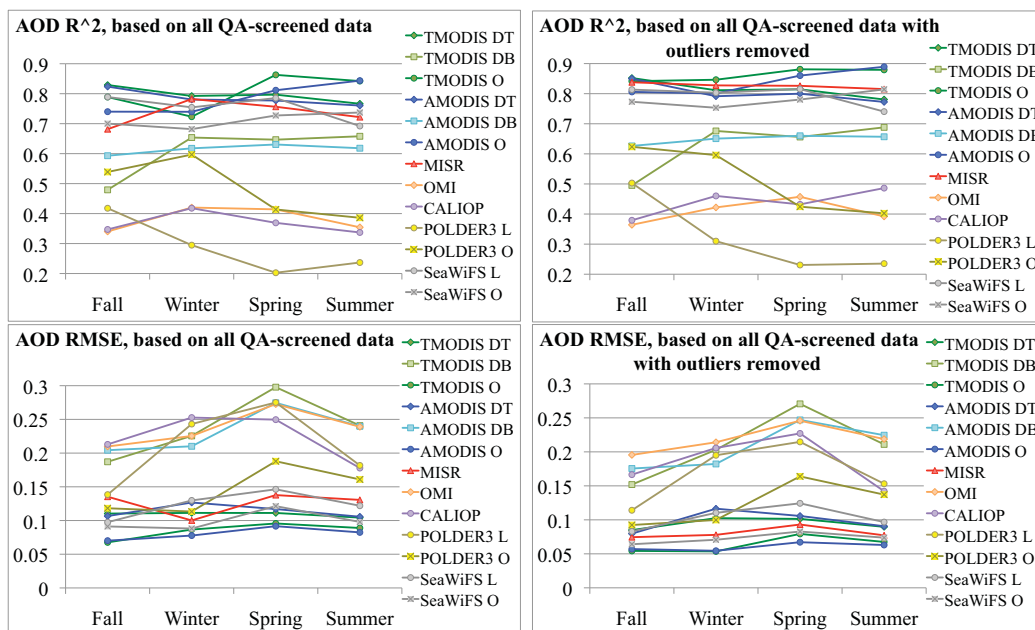
Fig. 10. Seasonal dependence of squared linear fit correlation coefficient (R^2) and root mean square error (RMSE) statistics between the collocated spaceborne and ground-based (AERONET) observations of AOD, based on the data in Table 3.

estimation of low-AOD events in the original Level 2 spaceborne data, even when considering the associated QA information, it is especially important to explore the behavior of each product across the complete range of AOD values.

Figure 10 charts the seasonal dependence of R^2 and RMSE of the spaceborne products based on the data in Table 3. While all of the products demonstrate the high seasonal variations in the statistical parameters, the OMI, CALIOP, POLDER, and MODIS Deep Blue are the most sensitive to the seasonal changes in the retrieval conditions, perhaps because of the uncertainties associated with cloud screening (Li et al., 2009), although collocating spaceborne observations with AERONET introduces certain bias towards cloud-free scenes because of the comprehensive AERONET cloud screening procedures (Smirnov et al., 2000). Furthermore, it can be seen that while removing the data outliers reduces the RMSE and sensitivity to the seasonal changes in the analyzed products, this reduction is not significant, indicating that the retrieval errors reflected by the RMSE of these products likely stem from the regular retrievals rather than the anomalous retrievals.

The accuracy of the spaceborne aerosol products might vary with the location of the retrieval and, depending on the location, some products might be significantly more accurate than others. The spatial dependence of the accuracy of the analyzed products is explored in Fig. 11 and Fig. 12, where it can be observed that no single sensor provides the best retrievals at all sites. Additionally, as indicated by the smaller relative sizes of certain markers in Fig. 11 and Fig. 12, al-

though some locations might be covered by highly accurate spaceborne retrievals from certain sensors, if such sensors offer limited coverage and data availability, their accuracy advantage may ultimately produce only limited impact, highlighting the auxiliary but still important role of the less precise but more spatially extensive products. Furthermore, as depicted by the lighter shading in Fig. 11 (e.g., southern Australia) and also in the histogram of R^2 inset in this figure, some sites are not covered by high-correlation (i.e., $R^2 \geq 0.75$) retrievals at all or have no collocated retrievals from the most accurate of the products.

Moreover, it can be observed that the best-performing aerosol products differ between Fig. 11 and Fig. 12, and the products providing the best RMSE are oftentimes those with the lower R^2. Therefore, when choosing an aerosol product for a specific analysis goal and at a specific region, it is necessary to consider a balance between a variety of seasonal, statistical, and spatial factors.

7 Accuracy of aerosol data products based on land cover type

Aerosol properties are derived from satellite observations based on a set of assumptions about the type and the optical properties of the underlying terrestrial surfaces. Therefore, it can be beneficial to compare the accuracy of the considered aerosol data products based on the land cover types of the sites over which the data subsets were extracted. As a reference for land cover types and their spatial extent, we used the
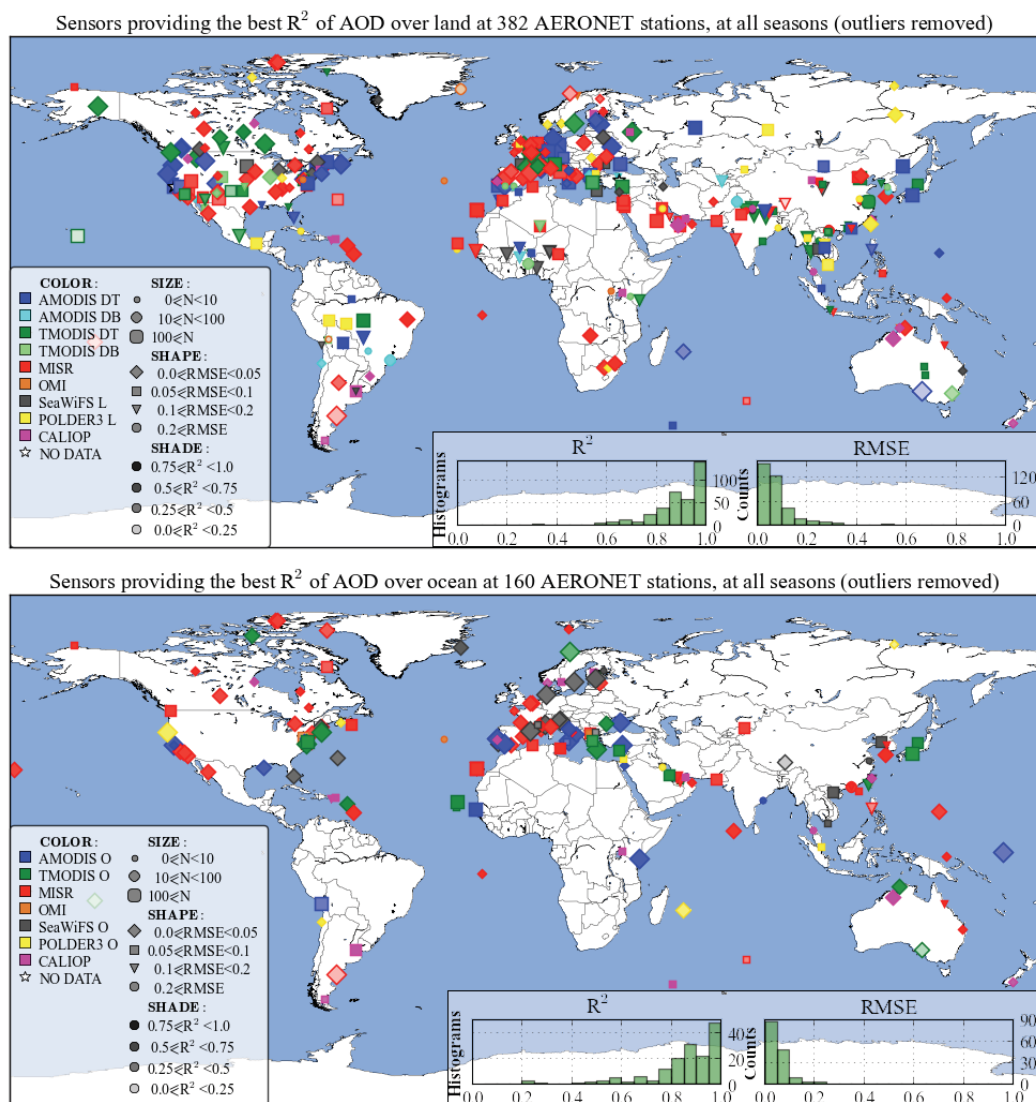
Fig. 11. Spaceborne datasets with the best correlation (R^2) of the retrieved AOD with the AOD measured by each individual inland (top) and coastal or island-based (bottom) AERONET site. The intensity of marker shading indicates the degree of correlation. Marker shape indicates the range of root mean square error (RMSE) associated with the displayed best R^2. Finally, marker size corresponds to the number of collocated data points used to compute the displayed statistics. Histograms in the bottom insets highlight the distribution of these statistics over all sites based on bins of 0.05 AOD. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5.

global dataset that is based on the International Geosphere-Biosphere Programme (IGBP) classification scheme and is available from the suite of MODIS products (Friedl et al., 2002). For each land cover type, we identified coincident AERONET stations and averaged their corresponding statistical results from Sect. 6. Tables 4 and 5 list the results of this aggregation, while Fig. 13 and Fig. 14 outline these results on a geographical map.

Generally, these aggregated results corroborate the findings of Sect. 6, and the aerosol products from MODIS and MISR sensors produce the most accurate results for the ma-

jority of the land cover types, although there are some peculiarities that should be discussed in greater detail in order to understand better the best areas of application of the analyzed aerosol products.

Specifically, IGBP *water* surface locations include 31 AERONET stations out of 160 stations with collocated ocean retrievals identified in Fig. 2. At these 31 locations, MODIS, MISR, and SeaWiFS demonstrate the best results with $R^2 \approx 0.7$. Furthermore, POLDER ocean dataset has a good RMSE = 0.07 (Deuzé et al., 1999) that is comparable to the best performing sensors in this region, albeit it

Sensors providing the best RMSE of AOD over land at 382 AERONET stations, at all seasons (outliers removed)

Sensors providing the best RMSE of AOD over ocean at 160 AERONET stations, at all seasons (outliers removed)
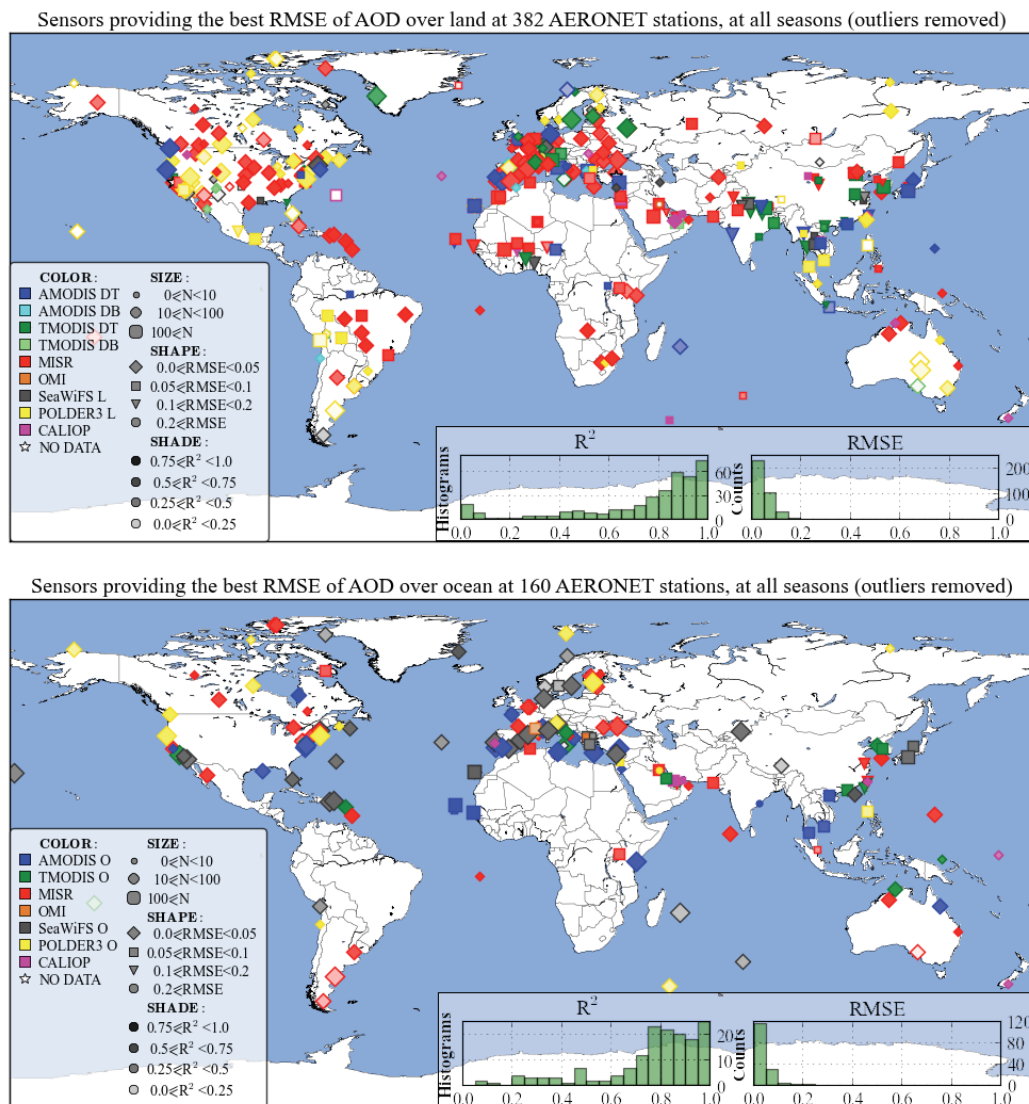

Fig. 12. Spaceborne datasets with the best root mean square error (RMSE) of the retrieved AOD to the AOD measured by each individual inland (top) and coastal or island-based (bottom) AERONET site. The symbols used are the same as the symbols in Fig. 11. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5.

has a somewhat lower squared correlation coefficient value of $R^2 = 0.62$; note that these statistics are different from POLDER ocean statistics in Fig. 6, which analyzes a more complete set of AERONET stations. It is interesting to note that the correlation between AERONET and Aqua MODIS AOD with $R^2 = 0.8$ is higher than the correlation between AERONET and Terra MODIS AOD with $R^2 = 0.72$. A detailed inspection of the data showed that this difference stems from several AERONET sites with relatively small numbers of collocated data points ($N < 35$) and the average AOD below 0.2. Under such low-AOD conditions, MODIS ocean algorithm has difficulty in retrieving the precise AOD values and, as a result, is subject to an increased rate of errors (Kleidman et al., 2005; Remer, 2002).

Evergreen broadleaf forest regions provide conditions that are favorable for retrieving AOD, and multiple sensors demonstrate the high correlation with AERONET, including MODIS Dark Target with $R^2 = 0.84$, MISR with $R^2 = 0.90$, and POLDER with $R^2 = 0.98$. However, since these regions are also susceptible to complex smoke events (e.g., Ji Parana, Brazil), sometimes combined with dust and pollution events (e.g., Anmyon, South Korea; Hong Kong, China), most of the sensors demonstrate a rather poor RMSE (Hyer et al., 2011). The important exception is POLDER dataset that has RMSE $= 0.07$, possibly because POLDER is especially sensitive to small particles produced by biomass burning and anthropogenic pollution sources (Fan et al., 2008), thereby retrieving fairly accurate AOD values at Ji Parana, Brazil, and
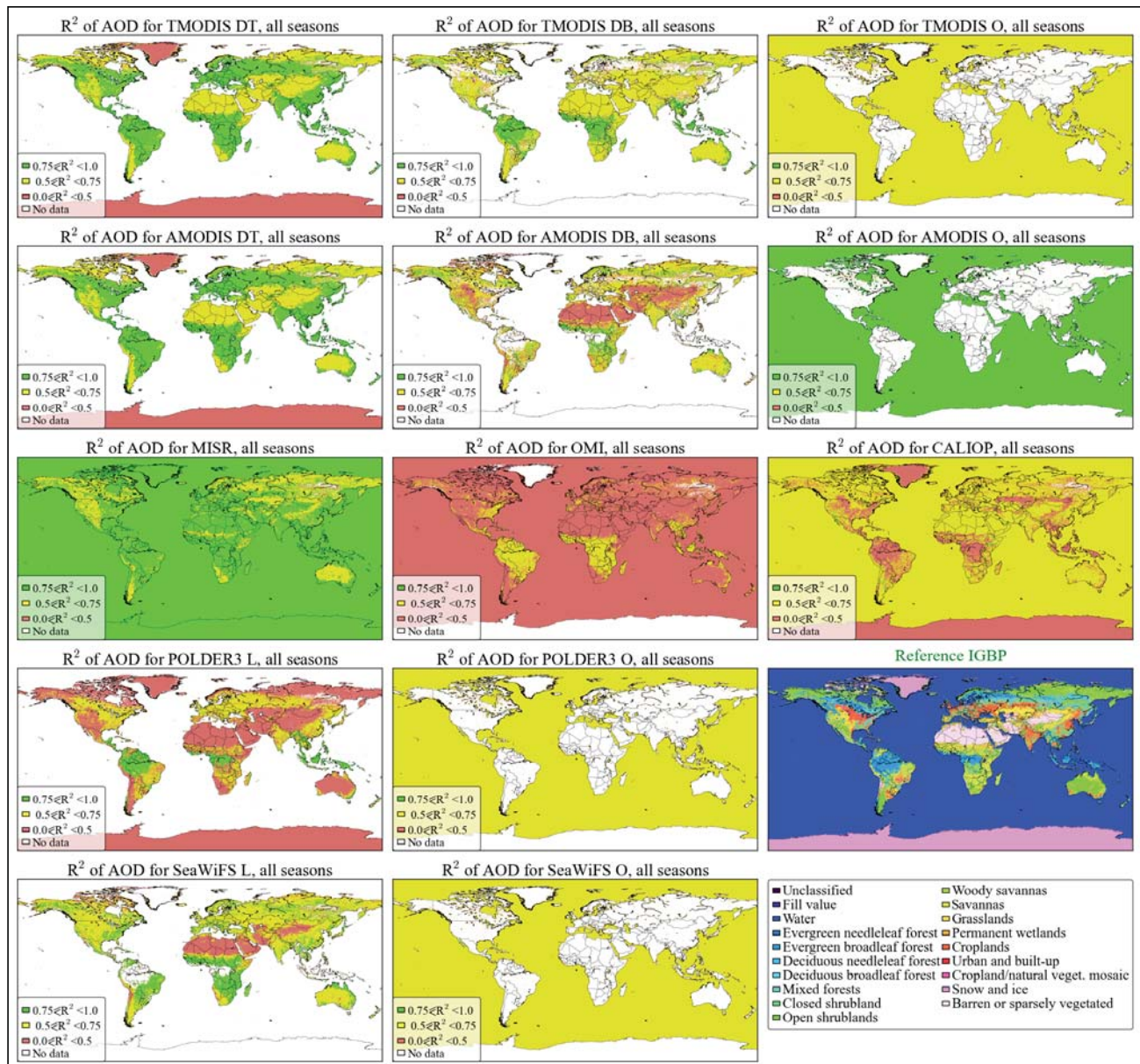
Fig. 13. Land cover type dependence of squared linear fit correlation coefficient (R^2) between the collocated spaceborne and ground-based (AERONET) observations of AOD. Areas corresponding to each IGBP land cover type (bottom right inset) are colored based on the average of the data from those AERONET sites that reside in these areas. The statistics were calculated based on data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5.

Lulin, Taiwan. It should be also noted that together with deciduous broadleaf forests and savannas, evergreen broadleaf forest is one of the three land cover types where POLDER demonstrates good results with $R^2 \approx 0.98$, indicating the advantage of polarization measurements for aerosol retrievals over these regions.

For *mixed forests*, MODIS Dark Target products provide the highest retrieval accuracy with $R^2 = 0.78$ for Terra and 0.82 for Aqua, while MISR data are somewhat less ac-

curate with $R^2 = 0.72$ as a result of underestimating high AODs during summertime biomass burning events (Kahn et al., 2010), although RMSE = 0.04 of MISR is better than RMSE = 0.05 of Terra MODIS and RMSE = 0.06 of Aqua MODIS. Sufficiently reliable aerosol data are also retrieved by SeaWiFS with $R^2 = 0.64$, by POLDER with $R^2 = 0.62$, and CALIOP with $R^2 = 0.61$.

For *closed shrubland*, MISR with $R^2 = 0.90$, CALIOP with $R^2 = 0.88$, and MODIS Deep Blue with $R^2 = 0.74$ for
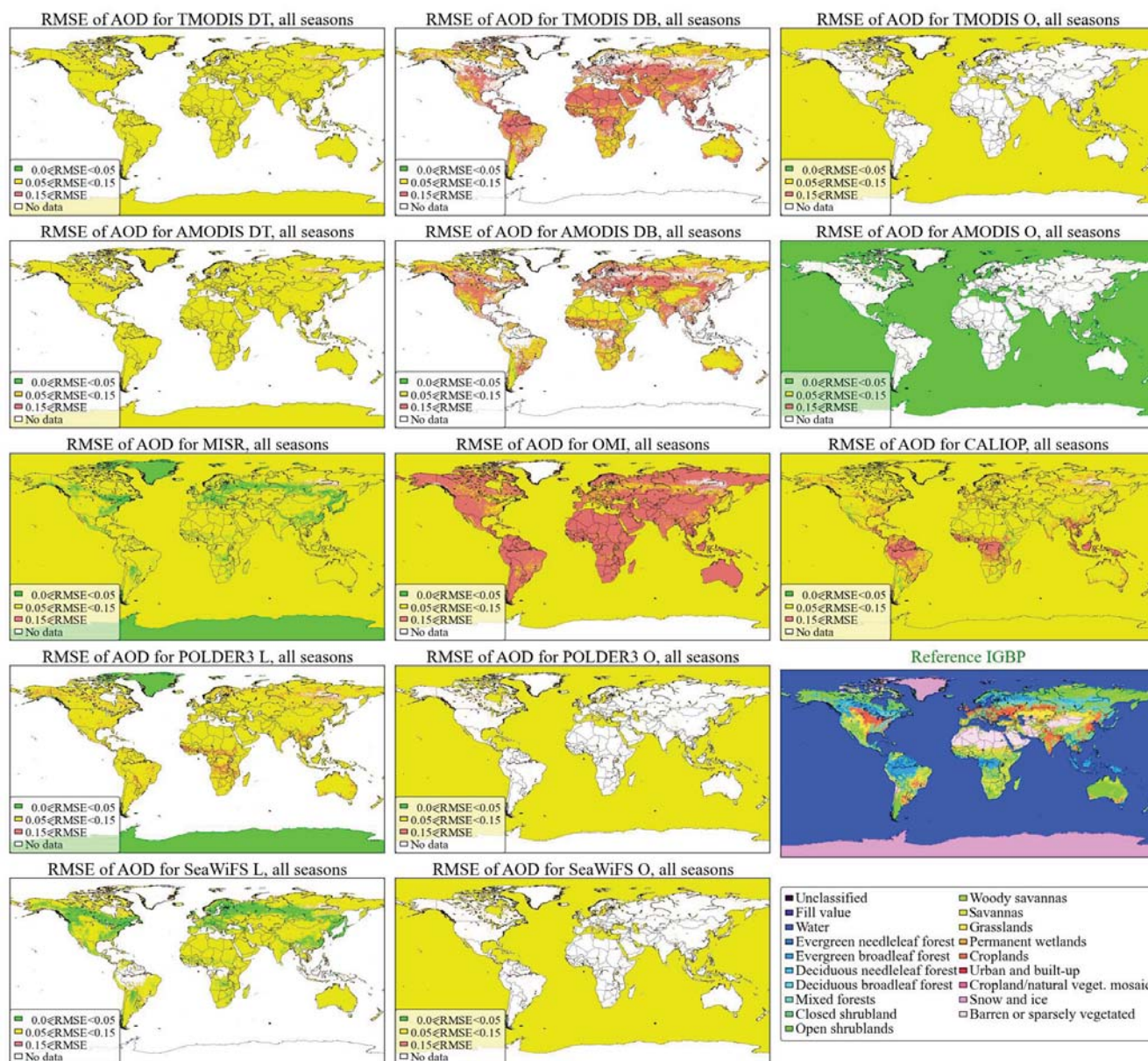
Fig. 14. Land cover type dependence of root mean square error (RMSE) between the collocated spaceborne and ground-based (AERONET) observations of AOD. Areas corresponding to each IGBP land cover type (bottom right inset) are colored based on the average of the data from those AERONET sites that reside in these areas. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5.

Terra and $R^2 = 0.84$ for Aqua produce the best results. Although MODIS Deep Blue shows a better performance than MODIS Dark Target for this land cover type, the Deep Blue products are retrieved only over a single Lake Argyle AERONET site in northern Australia, whereas Dark Target products are retrieved over 7 sites and have a significantly larger number of data points. Likewise, the good result demonstrated by CALIOP also originates exclusively from the Lake Argyle retrievals. The difference of 0.1 in R^2 between MODIS Terra Deep Blue and MODIS Aqua Deep

Blue can be partly explained by the difference in the data availability of these two datasets, as MODIS Terra Deep Blue at the time of this work was available only through 2007; this effect can be also observed for several other land cover types, where MODIS Aqua Deep Blue tends to have a lower correlation with AERONET and produces results that are closer to the results of SeaWiFS, probably because the latter is also based on the Deep Blue retrieval algorithm (Hsu et al., 2006). It should be noted that closed shrubland is the only area where CALIOP produces some of the best retrievals,
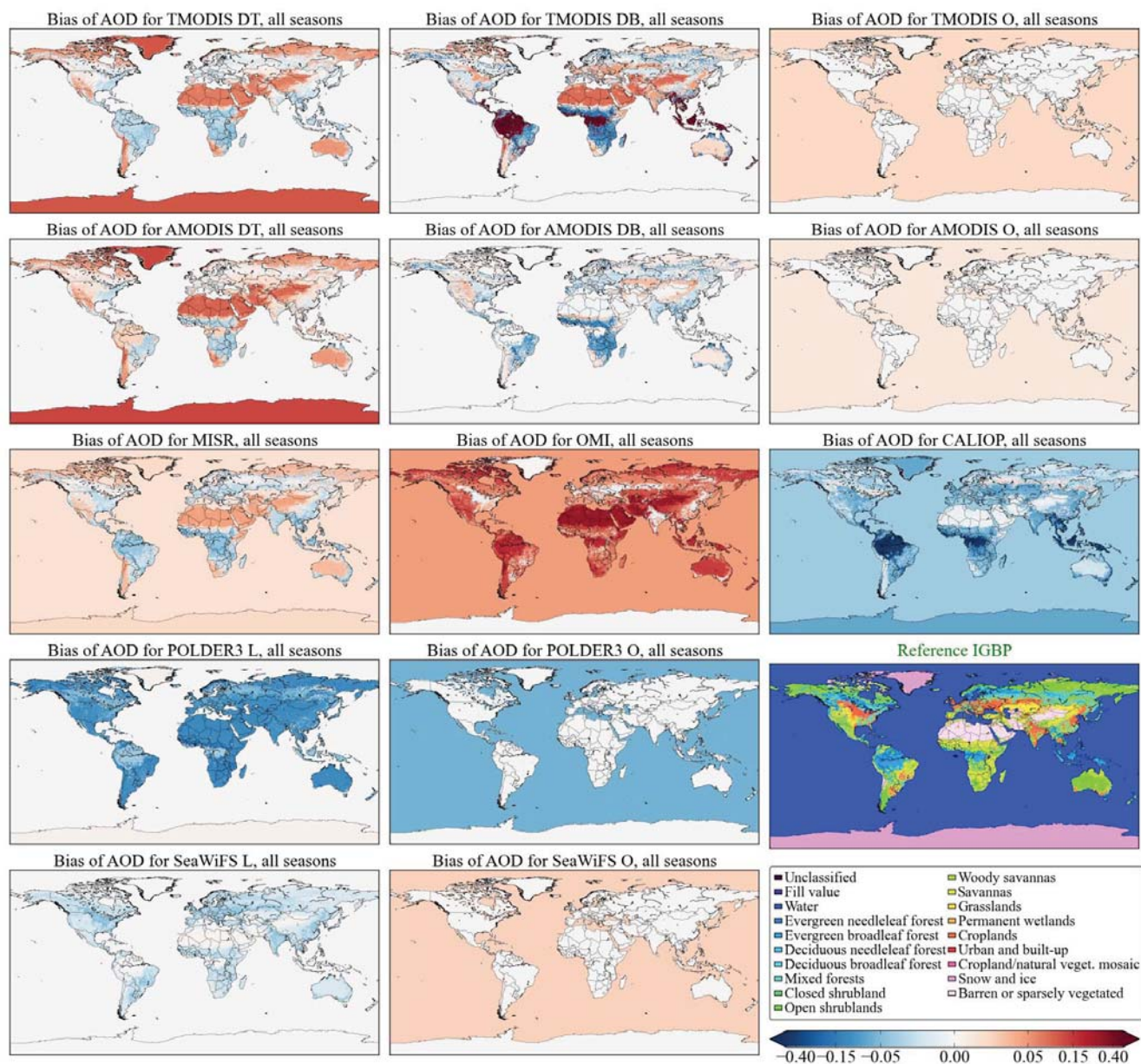
Fig. 15. Land cover type dependence of bias between the collocated spaceborne and ground-based (AERONET) observations of AOD. Areas corresponding to each IGBP land cover type (bottom right inset) are colored based on the average of the data from those AERONET sites that reside in these areas. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5.

possibly indicating the advantage of active aerosol sensing over this bright-surface region.

Over *wooded savannas*, both Dark Target and Deep Blue products from MODIS show very good results with R^2 values between 0.80 and 0.90. MISR with $R^2 = 0.79$ and SeaWiFS with $R^2 = 0.77$ produce lower but still reasonably good results. The reduced performance of MISR in this region can be explained by the lack of region-specific aerosol mixtures in its retrieval algorithm, a situation that is expected to be improved in future revisions of the product (Kahn et al., 2009).

It should be also noted that this region enables one of the two highest correlations between OMI and AERONET observations, probably as a result of favorable cloud-free conditions in sub-Saharan Africa (Ahn et al., 2008; Torres et al., 2007).

Open shrublands are very dry and sparsely vegetated regions that are characterized by bright surfaces. Such regions present a great challenge for remote retrieval of aerosol properties (Kahn et al., 2009), and none of the analyzed products exceeded the correlation coefficient of 0.7. Among the best-performing products are MODIS Dark Target with $R^2 = 0.68$
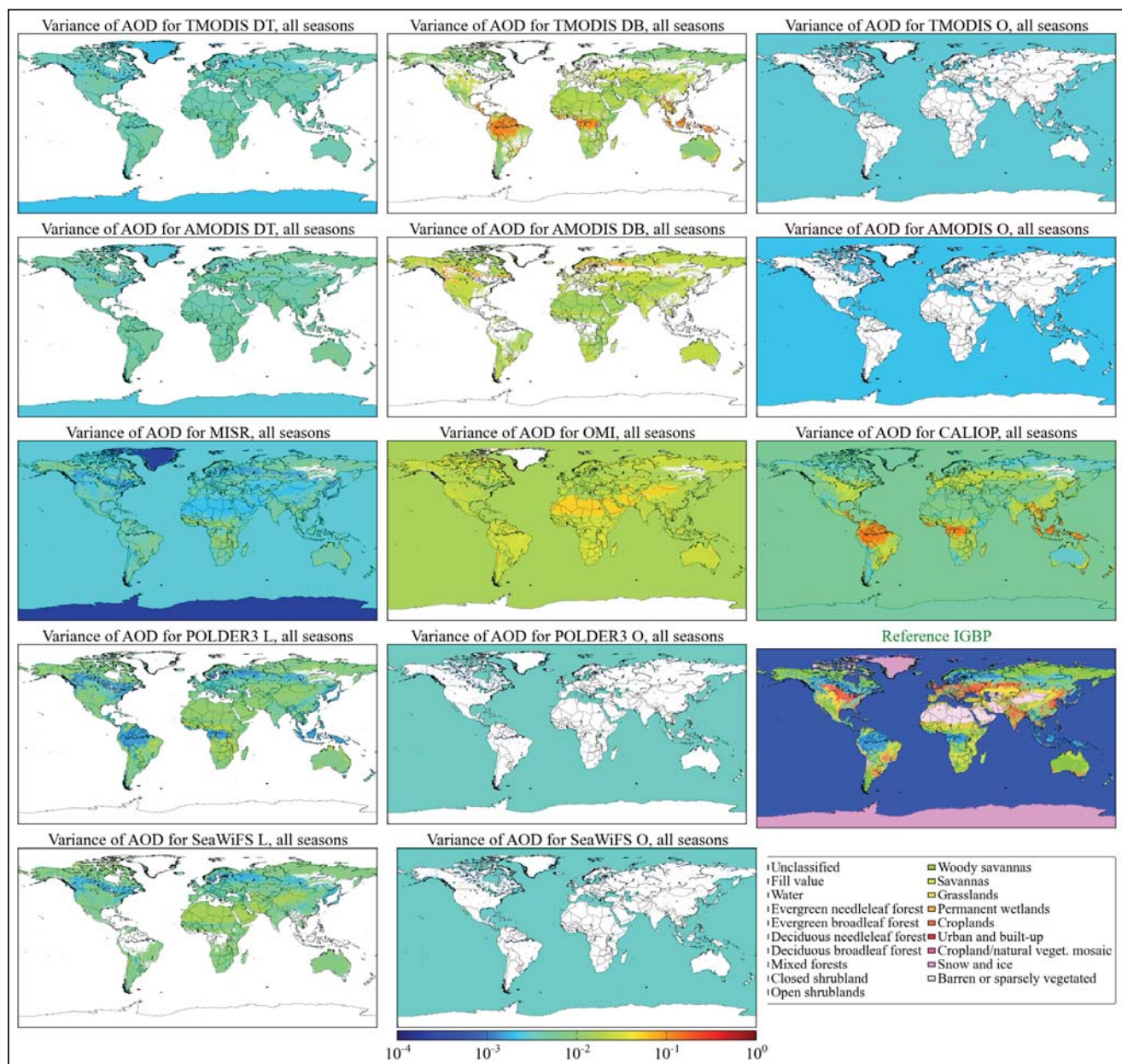
Fig. 16. Land cover type dependence of variance between the collocated spaceborne and ground-based (AERONET) observations of AOD. Areas corresponding to each IGBP land cover type (bottom right inset) are colored based on the average of the data from those AERONET sites that reside in these areas. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5.

for Terra and $R^2 = 0.61$ for Aqua, MISR with $R^2 = 0.65$, and MODIS Deep Blue with $R^2 = 0.52$ for Terra and $R^2 = 0.61$ for Aqua, as well as CALIOP with $R^2 = 0.59$.

Similar to open shrublands, *grasslands* were challenging to all of the sensors, where Terra MODIS Deep Blue, Aqua MODIS Dark Target, and MISR demonstrated the best results with R^2 values between 0.65 and 0.67. Even more challenging were *snow and ice* and also *barren or sparsely vegetated* areas, where MISR was the only highly accurate

aerosol product with $R^2 = 0.78$ for both land cover types, thanks to its multi-angle measurement capabilities that allow retrieving aerosol properties over bright surfaces and enable the advanced cloud and ice detection capabilities (Kahn et al., 2009).

Table 4. Linear fit correlation coefficient (R^2) between the collocated spaceborne and ground-based observations of AOD estimated at the stations that coincide with different IGBP land cover types. Empty cells indicate no collocated data available from a specific sensor over a specific land cover type. No AERONET stations are available at the areas occupied by deciduous needleleaf forest. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5. A graphical representation of this table is in Fig. 13. Please see the Supplement for a breakdown of the listed statistics based on the nominal boreal seasons.

	TMODIS DT	TMODIS DB	TMODIS O	AMODIS DT	AMODIS DB	AMODIS O	MISR	OMI	CALIOP	POLDER3 L	POLDER3 O	SeaWiFS L	SeaWiFS O
Water			0.72			0.80	0.78	0.41	0.59		0.62		0.72
Evergreen needleleaf forest	0.79			0.78	0.71		0.74	0.34	0.54	0.51		0.67	
Evergreen broadleaf forest	0.84	1.00		0.84			0.90	0.60	0.14	0.98			
Deciduous broadleaf forest	0.84			0.88			0.82	0.56	0.27	0.74		0.83	
Mixed forests	0.78			0.82			0.72	0.39	0.61	0.62		0.64	
Closed shrubland	0.51	0.74		0.64	0.84		0.90	0.50	0.88	0.55		0.60	
Open shrublands	0.68	0.52		0.61	0.61		0.65	0.31	0.59	0.31		0.53	
Woody savannas	0.80	0.91		0.85	0.86		0.79	0.59	0.34	0.42		0.77	
Savannas	0.76	0.56		0.82	0.67		0.78	0.55	0.63	0.73		0.80	
Grasslands	0.57	0.65		0.67	0.43		0.68	0.49	0.40	0.43		0.52	
Permanent wetlands	0.75			0.76			0.77	0.32	0.62	0.54		0.02	
Croplands	0.79	0.72		0.78	0.64		0.78	0.46	0.57	0.55		0.69	
Urban and built-up	0.69	0.65		0.70	0.57		0.76	0.43	0.51	0.46		0.63	
Cropland/natural veget. mosaic	0.75			0.79	0.50		0.72	0.54	0.46	0.52		0.72	
Snow and ice	0.26			0.27			0.78		0.03	0.22			
Barren or sparsely vegetated	0.60	0.56		0.62	0.34		0.78	0.30	0.58	0.22		0.37	

Table 5. Root mean square error (RMSE) between the collocated spaceborne and ground-based observations of AOD estimated at the stations that coincide with different IGBP land cover types. Empty cells indicate no collocated data available from a specific sensor over a specific land cover type. No AERONET stations are available at the areas occupied by deciduous needleleaf forest. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5. A graphical representation of this table is in Fig. 14. Please see the Supplement for a breakdown of the listed statistics based on the nominal boreal seasons.

	TMODIS DT	TMODIS DB	TMODIS O	AMODIS DT	AMODIS DB	AMODIS O	MISR	OMI	CALIOP	POLDER3 L	POLDER3 O	SeaWiFS L	SeaWiFS O
Water			0.06			0.05	0.06	0.14	0.11		0.07		0.07
Evergreen needleleaf forest	0.06			0.06	0.30		0.06	0.17	0.08	0.05		0.05	
Evergreen broadleaf forest	0.09	0.80		0.08			0.10	0.28	0.59	0.07			
Deciduous broadleaf forest	0.06			0.06			0.04	0.11	0.11	0.06		0.05	
Mixed forests	0.05			0.06			0.04	0.13	0.13	0.06		0.05	
Closed shrubland	0.09	0.06		0.08	0.04		0.06	0.19	0.04	0.11		0.06	
Open shrublands	0.10	0.10		0.10	0.13		0.08	0.22	0.06	0.13		0.09	
Woody savannas	0.08	0.27		0.09	0.23		0.13	0.21	0.18	0.24		0.11	
Savannas	0.10	0.15		0.09	0.14		0.07	0.22	0.11	0.13		0.09	
Grasslands	0.09	0.18		0.09	0.16		0.05	0.19	0.11	0.13		0.08	
Permanent wetlands	0.06			0.06			0.05	0.14	0.11	0.06		0.13	
Croplands	0.09	0.16		0.09	0.17		0.07	0.16	0.14	0.11		0.09	
Urban and built-up	0.09	0.14		0.09	0.15		0.07	0.20	0.13	0.10		0.10	
Cropland/natural veget. mosaic	0.07			0.08	0.13		0.06	0.15	0.22	0.08		0.11	
Snow and ice	0.11			0.14			0.02		0.10	0.01			
Barren or sparsely vegetated	0.12	0.17		0.13	0.13		0.06	0.37	0.10	0.14		0.13	

Table 6. Bias between the collocated spaceborne and ground-based observations of AOD estimated at the stations that coincide with different IGBP land cover types. Empty cells indicate no collocated data available from a specific sensor over a specific land cover type. No AERONET stations are available at the areas occupied by Deciduous needleleaf forest. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5. A graphical representation of this table is in Fig. 14. Please see the Supplement for a breakdown of the listed statistics based on the nominal boreal seasons.

	TMODIS DT	TMODIS DB	TMODIS O	AMODIS DT	AMODIS DB	AMODIS O	MISR	OMI	CALIOP	POLDER3 L	POLDER3 O	SeaWiFS L	SeaWiFS O
Water			0.02			0.01	0.02	0.05	−0.04		−0.06		0.02
Evergreen needleleaf forest	0.00			0.01	0.01		0.00	0.06	−0.01	−0.05		−0.03	
Evergreen broadleaf forest	−0.03	0.73		0.02			−0.04	0.18	−0.47	−0.04			
Deciduous broadleaf forest	−0.03			−0.03			−0.01	0.02	−0.04	−0.05		−0.02	
Mixed forests	−0.02			0.00			0.01	0.06	0.03	−0.05		−0.02	
Closed shrubland	−0.01	0.00		0.02	0.02		0.03	0.11	−0.03	−0.09		−0.03	
Open shrublands	0.05	0.02		0.05	0.01		0.03	0.14	−0.02	−0.11		−0.02	
Woody savannas	−0.04	−0.23		−0.03	−0.19		−0.08	0.03	−0.12	−0.20		−0.03	
Savannas	−0.05	−0.10		−0.04	−0.08		−0.02	0.08	−0.04	−0.11		−0.02	
Grasslands	0.01	−0.02		0.02	0.03		0.01	0.11	−0.06	−0.11		−0.01	
Permanent wetlands	−0.02			0.01			0.01	0.05	−0.09	−0.05		−0.01	
Croplands	0.00	0.04		0.01	−0.02		−0.03	−0.01	−0.07	−0.09		−0.04	
Urban and built-up	−0.01	−0.04		0.00	−0.02		−0.03	0.03	−0.05	−0.09		−0.03	
Cropland/natural veget. mosaic	−0.02			−0.03	0.00		−0.02	0.00	−0.02	−0.07		−0.07	
Snow and ice	0.10			0.13			0.02		−0.07	0.00			
Barren or sparsely vegetated	0.07	0.08		0.09	0.00		0.04	0.29	−0.01	−0.11		0.00	

Table 7. Variance between the collocated spaceborne and ground-based observations of AOD estimated at the stations that coincide with different IGBP land cover types. Empty cells indicate no collocated data available from a specific sensor over a specific land cover type. No AERONET stations are available at the areas occupied by Deciduous needleleaf forest. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sects. 4 and 5. A graphical representation of this table is in Fig. 14. Please see the Supplement for a breakdown of the listed statistics based on the nominal boreal seasons.

	TMODIS DT	TMODIS DB	TMODIS O	AMODIS DT	AMODIS DB	AMODIS O	MISR	OMI	CALIOP	POLDER3 L	POLDER3 O	SeaWiFS L	SeaWiFS O
Water			0.003			0.002	0.003	0.016	0.006		0.003		0.003
Evergreen needleleaf forest	0.003			0.003	0.088		0.001	0.021	0.004	0.001		0.002	
Evergreen broadleaf forest	0.004	0.112		0.005			0.004	0.027	0.125	0.001			
Deciduous broadleaf forest	0.003			0.002			0.001	0.012	0.010	0.001		0.002	
Mixed forests	0.002			0.003			0.001	0.013	0.022	0.002		0.002	
Closed shrubland	0.005	0.003		0.003	0.001		0.002	0.022	0.001	0.005		0.002	
Open shrublands	0.004	0.008		0.006	0.024		0.004	0.026	0.003	0.007		0.008	
Woody savannas	0.004	0.020		0.004	0.014		0.019	0.024	0.017	0.027		0.014	
Savannas	0.008	0.017		0.006	0.014		0.005	0.040	0.018	0.013		0.009	
Grasslands	0.005	0.029		0.003	0.026		0.002	0.020	0.008	0.009		0.006	
Permanent wetlands	0.003			0.004			0.002	0.018	0.004	0.001		0.016	
Croplands	0.006	0.027		0.007	0.028		0.005	0.021	0.022	0.005		0.006	
Urban and built-up	0.006	0.015		0.007	0.016		0.004	0.030	0.011	0.004		0.007	
Cropland/natural veget. mosaic	0.004			0.004	0.019		0.002	0.021	0.042	0.003		0.008	
Snow and ice	0.002			0.003			0.000		0.005	0.000			
Barren or sparsely vegetated	0.004	0.019		0.004	0.014		0.002	0.054	0.007	0.011		0.015	

8 Conclusions

In this paper, we analyzed and intercompared 11 spaceborne aerosol products from MODIS, MISR, OMI, SeaWiFS, POLDER, and CALIOP sensors, which were sampled fairly uniformly based on the MAPSS framework that was used to collocate these spaceborne observations with ground-based AERONET observations during the period of 2006-06-07 and 2010-12-11, when all the sensors were operational. Based on this analysis, for each of the AERONET stations, we identified products providing the best correlation coefficient (R^2) and root mean square error (RMSE). It was found that no single product provides the best retrieval over all sites, and certain sites are not covered by accurate retrievals at all. Furthermore, it was observed that a product providing the best R^2 at a certain location does not always provide the best RMSE at the same location. Therefore, to facilitate the multivariate analysis that is necessary when choosing the most suitable spaceborne aerosol product at a specific region, we plan to develop an interactive tool that would allow exploration of the multi-sensor collocated data on an interactive map.

Further, a statistical approach based on the statistical modified Z-score test has been used to identify automatically possible data outliers in the collocated datasets. The reported analysis shows that even though such atypical data points constitute a relatively minor portion (2–7 %) of the analyzed datasets, they can significantly bias the results of the statistical analysis. For this reason, it is suggested that such data points be set aside when analyzing collocated datasets and inspected separately, in order to develop appropriate mitigation measures in the retrieval algorithms or to design specific data screening strategies that could be used to identify outliers in spaceborne datasets independently and systematically.

Finally, we assessed the accuracy of the spaceborne aerosol products based on their spatial distribution relative to different surface types derived from MODIS using the IGBP land cover classification scheme. This analysis identified sensors that retrieve the most accurate aerosol properties over each of the defined land cover types and highlighted the differences that exist between the sensors, providing an advantage or disadvantage in retrieving AOD over a particular land cover type. Notably, some of the land cover types, including open shrublands and grasslands, had only moderately accurate retrievals, indicating the need for improved spaceborne aerosol remote sensing instrumentation/approaches and/or retrieval algorithms.

Supplementary material related to this article is available online at: http://www.atmos-chem-phys.net/13/6777/2013/acp-13-6777-2013-supplement.pdf.

References

Ahn, C., Torres, O., and Bhartia, P. K.: Comparison of Ozone Monitoring Instrument UV Aerosol Products with Aqua/Moderate Resolution Imaging Spectroradiometer and Multiangle Imaging Spectroradiometer observations in 2006, J. Geophys. Res., 113, D16S27, doi:10.1029/2007JD008832, 2008.

Bréon, F. M., Buriez, J. C., Couvert, P., Deschamps, P. Y., Deuzé, J. L., Herman, M., Goloub, P., Leroy, M., Lifermann, A., Moulin, C., Parol, F., Sèze, G., Tanré, D., Vanbauce, C., and Vesperini, M.: Scientific results from the Polarization and Directionality of the Earth's Reflectances (POLDER), Adv. Space. Res., 30, 23832386, doi:10.1016/S0273-1177(02)80282-4, 2002.

Bréon, F.-M., Vermeulen, A., and Descloitres, J.: An evaluation of satellite aerosol products against sunphotometer measurements, Remote Sens. Environ., 115, 3102–3111, doi:10.1016/j.rse.2011.06.017, 2011.

Cheng, T., Chen, H., Gu, X., Yu, T., Guo, J., and Guo, H.: The inter-comparison of MODIS, MISR and GOCART aerosol products against AERONET data over China, Journal of Quantitative Spectroscopy and Radiative Transfer, 113, 21352145, doi:10.1016/j.jqsrt.2012.06.016, 2012.

Chu, D. A., Kaufman, Y. J., Ichoku, C., Remer, L. A., Tanré, D., and Holben, B. N.: Validation of MODIS aerosol optical

depth retrieval over land, Geophys. Res. Lett., 29, 8007, doi:10.1029/2001GL013205, 2002.

Deuzé, J. L., Herman, M., Goloub, P., Tanré, D., and Marchand, A.: Characterization of aerosols over ocean from POLDER/ADEOS-1, Geophys. Res. Lett., 26, 1421–1424, doi:10.1029/1999GL900168, 1999.

Deuzé, J. L., Bréon, F. M., Devaux, C., Goloub, P., Herman, M., Lafrance, B., Maignan, F., Marchand, A., Nadal, F., Perry, G., and Tanré, D.: Remote sensing of aerosols over land surfaces from POLDER-ADEOS-1 polarized measurements, J. Geophys. Res., 106, 4913–4926, doi:10.1029/2000JD900364, 2001.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D., and Slutsker, I.: Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations, J. Atmos. Sci., 59, 590–608, doi:10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2, 2002.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, J. Geophys. Res., 104, 31333–31349, doi:10.1029/1999JD900923, 1999.

Fan, X., Goloub, P., Deuzé, J.-L., Chen, H., Zhang, W., Tanré, D., and Li, Z.: Evaluation of PARASOL aerosol retrieval over North East Asia, Remote Sens. Environ., 112, 697–707, doi:10.1016/j.rse.2007.06.010, 2008.

Fox, J.: Applied Regression Analysis, Linear Models, and Related Methods, Sage Publications, Inc, Newbury Park, CA, 1997.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.: Global land cover mapping from MODIS: algorithms and early results, Remote Sensing of Environment, 83, 287302, doi:10.1016/S0034-4257(02)00078-0, 2002.

Gérard, B., Deuzé, J.-L., Herman, M., Kaufman, Y. J., Lallart, P., Oudard, C., Remer, L. A., Roger, B., Six, B., and Tanré, D.: Comparisons between POLDER 2 and MODIS/Terra aerosol retrievals over ocean, J. Geophys. Res., 110, 1–15, doi:10.1029/2005JD006218, 2005.

Guo, J., Xue, Y., Cao, C., Zhang, H., Guang, J., Zhang, X., and Li, X.: A synergic algorithm for retrieval of aerosol optical depth over land, Adv. Atmos. Sci., 26, 973–983, doi:10.1007/s00376-009-7218-4, 2009.

Herman, M., Deuzé, J. L., Devaux, C., Goloub, P., Bréon, F. M., and Tanré, D.: Remote sensing of aerosols over land surfaces including polarization measurements and application to POLDER measurements, J. Geophys. Res., 102, 17039–17049, doi:10.1029/96JD02109, 1997.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sens. Environ., 66, 1–16, doi:10.1016/S0034-4257(98)00031-5, 1998.

Holben, B. N., Tanré, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Castle, J. Vande, Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G.: An emerging ground-based aerosol climatology: Aerosol op-

tical depth from AERONET, J. Geophys. Res., 106, 12067–12097, doi:10.1029/2001JD900014, 2001.

Hsu, N. C., Tsay, S.-C., King, M. D., and Herman, J. R.: Aerosol Properties Over Bright-Reflecting Source Regions, IEEE T. Geosci. Remote, 42, 557–569, doi:10.1109/TGRS.2004.824067, 2004.

Hsu, N. C., Gautam, R., Sayer, A. M., Bettenhausen, C., Li, C., Jeong, M. J., Tsay, S.-C., and Holben, B. N.: Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010, Atmos. Chem. Phys., 12, 8037–8053, doi:10.5194/acp-12-8037-2012, 2012.

Hyer, E. J., Reid, J. S., and Zhang, J.: An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals, Atmos. Meas. Tech., 4, 379–408, doi:10.5194/amt-4-379-2011, 2011.

Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., Slutsker, I., and Holben, B. N.: A spatio-temporal approach for global validation and analysis of MODIS aerosol products, Geophys. Res. Lett., 29, 1–4, doi:10.1029/2001GL013206, 2002.

Ichoku, C., Remer, L. A., Kaufman, Y. J., Levy, R., Chu, D. A., Tanré, D., and Holben, B. N.: MODIS observation of aerosols and estimation of aerosol radiative forcing over southern Africa during SAFARI 2000, J. Geophys. Res., 108, 8499, doi:10.1029/2002JD002366, 2003.

Ichoku, C., Remer, L. A., and Eck, T. F.: Quantitative evaluation and intercomparison of morning and afternoon Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol measurements from Terra and Aqua, J. Geophys. Res., 110, D10S03, doi:10.1029/2004JD004987, 2005.

Iglewicz, B. and Hoaglin, D. C.: How to Detect and Handle Outliers, American Society for Quality Control, Milwaukee, WI, 1993.

Kacenelenbogen, M., Vaughan, M. A., Redemann, J., Hoff, R. M., Rogers, R. R., Ferrare, R. A., Russell, P. B., Hostetler, C. A., Hair, J. W., and Holben, B. N.: 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, Atmos. Chem. Phys., 11, 3981—4000, doi:10.5194/acp-11-3981-2011, 2011.

Kahn, R. A.: Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations, J. Geophys. Res., 110, 1–16, doi:10.1029/2004JD004706, 2005.

Kahn, R. A., Garay, M. J., Nelson, D. L., Yau, K. K., Bull, M. A., Gaitley, B. J., Martonchik, J. V, and Levy, R. C.: Satellite-derived aerosol optical depth over dark water from MISR and MODIS: Comparisons with AERONET and implications for climatological studies, J. Geophys. Res., 112, D18205, doi:10.1029/2006JD008175, 2007.

Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V, Paradise, S. R., Hansen, E. G., and Remer, L. A.: MISR Aerosol Product Attributes and Statistical Comparisons With MODIS, IEEE T. Geosci. Remote, 47, 4095–4114, doi:10.1109/TGRS.2009.2023115, 2009.

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A., and Holben, B. N.: Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network, J. Geophys. Res., 115,

doi:10.1029/2010JD014601, 2010.

Kahn, R. A., Garay, M. J., Nelson, D. L., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V, Hansen, E. G., Remer, L. A., and Tanré, D.: Response to "Toward unified satellite climatology of aerosol properties. 3. MODIS versus MISR versus AERONET", J. Quant. Spectrosc. Ra., 112, 901–909, doi:10.1016/j.jqsrt.2010.11.001, 2011.

Kittaka, C., Winker, D. M., Vaughan, M. A., Omar, A., and Remer, L. A.: Intercomparison of column aerosol optical depths from CALIPSO and MODIS-Aqua, Atmos. Meas. Tech., 4, 131–141, doi:10.5194/amt-4-131-2011, 2011.

Kleidman, R. G., O'Neill, N. T., Remer, L. A., Kaufman, Y. J., Eck, T. F., Tanré, D., Dubovik, O., and Holben, B. N.: Comparison of Moderate Resolution Imaging Spectroradiometer (MODIS) and Aerosol Robotic Network (AERONET) remote-sensing retrievals of aerosol fine mode fraction over ocean, J. Geophys. Res., 110, D22205, doi:10.1029/2005JD005760, 2005.

Kokhanovsky, A. A., Breon, F.-M., Cacciari, A., Carboni, E., Diner, D., Di Nicolantonio, W., Grainger, R. G., Grey, W. M. F., Höller, R., Lee, K.-H., Li, Z., North, P. R. J., Sayer, A. M., Thomas, G. E., and Von Hoyningen-Huene, W.: Aerosol remote sensing over land: A comparison of satellite retrievals using different algorithms and instruments, Atmos. Res., 85, 372–394, doi:10.1016/j.atmosres.2007.02.008, 2007.

Kokhanovsky, A. A., Deuzé, J. L., Diner, D. J., Dubovik, O., Ducos, F., Emde, C., Garay, M. J., Grainger, R. G., Heckel, A., Herman, M., Katsev, I. L., Keller, J., Levy, R., North, P. R. J., Prikhach, A. S., Rozanov, V. V., Sayer, A. M., Ota, Y., Tanré, D., Thomas, G. E., and Zege, E. P.: The inter-comparison of major satellite aerosol retrieval algorithms using simulated intensity and polarization characteristics of reflected light, Atmos. Meas. Tech., 3, 909–932, doi:10.5194/amt-3-909-2010, 2010.

Lee, K., Li, Z., and Kim, Y.: Atmospheric Aerosol Monitoring from Satellite Observations: A History of Three Decades, in: Proceedings of Atmospheric and Biological Environmental Monitoring, edited by: Kim, Y. J., Platt, U., Gu, M. B., and Iwahashi, H., 13–38, doi:10.1007/978-1-4020-9674-7, 2009.

Levy, R. C., Leptoukh, G. G., Kahn, R., Zubko, V., Gopalan, A., and Remer, L. A.: A Critical Look at Deriving Monthly Aerosol Optical Depth From Satellite Data, IEEE T. Geosci. Remote, 47, 29422956, doi:10.1109/TGRS.2009.2013842, 2009.

Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, Atmos. Chem. Phys., 10, 10399–10420, doi:10.5194/acp-10-10399-2010, 2010.

Li, Z., Zhao, X., Kahn, R., Mishchenko, M., Remer, L., Lee, K.-H., Wang, M., Laszlo, I., Nakajima, T., and Maring, H.: Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: a review and perspective, Ann. Geophys., 27, 2755–2770, doi:10.5194/angeo-27-2755-2009, 2009.

Martonchik, J. V., Kahn, R. A., and Diner, D. J.: Retrieval of aerosol properties over land using MISR observations, in: Proceedings of Satellite Aerosol Remote Sensing over Land, edited by: Kokhanovsky, A. A. and Leeuw, G., 267–293, doi:10.1007/978-3-540-69397-0_9, 2009.

Mishchenko, M. I., Liu, L., Geogdzhayev, I. V., Travis, L. D., Cairns, B., and Lacis, A. A.: Toward unified satellite climatology of aerosol properties., J. Quant. Spectrosc. Ra., 111, 540–552, doi:10.1016/j.jqsrt.2009.11.003, 2010.

Myhre, G., Stordal, F., Johnsrud, M., Diner, D. J., Geogdzhayev, I. V., Haywood, J. M., Holben, B. N., Holzer-Popp, T., Ignatov, A., Kahn, R. A., Kaufman, Y. J., Loeb, N., Martonchik, J. V., Mishchenko, M. I., Nalli, N. R., Remer, L. A., Schroedter-Homscheidt, M., Tanré, D., Torres, O., and Wang, M.: Intercomparison of satellite retrieved aerosol optical depth over ocean during the period September 1997 to December 2000, Atmos. Chem. Phys., 5, 1697–1719, doi:10.5194/acp-5-1697-2005, 2005.

National Institute of Standards and Technology: Engineering Statistics Handbook, 2012.

Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. R., Ferrare, R. A., Lee, K.-P., Hostetler, C. A., Kittaka, C., Rogers, R. R., Kuehn, R. E., and Liu, Z.: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, J. Atmos. Ocean. Tech., 26, 1994–2014, doi:10.1175/2009JTECHA1231.1, 2009.

O'Neill, N. T., Ignatov, A., Holben, B. N., and Eck, T. F.: The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET sunphotometer data, Geophys. Res. Lett, 27, 3333–3336, doi:10.1029/2000GL011581, 2000.

Oo, M. and Holz, R.: Improving the CALIOP aerosol optical depth using combined MODIS-CALIOP observations and CALIOP integrated attenuated total color ratio, J. Geophys. Res., 116, D14201, doi:10.1029/2010JD014894, 2011.

Petrenko, M., Ichoku, C., and Leptoukh, G.: Multi-sensor Aerosol Products Sampling System (MAPSS), Atmos. Meas. Tech., 5, 913–926, doi:10.5194/amt-5-913-2012, 2012.

Redemann, J., Vaughan, M. A., Zhang, Q., Shinozuka, Y., Russell, P. B., Livingston, J. M., Kacenelenbogen, M., and Remer, L. A.: The comparison of MODIS-Aqua (C5) and CALIOP (V2 & V3) aerosol optical depth, Atmos. Chem. Phys., 12, 3025–3043, doi:10.5194/acp-12-3025-2012, 2012.

Remer, L. A.: Validation of MODIS aerosol retrieval over ocean, Geophys. Res. Lett., 29, 8008, doi:10.1029/2001GL013204, 2002.

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote, E., and Holben, B. N.: The MODIS Aerosol Algorithm, Products, and Validation, J. Atmos. Sci., 62, 947–973, doi:10.1175/JAS3385.1, 2005.

Remer, L. A., Kleidman, R. G., Levy, R. C., Kaufman, Y. J., Tanré, D., Mattoo, S., Martins, J. V., Ichoku, C., Koren, I., Yu, H., and Holben, B. N.: Global aerosol climatology from the MODIS satellite sensors, J. Geophys. Res., 113, 1–18, doi:10.1029/2007JD009661, 2008.

Sinnott, R. W.: Virtues of the Haversine, Sky and Telescope, 68, 159, 1984.

Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., and Slutsker, I.: Cloud-Screening and Quality Control Algorithms for the AERONET Database, Remote Sens. Environ., 73, 337–349, doi:10.1016/S0034-4257(00)00109-7, 2000.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J.: Derivation of aerosol properties from satellite measurements, J. Geophys. Res., 103, 17099–17110, doi:10.1029/98JD00900, 1998.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument

observations: An overview, J. Geophys. Res., 112, 1–14, doi:10.1029/2007JD008809, 2007.

Torres, O., Jethva, H., and Bhartia, P. K.: Retrieval of Aerosol Optical Depth above Clouds from OMI Observations: Sensitivity Analysis and Case Studies, J. Atmos. Sci., 69, 1037–1053, doi:10.1175/JAS-D-11-0130.1, 2012.

Winker, D. M., Hunt, W. H., and McGill, M. J.: Initial performance assessment of CALIOP, Geophys. Res. Lett., 34, 1–5, doi:10.1029/2007GL030135, 2007.

Yu, H., Kaufman, Y. J., Chin, M., Feingold, G., Remer, L. A., Anderson, T. L., Balkanski, Y., Bellouin, N., Boucher, O., Christopher, S., DeCola, P., Kahn, R., Koch, D., Loeb, N., Reddy, M. S., Schulz, M., Takemura, T., and Zhou, M.: A review of measurement-based assessments of the aerosol direct radiative effect and forcing, Atmos. Chem. Phys., 6, 613–666, doi:10.5194/acp-6-613-2006, 2006.

Yu, H., Zhang, Y., Chin, M., Liu, Z., Omar, A., Remer, L. A., Yang, Y., Yuan, T., and Zhang, J.: An integrated analysis of aerosol above clouds from A-Train multi-sensor measurements, Remote Sens. Environ., 121, 125–131, doi:10.1016/j.rse.2012.01.011, 2012.

Zhang, J. and Reid, J. S.: A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products, Atmos. Chem. Phys., 10, 10949–10963, doi:10.5194/acp-10-10949-2010, 2010.

Zhang, J., Campbell, J. R., Reid, J. S., Westphal, D. L., Baker, N. L., Campbell, W. F., and Hyer, E. J.: Evaluating the impact of assimilating CALIOP-derived aerosol extinction profiles on a global mass transport model, Geophys. Res. Lett., 38, 1–6, doi:10.1029/2011GL047737, 2011.

Coherent uncertainty analysis of aerosol measurements from multiple satellite sensors

M. Petrenko[1,2], C. Ichoku[2]

[1]{Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland}

[2]{NASA Goddard Space Flight Center, Greenbelt, Maryland}

Correspondence to: M. Petrenko (maksym.petrenko@nasa.gov)

Abstract

Aerosol retrievals from multiple spaceborne sensors, including MODIS (on Terra and Aqua), MISR, OMI, POLDER, CALIOP, and SeaWiFS – altogether, a total of 11 different aerosol products – were comparatively analyzed using data collocated with ground-based aerosol observations from the Aerosol Robotic Network (AERONET) stations within the Multi-sensor Aerosol Products Sampling System (MAPSS, http://giovanni.gsfc.nasa.gov/mapss/ and http://giovanni.gsfc.nasa.gov/aerostat/). The analysis was performed by comparing quality-screened satellite aerosol optical depth or thickness (AOD or AOT) retrievals during 2006-2010 to available collocated AERONET measurements globally, regionally, and seasonally, and deriving a number of statistical measures of accuracy. We used a robust statistical approach to detect and remove possible outliers in the collocated data that can bias the results of the analysis. Overall, the proportion of outliers in each of the quality-screened AOD products was within 12%. Squared correlation coefficient (R^2) values of the satellite AOD retrievals relative to AERONET exceeded 0.6, with R^2 for most of the products exceeding 0.7 over land and 0.8 over ocean. Root mean square error (RMSE) values for most of the AOD products were within 0.15 over land and 0.09 over ocean. We have been able to generate global maps showing regions where the different products present advantages over the others, as well as the relative performance of each product over different landcover types. It was observed that while MODIS, MISR, and SeaWiFS provide accurate retrievals over most of the landcover types, multi-angle capabilities make MISR the only sensor to retrieve reliable AOD over barren and snow / ice surfaces. Likewise, active sensing enables CALIOP to

retrieve aerosol properties over bright-surface shrublands more accurately than the other sensors, while POLDER, which is the only one of the sensors capable of measuring polarized aerosols, outperforms other sensors in certain smoke-dominated regions, including broadleaf evergreens in Brazil and South-East Asia.

1   Introduction

Remote sensing of aerosols from space has been a subject of extensive research, with multiple sensors retrieving global aerosol properties on a daily or weekly basis. During the past decade, the retrievals of atmospheric aerosol parameters have been available from a multitude of spaceborne sensors (Lee et al., 2009; Yu et al., 2006). The diverse algorithms used for these retrievals operate on different types of the remotely-sensed signals and rely on different assumptions about the underlying physical phenomena. Significant effort has been made by the various aerosol algorithm teams to progressively refine these assumptions, from algorithm version to version, in order to derive and provide the most accurate products possible. However despite these efforts, measurements of identical aerosol parameters from different sensors, including the most common observable and widely used aerosol optical depth or thickness (AOD or AOT or τ_a) parameter, often disagree with each other due to a variety of reasons including differences in the underlying surface properties at different locations, intrinsic sensor observation characteristics and retrieval approaches (Li et al., 2009). Therefore, it has become necessary to consistently analyze the available aerosol products wherever possible in order to establish the geographical locations where and under what circumstances each of these products provide the greatest accuracy.

The unique attributes of a particular sensor may be advantageous for aerosol retrievals, depending on the parameter(s) being retrieved, especially under favorable atmospheric conditions. However, aerosol retrieval accuracy can also be affected by numerous other factors, including the retrieval algorithm's assumptions and parameterizations, the instrument characteristics (intrinsic design, calibration, and time-dependent degradation), the measurement configurations (solar and view geometry), the atmospheric conditions (cloudiness, aerosol mixing, layer height, and humidity), the surface background (vegetated, bare, snow-covered, inundated, or simply just dark or bright land surface or ocean), and others (Kokhanovsky et al., 2007).

1 Since the accuracy of aerosol retrieval from a sensor may be affected positively or negatively

2 by these factors and conditions in different ways and to varying degrees, a synergetic use of

3 similar aerosol parameters across the sensors is non-trivial and the data synergy research is

4 instead focused on combining orthogonal (i.e., non-conflicting) aerosol measurements. For

5 example, the aerosol layer height information from the Cloud-Aerosol Lidar with Orthogonal

6 Polarization (CALIOP) has been used to enhance aerosol retrievals from other sensors (Oo

7 and Holz, 2011; Torres et al., 2012; Zhang et al., 2011), while the geometry information from

8 the Advanced Along Track Scanning Radiometer (AATSR) was used to initialize the

9 Moderate Resolution Imaging Spectroradiometer (MODIS) Bi-Directional Reflection

10 Distribution Function (BRDF) in order to derive AATSR AOD (Guo et al., 2009).

11 To better characterize the differences and uncertainties that exist between the aerosol

12 retrievals from different sensors, several studies compared a limited number of sensors, e.g.,

13 AOD retrievals from MODIS were separately compared to retrievals from the MISR Multi-

14 angle Imaging Spectroradiometer (Kahn et al., 2007, 2011; Mishchenko et al., 2010; Zhang

15 and Reid, 2010), the POLDER POLarization and Directionality of the Earth's Reflectances

16 sensor (Gérard et al., 2005), and CALIOP (Kittaka et al., 2011; Redemann et al., 2011). A

17 larger set of sensors was intercompared using a synthetic benchmark (Kokhanovsky et al.,

18 2010), and also based on a detailed analysis of limited geographical regions (Cheng et al.,

19 2012; Yu et al., 2012). In addition, a set of 9 aerosol products was evaluated over ocean and

20 coastal AERONET sites during the period of 1997-2000, highlighting regions of the high

21 retrieval agreement and disagreement (Myhre et al., 2005). However, all the satellite data used

22 in that study had already undergone post-retrieval spatio-temporal aggregation at 1×1 degree

23 grid resolution on a monthly mean basis (so-called Level 3 products) before they were used in

24 the comparisons.

25 In this work, eleven retrieval-scale (Level 2) aerosol products from multiple spaceborne

26 sensors are intercompared during the recent 'golden' period of 2006-2010 (see Fig. 1), when

27 as many as seven major sensors were in operation and measuring aerosols concurrently.

28 Specifically, we focus on aerosol products retrieved over land and ocean from MODIS on

29 Terra and Aqua, MISR on Terra, the Ozone Monitoring Instrument (OMI) on Aura, POLDER

30 on PARASOL, CALIOP on CALIPSO, and the Sea-viewing Wide Field of view Sensor

31 (SeaWiFS) aboard the SeaStar spacecraft. At the time of this study (January 2013), all of the

32 studied sensors are still active, with the exception of SeaWiFS, whose operation ended in

December 2010. The analysis is based on the collocation of the satellite data products using the Multi-sensor Aerosol Products Sampling System (MAPSS) framework (Petrenko et al., 2012) that samples these satellite products relatively uniformly over the global AERosol Robotic NETwork (AERONET) of sun-photometers and other important ground-based stations both over land and ocean.

The details of the MAPSS sampling approach are explained in Sect. 2, while the relevant characteristics of the aerosol data products from the different sensors and the corresponding data quality screening techniques are described in Sect. 3 and Sect. 4. Section 5 describes a novel statistical approach for detecting and removing possible data outliers that can exist in the collocated data and, as a result, bias the statistical analysis of these data. Section 6 presents the detailed analysis of the compared aerosol products, while Sect. 7 examines the accuracy of these products based on land cover type. Conclusions are presented in Sect. 8.

2   Sampling Method

The different aerosol-measuring sensors have different spatial resolutions, of which some have square-shaped footprints while others have rectangular pixel shapes. The nominal ground pixel sizes of the analyzed aerosol products at nadir are summarized in Table 1 and these sizes become progressively larger away from nadir. To ensure a uniform and fair sampling of the aerosol products for cross-evaluation with AERONET and for comparison with one another, we used the framework of Multi-sensor Aerosol Products Sampling System (MAPSS) that was originally developed by Ichoku et al. (Ichoku et al., 2002) for validation and analysis of MODIS aerosol products (Chu et al., 2002; Ichoku et al., 2003, 2005; Levy et al., 2010; Remer, 2002) and later expanded to support aerosol products retrieved by other spaceborne sensors (Petrenko et al., 2012).   MAPSS subsets the aerosol products by extracting pixels covering approximately the same area on the ground centered over AERONET sun photometer measurement sites and over certain other point locations that are not addressed in this study.

Assuming an imaginary circle of 55-km diameter whose center coincides with each AERONET station, all spaceborne aerosol product pixels falling within the circle are extracted. An aerosol pixel is regarded as being within the circle if the coordinates of the pixel center fall within 27.5km from the coordinates of the circle center, where the distance between the coordinates of the two points is determined using the Haversine formula (Sinnott,

1 1984). Based on the nominal spatial resolution of the sensors in Table 1, the approximate

2 maximum number of pixels within the 55-km diameter sample space at nadir for the different

3 sensors is as follows: MODIS - 25, MISR - 9, OMI - 8, POLDER - 9, CALIOP - 11, and

4 SeaWiFS - 16. The actual number of pixels within the sampling circle decreases for the

5 aerosol retrievals away from the nadir of the satellite scene, and can be further reduced in the

6 presence of clouds or other factors preventing retrieval of aerosol parameters. Based on the

7 extracted sample, statistics of each aerosol parameter retrieved within the sampling areas are

8 calculated and include mean, median, standard deviation, as well as the value of the central

9 pixel over the ground station. In this paper, results are reported based on the analysis of the

10 mean values; although not reported in this paper because of the space considerations, a similar

11 analysis was performed based on the central values and is reported in the digital supplement

12 to this paper. It is appropriate to use the mean values in this paper, so as to maintain the

13 uniform sampling criterion across the different sensors to facilitate a fair intercomparison.

14 Analysis based on central pixel values can provide further details on the effect of difference in

15 sampling aerosol products from individual sensors.

16 To collocate AERONET data in time and space with the satellite data, AERONET

17 measurements acquired within ±30-min of each satellite sensor overpass are also extracted

18 and the corresponding statistics are derived. Additionally, for the convenience of aerosol data

19 intercomparison and validation, AERONET AOD are interpolated or (where necessary)

20 extrapolated to the wavelengths of spaceborne sensors in Table 1 based on the established

21 wavelength dependence of AOD (Eck et al., 1999). It is pertinent to note that this

22 interpolation (and particularly) extrapolation process might introduce an additional source of

23 uncertainty when intercomparing the aerosol products, especially for certain stations, where

24 AERONET AOD observations in the range of 440nm-1200nm have to be extrapolated by

25 52nm to match OMI AOD at 388nm.

26 Each AERONET station has a different period of operation and the quantity of available AOD

27 data points is not uniform across all stations; while many stations are still active, certain

28 stations were active in the past and only for a short period of time. The overall availability of

29 the collocated data during the analysis period of 2006-06-07 to 2010-12-11 is shown in Fig. 2,

30 where for the purposes of this study the stations are classified as land-only, ocean-only, or

31 land-and-ocean. This classification is based on analyzing collocated data of separate aerosol

32 retrievals over land and ocean from the MODIS, SeaWiFS, and POLDER sensors and

identifying stations that have AOD data points from the land datasets, ocean datasets, or both; note that the MISR, OMI, and CALIOP sensors provide only joint land-and-ocean datasets.

3   Aerosol Products

The key properties of the 11 analyzed aerosol products are summarized in Table 1, while the original science data set (SDS) names of the spaceborne aerosol products are outlined in the first column of Table 2, except for the POLDER products that do not have an established SDS product naming convention. The sampled satellite data products are derived directly from the retrieval level aerosol products (Level 2) that represent the highest available spatial resolution for each product/sensor combination and are free of aggregation artifacts that can be present in data at Level 3 (Hyer, 2011; Levy et al., 2009; Zhang and Reid, 2010).

Of the 11 sampled products, 3 are combined land-and-ocean products, 6 are land-only products, and 4 are ocean-only products. Furthermore, 6 aerosol products are retrieved from the twin MODIS-Terra and MODIS-Aqua sensors using the same set of 3 algorithms: the ocean algorithm is used for the retrievals over oceans and other large bodies of water, the land Dark Target (DT) algorithm is used over vegetated regions and other dark surfaces (Remer et al., 2005), and the land Deep Blue (DB) algorithm is used for deserts and barren lands (Hsu et al., 2004). Although the results between the two MODIS sensors are expected to be very close, they might still differ due to the different times of scene observation during the day and other factors summarized in (Ichoku et al., 2005; Remer et al., 2008).

The remainder of this section provides a brief description of the analyzed products and highlights some of the unique aerosol properties reported in these products. A more detailed overview can be found in the theoretical and validation works of the respective science teams of the products as cited below, while a general comparative overview of multiple products and retrieval algorithms are in (Kokhanovsky et al., 2007; Lee et al., 2009; Li et al., 2009; Yu et al., 2006).

AERONET (http://aeronet.gsfc.nasa.gov) sun-photometers measure aerosol properties using ground-based observations of solar direct and diffuse irradiances. In this work, the AERONET product used is the aerosol optical depth or thickness (AOD or AOT), which is retrieved from the AERONET direct measurements of solar irradiance. Since AERONET measurements are made from the ground looking up, they present a distinct advantage over

1   spaceborne retrievals in that they are not affected by uncertainties associated with the effects

2   of surface properties as much as satellite measurements are (Dubovik et al., 2002; Holben et

3   al., 1998, 2001). Furthermore, the Level 2.0 AERONET data used in this work are carefully

4   calibrated, cloud screened, and quality assured (Smirnov et al., 2000) and therefore are

5   especially suitable for use as the reference standard against which the satellite aerosol remote-

6   sensing data are evaluated.

7   The MODIS (http://modis.gsfc.nasa.gov) aerosol product (MOD04 and MYD04) comprises

8   the column aerosol optical thickness and other physical properties of aerosols retrieved

9   globally over land and ocean (Chu et al., 2002; Hsu et al., 2004; Ichoku et al., 2005; Levy et

10   al., 2010; Remer, 2002; Remer et al., 2005).

11   The MISR (http://www-misr.jpl.nasa.gov) aerosol product (MIL2ASAE) features aerosol

12   retrievals based on observations from 9 independent camera angles. Multiple viewing angles

13   allow MISR to measure certain aerosol properties that are not available from the other

14   instruments (e.g., aerosol particle size). Furthermore, MISR multiple cameras enable

15   retrievals under conditions that are unfavorable to single-view (e.g., nadir) instruments, such

16   as over bright surfaces or sun glint, where the other instruments are unable to make reliable

17   retrievals in the visible wavelengths (Kahn, 2005; Kahn et al., 2010a; Martonchik et al.,

18   2009).

19   The   OMI   (http://www.knmi.nl/omi/research/instrument/index.php)   aerosol   product

20   (OMAERUV) measures the near-UV (near ultraviolet) aerosol absorption and extinction

21   optical depth, as well as single scattering albedo, among other aerosol properties (Torres et

22   al., 1998, 2007). Moreover, OMI is capable of retrieving absorption optical depth in partially

23   cloudy conditions that usually pose a challenge to other aerosol instruments.

24   The POLDER onboard PARASOL (http://www.icare.univ-lille1.fr/parasol) aerosol land

25   product (P3L2TLGC) and aerosol ocean product (P3L2TOGC) are derived from measuring

26   spectral, directional, and polarized properties of reflected solar radiation. One of the main

27   features of the POLDER instrument is its utilization of polarization properties of the

28   measured radiation for retrieving anthropogenic aerosol optical depth (Bréon et al., 2002;

29   Buriez et al., 2002; Deuzé et al., 1999, 2001; Herman et al., 1997). It is important to note that

30   the POLDER operational algorithm retrieves AOD at 2 wavelengths (670 and 865 nm) over

31   ocean and only at 1 wavelength (865 nm) over land. Furthermore over land, the POLDER

1 algorithm retrieves only AOD that corresponds to polarized particles, i.e., mainly fine mode

2 particles originating from anthropogenic activities.

3 The CALIOP (http://www-calipso.larc.nasa.gov) aerosol product (05kmALay) represents

4 daytime and nighttime atmospheric curtain slices portraying the vertical distribution of

5 aerosols and clouds in the atmosphere, including the density and certain properties of

6 individual aerosol layers (Omar et al., 2009; Winker et al., 2007).

7 The SeaWiFS (http://disc.sci.gsfc.nasa.gov/dust/) aerosol product (SWDB) uses the Deep

8 Blue algorithm to derive aerosol optical thickness and Ångström exponent. The key features

9 of this product are the retrievals of aerosol properties over both bright desert and vegetated

10 surfaces, avoidance of sun glint that improves aerosol retrievals over ocean, and a highly

11 precise calibration of the SeaWiFS sensor (Hsu et al., 2004, 2012).

12 Since each of the foregoing data sets has a few versions because of the periodic revisions and

13 updates of their retrieval algorithms over time, the data versions that were current at the time

14 of writing this paper (January 2013) were sampled, although the study has been designed in a

15 highly flexible way to enable rapid re-analysis as the new versions become available. The

16 respective data versions used in this paper are: Terra and Aqua MODIS (Collection 051),

17 MISR (Version 002), OMI (Version 003), CALIOP (Version 3-01), POLDER (Versions L

18 and K), SeaWiFS (Version 003), and AERONET AOD (Version 2). Therefore, all of the

19 illustrations and analyses shown in this paper are based on these data versions for the

20 respective aerosol sensors.

21

22 **4   Data Quality Screening**

23 While the AERONET Level-2 data are manually inspected to be free of retrieval defects and

24 anomalies (Smirnov et al., 2000), such approach is not feasible for the voluminous spaceborne

25 data. Instead, all Level-2 aerosol products analyzed in this paper assign to AOD pixels one or

26 more quality assurance (QA) flags that indicate a degree of 'confidence' of the retrieval

27 algorithms in their results. For MODIS and SeaWiFS, aerosol QA flags are integer numbers

28 ranging from 0 to 3, with 3 representing the highest quality. For MISR and OMI data, the

29 reverse is the case (i.e., 0 is the highest quality). Finally, for POLDER and CALIOP, QA data

30 are a combination of one or more flags, most of which are real numbers ranging between 0

31 and 1, where 1 indicates the highest quality. By means of these QA flags, certain aerosol

1   retrievals are identified as 'bad quality' and are considered to be not trustworthy enough for

2   certain analyses. Therefore, users of these aerosol products have been advised to choose data

3   corresponding to a range of QA values that is most appropriate for their specific needs.

4   To establish similar yet valid QA thresholds for the analyzed products, we consulted science

5   teams of the analyzed products as well as data product validation results reported by these

6   teams and other research groups. Based on this inquiry, we chose the acceptable QA values as

7   described in Table 2. For the majority of the products, the thresholds are set based on

8   selecting a limited subset of the possible QA values. An important exception is the POLDER

9   aerosol products, where the QA flags are expressed as real numbers between 0 ('bad

10 retrieval') and 1 ('excellent retrieval'). Since there are no formal recommendations on the

11 acceptable range of the flag values, we empirically set its threshold to ≥0.7, which selects data

12 of a reasonable quality yet discards the minimal number of data pixels (see Fig. 3).

13 The original MAPSS framework was designed to facilitate data analysis experiments based on

14 different values of QA flags. For this, MAPSS extracts QA flags over the sampling area and

15 computes the statistical mode for integer QA flags and mean for real QA flags. These

16 statistical modes of the integer QA flags and means of the real QA flags provide a single

17 number for the quality assessment of each sample set, and can be used to screen the

18 corresponding subset statistics while providing a convenient alternative compared to

19 screening individual pixels, e.g., see (Levy et al., 2010; Remer et al., 2008). However, it was

20 observed that this approach has an unequal impact on the statistical properties of the different

21 aerosol products (Petrenko et al., 2012).

22 As an example, consider Fig. 4, where the global collocated subset mean AOD values from

23 OMI and Terra MODIS Deep Blue (TMODIS DB) are compared to the corresponding subset

24 mean AOD values from AERONET. It can be observed that while filtering the mean

25 TMODIS DB AOD values by the mode of QA flags improves the R^2 and RMSE statistics,

26 when compared to computing the mean values based on individually screened TMODIS DB

27 AOD pixels, this filtering significantly changes the distribution of the collocated data.

28 Specifically, compared to screening individual pixels, QA mode filtering removes 50% more

29 of the collocated data points and degrades the slope of the fitted regression line as a result of

30 removing certain high-biased points. The opposite behavior can be observed in the collocated

31 OMI AOD and AERONET AOD datasets, where screening by QA mode degrades R^2 of the

32 collocated data when compared to screening individual pixels, although RMSE is still

1 improved and the slope of the fitted regression line remains the same, since both screening
2 approaches produce approximately the same number of the OMI subset data points.

3 This observation indicates a certain inhomogeneity in the uncertainties that are present in the
4 aerosol products, as in some cases high biases in individual pixels might overwhelm the
5 statistics derived from the sample set. Therefore, to avoid such biases and ensure a fair
6 comparison between the analyzed aerosol products, the rest of this study is based on the QA
7 'pre-filtering' approach, where individual pixels in a spatial sample are screened by their QA
8 values before computing the statistics of this sample. This approach also closely models a
9 typical use of the spaceborne aerosol data, where data users screen each pixel individually and
10 do not consider QA values of its neighboring pixels. The data quantity impact of the described
11 QA screening approach can be observed in Table 3 that provides the sizes of the analyzed
12 datasets before and after the screening. It is noticeable that, depending on the product, the
13 impact is quite different, with the two MODIS ocean AOD datasets and the MISR AOD
14 dataset retaining almost all of their available datasets whereas the two MODIS DT datasets
15 retained only one-fourth of the complete collocated datasets.

16 It is important however to keep in mind that the QA values reported by the retrieval
17 algorithms are to a large degree subjective to these algorithms and do not always reflect the
18 actual quality of the retrievals. For example, in an absence of a proper aerosol or surface
19 model, an algorithm can in certain cases use a wrong model to retrieve aerosol properties and
20 mistakenly assign this retrieval a 'good' QA flag, e.g., (Kahn et al., 2010a; Levy et al., 2010).
21 Furthermore, a retrieval algorithm used might not have enough skill or even the possibility to
22 correctly recognize certain conditions that are unfavorable for aerosol retrieval, e.g., sub-pixel
23 cloud contamination in OMI retrievals (Torres et al., 1998). In yet another situation, an
24 aerosol scene can be observed in only a portion of the available observation modes of a
25 sensor, e.g., in only a few of the available observation directions in POLDER (Herman et al.,
26 1997), which can lead to more confident yet less reliable results. The opposite case can also
27 be true, where an algorithm correctly retrieves aerosol properties but is not confident about
28 the retrieval. As an example, consider Fig. 5 that explores how QA screening degrades the
29 statistics of OMI AOD and Aqua MODIS Deep Blue AOD datasets when compared to
30 AERONET AOD over Djougou, Benin, as a result of assigning a 'bad' QA flag to sufficiently
31 'good' retrievals.

32

5 Possible Data Outliers

Under rare circumstances, aerosol retrievals from spaceborne observations can produce aerosol properties that do not reflect the actual physical state of aerosol in the atmosphere. Some of the reasons for such retrievals were discussed in the previous section and might include, but are not limited to, such factors as the lack of a proper aerosol model, incorrect assumptions about boundary conditions, cloud contamination, and several other factors. In Fig. 6, the possible abnormal retrievals can be visually identified by observing points that have a minimal data density and lie abnormally far from the fitted regression lines. Even though an actual fraction of such data points in a complete collocated data set can be relatively minor, the extreme deviations of such points from the overall trend might significantly bias and misrepresent the overall statistics of the data. Therefore, when computing the overall statistics and inter-comparing the aerosol products, such data points should be treated as possible outliers and analyzed separately from the rest of the data.

In order to identify and separate the possible data outliers, we analyzed AOD residuals, i.e., the difference between spaceborne AOD and AERONET AOD observations, using the Modified Z-Score test (Iglewicz and Hoaglin, 1993; National Institute of Standards and Technology, 2012). This test is designed for testing data for multiple outliers in approximately normal data sets and works by finding data points that differ from the mean value by more than 5 median absolute deviations. Unlike the standard deviation used in the traditional Z-Score test, the median absolute deviation in the Modified Z-Score test is calculated based on the median of the data and is less sensitive to extreme values.

It is pertinent to note that even though AOD data are known to follow the lognormal distribution (O'Neill et al., 2000), the AOD residuals of the analyzed products follow an approximately normal distribution as shown in Fig. 7 and Fig. 8, with the exception of the POLDER products that mostly underestimate AOD, because their retrievals focus on anthropogenic aerosols, and thus represent only the negative portion of the distribution. In the figures, it can be seen that the distributions have long tails, strongly indicating a presence of outliers. Furthermore, it can be observed that the slopes of the fitted lines are different from the slope of the 1:1 line. This indicates that the standard deviation of the analyzed residuals is different from 1, showing that these data do not follow the standard normal distribution, although this difference does not affect the test since the Modified Z-Score test normalizes residuals by the median absolute deviation of the data.

The overall effect of removing the possible outliers can be observed in the bottom-right sub-plots of Fig. 6, Fig. 7, and Fig. 8, as highlighted by the green frames, showing that 926 (6.9%) outliers are removed from the SeaWiFS Ocean AOD dataset. The total numbers of the removed outliers are provided in Table 3 and do not exceed 12% of the total QA-screened data for any of the datasets when considering the all-season data. The global distribution of the possible data outliers is depicted in Fig. 9 and generally corresponds to the outlier locations reported by the science teams of the aerosol products, e.g., outliers around the coastal areas where the significant subpixel surface variations, shallow waters, sediments, and complex marine / inland aerosol mixtures complicate the retrievals, and also data outliers associated with uncertain retrievals by the MODIS and MISR algorithms in Amazon basin and near the Sahara desert (Kahn et al., 2010a; Levy et al., 2010), although a more detailed study is needed to determine the specific factors that lead to these outliers and their spatio-temporal distributions, in order to develop appropriate mitigation measures in the retrieval algorithms for each of the products. In the remainder of this paper, the reported results are based on the QA-screened data with the outliers removed.

6   Analysis

The overall data distribution for the analyzed spaceborne aerosol products is presented in Fig. 6, whereas the detailed linear regression fit statistics (Fox, 1997) for the products based on the treatment of the possible data outliers and the nominal delimiters of the four boreal seasons, namely, spring (March-May), summer (June-August), autumn (September-November), and winter (December to February), are listed in Table 3. The statistics are presented based on a seasonal timeframes rather than monthly or shorter time periods because there may not be sufficient coincident data for a scatter plot over such shorter time periods, due to the infrequency of satellite aerosol retrieval caused by cloud cover and other issues. Fortunately, many climatic events that are relevant to aerosol emission, transport, and distribution are often roughly aligned with these seasons.

In the presented statistics, the slope value indicates by how much the satellite retrieval for the parameter under consideration is relatively underestimated or overestimated across different magnitudes, depending on whether the slope value is less than or greater than unity. The offset parameter indicates the extent to which the satellite retrieval is biased. The squared linear correlation coefficient (R^2) indicates how consistent the parameter retrieval is across its

1   magnitude range, that is how tightly the points are aligned close to the 1-to-1 line. Finally, the

2   root mean square error (RMSE) indicates the accuracy of the retrievals measured as the

3   average error in the spaceborne retrievals as compared to the ground-based AERONET

4   retrievals.

5   In Table 3, it can be seen that all MODIS, MISR, and SeaWiFS aerosol products correlate to

6   AERONET observations with $R^2 \geq 0.6$. Furthermore, MISR, SeaWiFS Land, and MODIS Dark

7   Target products have $R^2 \geq 0.7$ and MODIS Ocean products have $R^2 \geq 0.8$. Also, once the

8   possible outliers are removed, the SeaWiFS, MISR, and MODIS Dark Target products reach

9   $R^2 \geq 0.8$, while the MODIS Deep Blue products exceed $R^2 \geq 0.7$. All the best-performing Land

10 products have RMSE values of about 0.15 (measured in the same units as AOD), with the

11 exception of MODIS Deep Blue products that have RMSE of 0.23. Removing possible

12 outliers improves (reduces) the RMSE of all products by 25%-50%, with the exception of

13 OMI for which the improvement is the smallest. This indicates an opportunity for

14 improvement of the aerosol data products by adjusting the retrieval algorithms in the areas

15 with the highest concentrations of the possible outliers.

16 Fig. 10 charts the seasonal dependence of R^2 and RMSE of the spaceborne products based on

17 the data in Table 3. While all of the products demonstrate the high seasonal variations in the

18 statistical parameters, the OMI, CALIOP, POLDER, SeaWiFS, and MODIS Deep Blue are

19 the most sensitive to the seasonal changes in the retrieval conditions, perhaps because of the

20 uncertainties associated with cloud screening (Li et al., 2009), although collocating

21 spaceborne observations with AERONET introduces certain bias towards cloud-free scenes

22 because of the comprehensive AERONET cloud screening procedures (Smirnov et al., 2000).

23 Furthermore, it can be seen that while removing the data outliers greatly reduces the RMSE

24 and removes sensitivity to the seasonal changes in CALIOP, POLDER, and SeaWiFS, the

25 sensitivity remains the same for OMI and MODIS Deep Blue indicating that the retrieval

26 errors reflected by the RMSE of these products likely stem from the regular retrievals rather

27 than the anomalous retrievals.

28 The accuracy of the spaceborne aerosol products might vary with the location of the retrieval

29 and, depending on the location, some products might be significantly more accurate than

30 others. The spatial dependence of the accuracy of the analyzed products is explored in Fig. 11

31 and Fig. 12, where it can be observed that no single sensor provides the best retrievals at all

32 sites. Furthermore, as indicated by the lighter shading in Fig. 11 (e.g., Southern Australia) and

13

1 also in the histogram of R^2 inset in this figure, some sites are not covered by high-correlation

2 (i.e., $R^2 \geq 0.75$) retrievals at all or have no collocated retrievals from the most accurate of the

3 products.

4 Furthermore, it can be observed that the best-performing aerosol products differ between Fig.

5 11 and Fig. 12 and the products providing the best RMSE are oftentimes those with the lower

6 R^2. Therefore, when choosing an aerosol product for a specific analysis goal and at a specific

7 region, it is necessary to consider a balance between a variety of seasonal, statistical, and

8 spatial factors.

9

10 7 Accuracy of aerosol data products based on land cover type

11 Aerosol properties are derived from satellite observations based on a set of assumptions about

12 the type and the optical properties of the underlying terrestrial surfaces. Therefore, it can be

13 beneficial to compare the accuracy of the considered aerosol data products based on the land

14 cover types of the sites over which the data subsets were extracted. As a reference for land

15 cover types and their spatial extent, we used the global data set that is based on the

16 International Geosphere-Biosphere Programme (IGBP) classification scheme and is available

17 from the suite of MODIS products (Friedl et al., 2002). For each land cover type, we

18 identified coincident AERONET stations and averaged their corresponding statistical results

19 from Section 6. Tables 4 and 5 list the results of this aggregation, while Fig. 13 and Fig. 14

20 outline these results on a geographical map.

21 Generally, these aggregated results corroborate the findings of Section 6 and the aerosol

22 products from MODIS and MISR sensors produce the most accurate results for the majority

23 of the land cover types, although there are some peculiarities that should be discussed in

24 greater detail in order to better understand the best areas of application of the analyzed aerosol

25 products.

26 Specifically, IGBP *water* surface locations include 36 AERONET stations out of 154 stations

27 with collocated ocean retrievals identified in Fig. 2. At these 36 locations, MODIS, MISR,

28 and SeaWiFS demonstrate the best results with $R^2 \approx 0.7$. Furthermore, POLDER Ocean data

29 set has a good RMSE=0.08 (Deuzé et al., 1999) that is comparable to the best performing

30 sensors in this region, albeit it has a relatively low squared correlation coefficient value of

31 $R^2 = 0.55$; note that these statistics are different from POLDER Ocean statistics in Fig. 6 that

1   analyzes a more complete set of AERONET stations. It is interesting to note that the

2   correlation between AERONET and Aqua MODIS AOD with R^2=0.8 is higher than the

3   correlation between AERONET and Terra MODIS AOD with R^2=0.74. A detailed inspection

4   of the data showed that this difference stems from several AERONET sites with relatively

5   small numbers of collocated data points (N<35) and the average AOD below 0.2. Under such

6   low-AOD conditions, MODIS Ocean algorithm has difficulty in retrieving the precise AOD

7   values and, as a result, is subject to an increased rate of errors (Kleidman et al., 2005; Remer,

8   2002).

9   *Evergreen broadleaf forest* regions provide conditions that are favorable for retrieving AOD

10   and multiple sensors demonstrate the high correlation with AERONET, including MODIS

11   Dark Target with R^2=0.85, MISR with R^2=0.89, SeaWiFS with R^2=0.94, and POLDER with

12   R^2=0.7. However, since these regions are also susceptible to complex smoke events (e.g., Ji

13   Parana, Brazil), sometimes combined with dust and pollution events (e.g., Anmyon, S. Korea,

14   Hong Kong, China), most of the sensors demonstrate a rather poor RMSE (Hyer et al., 2011).

15   The important exception is POLDER dataset that has RMSE=0.07, possibly because

16   POLDER is especially sensitive to small particles produced by biomass burning and

17   anthropogenic pollution sources (Fan et al., 2008), thereby retrieving fairly accurate AOD

18   values at Ji Parana and Lulin, Taiwan. It should be also noted that together with deciduous

19   broadleaf forests and savannas, evergreen broadleaf forest is one of the 3 land cover types

20   where POLDER demonstrates very good results with $R^2 \approx 0.7$, indicating the advantage of

21   polarization measurements for aerosol retrievals over these regions.

22   For *mixed forests*, MODIS Dark Target products provide the highest retrieval accuracy with

23   R^2=0.78 for Terra and 0.82 for Aqua, while MISR data is somewhat less accurate with R^2=0.7

24   as a result of underestimating high AODs during summertime biomass burning events (Kahn

25   et al., 2010b), although RMSE=0.04 of MISR is almost a factor of two better than

26   RMSE=0.08 of Terra MODIS and RMSE=0.07 of Aqua MODIS. Sufficiently reliable aerosol

27   data are also retrieved by SeaWiFS with R^2=0.69 and by POLDER with R^2=0.65.

28   For *closed shrubland*, CALIOP with R^2=0.88 MISR with R^2=0.81, and MODIS Deep Blue

29   with R^2=0.74 for Terra and R^2=0.85 for Aqua produce the best results. Although MODIS

30   Deep Blue shows a better performance than MODIS Dark Target for this land-cover type, the

31   Deep Blue products are retrieved only over a single Lake Argyle AERONET site in northern

32   Australia, whereas Dark Target products are retrieved over 7 sites and have a significantly

1 larger number of data points. Likewise, the best result demonstrated by CALIOP also

2 originates exclusively from the Lake Argyle retrievals. The difference of 0.1 in R^2 between

3 MODIS Terra Deep Blue and MODIS Aqua Deep Blue can be partly explained by the

4 difference in the data availability of these two data sets, as MODIS Terra Deep Blue at the

5 time of this work is availably only through 2007; this effect can be also observed for several

6 other land cover types, where MODIS Aqua Deep Blue tends to have a lower correlation to

7 AERONET and produces results that are closer to the results of SeaWiFS, probably because

8 the latter is also based on the Deep Blue retrieval algorithm (Hsu et al., 2006).

9 Over *wooded savannas*, both Dark Target and Deep Blue products from MODIS, and

10 SeaWiFS produce very good results with $R^2{\approx}0.85$. MISR with $R^2{=}0.63$ and OMI with

11 $R^2{=}0.66$ produce lower, but still reasonable results. The reduced performance of MISR in this

12 region can be explained by the lack of region-specific aerosol mixtures in its retrieval

13 algorithm, a situation that is expected to be improved in future revisions of the product (Kahn

14 et al., 2009). It should be also noted that this region enables the highest correlation between

15 OMI and AERONET observations, probably as a result of favorable cloud-free conditions in

16 sub-Saharan Africa (Ahn et al., 2008; Torres et al., 2007).

17 *Open shrublands* are very dry and sparsely vegetated regions that are characterized by bright

18 surfaces. Such regions present a great challenge for remote retrieval of aerosol properties

19 (Kahn et al., 2009) and none of the analyzed products exceeded the correlation coefficient of

20 0.7. Among the best-performing products, CALIOP produced the best results with $R^2{=}0.68$,

21 closely followed by MODIS Dark Target with $R^2{=}0.67$ for Terra and $R^2{=}0.62$ for Aqua,

22 MISR with $R^2{=}0.64$, and MODIS Deep Blue with $R^2{=}0.52$ for Terra and $R^2{=}0.65$ for Aqua. It

23 should be noted that open shrublands and closed shrublands are the two areas where CALIOP

24 outperforms other sensors, possibly indicating the advantage of active aerosol sensing over

25 these bright-surface regions.

26 Similar to open shrublands, *grasslands* were challenging to all of the sensors, where Terra

27 MODIS Deep Blue with $R^2{=}0.73$ and MISR with $R^2{=}0.7$ demonstrated the best results. Even

28 more challenging were *snow and ice* and also *barren or sparsely vegetated* areas, where

29 MISR was the only highly accurate aerosol product with $R^2{=}0.83$ for snow / ice and $R^2{=}0.78$

30 for barren lands, thanks to its multi-angle measurement capabilities that allow retrieving

31 aerosol properties over bright surfaces and enable the advanced cloud and ice detection

32 capabilities (Kahn et al., 2009).

16

1

8   Conclusions

3   In this paper, we analyzed and intercompared 11 spaceborne aerosol products from MODIS,
4   MISR, OMI, SeaWiFS, POLDER, and CALIOP sensors, which were sampled fairly
5   uniformly based on the MAPSS framework that was used to collocate these spaceborne
6   observations with ground-based AERONET observations during the period of 2006-06-07
7   and 2010-12-11, when all the sensors were operational. Based on this analysis, for each of the
8   AERONET stations, we identified products providing the best correlation coefficient (R^2) and
9   root mean square error (RMSE). It was found that no single product provides the best retrieval
10   over all sites, and certain sites are not covered by accurate retrievals at all. Furthermore, it was
11   observed that a product providing the best R^2 at a certain location does not always provide the
12   best RMSE at the same location. Therefore, to facilitate the multivariate analysis that is
13   necessary when choosing the most suitable spaceborne aerosol product at a specific region,
14   we plan to develop an interactive tool that would allow exploration of the multi-sensor
15   collocated data on an interactive map.

16   Further, a statistical approach based on the statistical Modified Z-Score test has been used to
17   automatically identify possible data outliers in collocated data sets. The reported analysis
18   shows that even though such atypical data points constitute a relatively minor portion (3%-
19   12%) of the analyzed data sets, they can significantly bias the results of the statistical
20   analysis. For this reason, it is suggested that such data points be set aside when analyzing
21   collocated data sets and inspected separately.

22   Finally, we assessed the accuracy of the spaceborne aerosol products based on IGBP land
23   cover classification scheme. This analysis identified sensors that retrieve the most accurate
24   aerosol properties over each of the defined land cover types and highlighted the differences
25   that exists between the sensors, providing an advantage or disadvantage in retrieving AOD
26   over the areas of a particular land cover type. Notably, some of the land cover types,
27   including open shrublands and grasslands, had only moderately accurate retrievals, indicating
28   the need for improved spaceborne aerosol remote sensing instrumentation/approaches and/or
29   retrieval algorithms.

30   **Acknowledgements**

References

Ahn, C., Torres, O., and Bhartia, P. K.: Comparison of Ozone Monitoring Instrument UV Aerosol Products with Aqua/Moderate Resolution Imaging Spectroradiometer and Multiangle Imaging Spectroradiometer observations in 2006, Journal of Geophysical Research, 113, D16S27, doi:10.1029/2007JD008832, 2008.

Bréon, F. M. M., Buriez, J. C. C., Couvert, P., Deschamps, P. Y. Y., Deuzé, J. L. L., Herman, M., Goloub, P., Leroy, M., Lifermann, A., Moulin, C., Parol, F., Sèze, G., Tanré, D., Vanbauce, C., and Vesperini, M.: Scientific results from the Polarization and Directionality of

the Earth's Reflectances (POLDER), Advances in Space Research, 30, 2383–2386, doi:10.1016/S0273-1177(02)80282-4, 2002.

Buriez, J. C., Couvert, P., Deschamps, P. Y., Deuz, J. L., Herman, M., and Goloub, P.: Scientific results from the Polarization and Directionality of the Earth's Reflectances (POLDER), , 30, 2383–2386, 2002.

Cheng, T., Chen, H., Gu, X., Yu, T., Guo, J., and Guo, H.: The inter-comparison of MODIS, MISR and GOCART aerosol products against AERONET data over China, Journal of Quantitative Spectroscopy and Radiative Transfer, doi:10.1016/j.jqsrt.2012.06.016, 2012.

Chu, D. A., Kaufman, Y. J., Ichoku, C., Remer, L. A., Tanré, D., and Holben, B. N.: Validation of MODIS aerosol optical depth retrieval over land, Geophysical Research Letters, 29, 8007, doi:10.1029/2001GL013205, 2002.

Deuzé, J. L., Bréon, F. M., Devaux, C., Goloub, P., Herman, M., Lafrance, B., Maignan, F., Marchand, A., Nadal, F., Perry, G., and Tanré, D.: Remote sensing of aerosols over land surfaces from POLDER-ADEOS-1 polarized measurements, Journal of Geophysical Research, 106, 4913–4926, doi:10.1029/2000JD900364, 2001.

Deuzé, J. L., Herman, M., Goloub, P., Tanré, D., and Marchand, A.: Characterization of aerosols over ocean from POLDER/ADEOS-1, Geophysical Research Letters, 26, 1421–1424, doi:10.1029/1999GL900168, 1999.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D., and Slutsker, I.: Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations, Journal of the Atmospheric Sciences, 59, 590–608, doi:10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2, 2002.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, Journal of Geophysical Research, 104, 31333–31349, doi:10.1029/1999JD900923, 1999.

Fan, X., Goloub, P., Deuzé, J.-L., Chen, H., Zhang, W., Tanré, D., and Li, Z.: Evaluation of PARASOL aerosol retrieval over North East Asia, Remote Sensing of Environment, 112, 697–707, doi:10.1016/j.rse.2007.06.010, 2008.

1   Fox, J.: Applied Regression Analysis, Linear Models, and Related Methods, Sage
2   Publications, Inc, 1997.

3   Friedl, M. ., McIver, D. ., Hodges, J. C. ., Zhang, X. ., Muchoney, D., Strahler, A. .,
4   Woodcock, C. ., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.:
5   Global land cover mapping from MODIS: algorithms and early results, Remote Sensing of
6   Environment, 83, 287–302, doi:10.1016/S0034-4257(02)00078-0, 2002.

7   Guo, J., Xue, Y., Cao, C., Zhang, H., Guang, J., Zhang, X., and Li, X.: A synergic algorithm
8   for retrieval of aerosol optical depth over land, Advances in Atmospheric Sciences, 26, 973–
9   983, doi:10.1007/s00376-009-7218-4, 2009.

10   Gérard, B., Deuzé, J.-L., Herman, M., Kaufman, Y. J., Lallart, P., Oudard, C., Remer, L. a.,
11   Roger, B., Six, B., and Tanré, D.: Comparisons between POLDER 2 and MODIS/Terra
12   aerosol retrievals over ocean, Journal of Geophysical Research, 110, 1–15,
13   doi:10.1029/2005JD006218, 2005.

14   Herman, M., Deuzé, J. L., Devaux, C., Goloub, P., Bréon, F. M., and Tanré, D.: Remote
15   sensing of aerosols over land surfaces including polarization measurements and application to
16   POLDER measurements, Journal of Geophysical Research, 102, 17039–17049,
17   doi:10.1029/96JD02109, 1997.

18   Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan,
19   J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.:
20   AERONET—A Federated Instrument Network and Data Archive for Aerosol
21   Characterization, Remote Sensing of Environment, 66, 1–16, doi:10.1016/S0034-
22   4257(98)00031-5, 1998.

23   Holben, B. N., Tanré, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W.
24   W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Castle, J. Vande, Setzer, A.,
25   Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C.,
26   Pinker, R. T., Voss, K., and Zibordi, G.: An emerging ground-based aerosol climatology :
27   Aerosol optical depth from AERONET, Journal of Geophysical Research, 106, 12067–12097,
28   doi:10.1029/2001JD900014, 2001.

29   Hsu, N. C., Gautam, R., Sayer, a. M., Bettenhausen, C., Li, C., Jeong, M. J., Tsay, S.-C., and
30   Holben, B. N.: Global and regional trends of aerosol optical depth over land and ocean using

SeaWiFS measurements from 1997 to 2010, Atmospheric Chemistry and Physics Discussions, 12, 8465–8501, doi:10.5194/acpd-12-8465-2012, 2012.

Hsu, N. C., Tsay, S.-C., King, M. D., and Herman, J. R.: Aerosol Properties Over Bright-Reflecting Source Regions, IEEE Transactions on Geoscience and Remote Sensing, 42, 557–569, doi:10.1109/TGRS.2004.824067, 2004.

Hyer, E.: Satellite inputs for aerosol data assimilation NAVDAS-AOD Data Assimilation, 2011.

Hyer, E. J., Reid, J. S., and Zhang, J.: An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals, Atmospheric Measurement Techniques, 4, 379–408, doi:10.5194/amt-4-379-2011, 2011.

Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., Slutsker, I., and Holben, B. N.: A spatio-temporal approach for global validation and analysis of MODIS aerosol products, Geophysical Research Letters, 29, 1–4, doi:10.1029/2001GL013206, 2002.

Ichoku, C., Remer, L. A., and Eck, T. F.: Quantitative evaluation and intercomparison of morning and afternoon Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol measurements from Terra and Aqua, Journal of Geophysical Research, 110, doi:10.1029/2004JD004987, 2005.

Ichoku, C., Remer, L. A., Kaufman, Y. J., Levy, R., Chu, D. A., Tanré, D., and Holben, B. N.: MODIS observation of aerosols and estimation of aerosol radiative forcing over southern Africa during SAFARI 2000, Journal of Geophysical Research, 108, 8499, doi:10.1029/2002JD002366, 2003.

Iglewicz, B. and Hoaglin, D. C.: How to Detect and Handle Outliers, American Society for Quality Control, Milwaukee, WI, 1993.

Kahn, R. A.: Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations, Journal of Geophysical Research, 110, 1–16, doi:10.1029/2004JD004706, 2005.

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A., and Holben, B. N.: Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison

with the Aerosol Robotic Network, Journal of Geophysical Research, 115, doi:10.1029/2010JD014601, 2010a.

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

Kahn, R. A., Garay, M. J., Nelson, D. L., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V, Hansen, E. G., Remer, L. A., and Tanré, D.: Response to "Toward unified satellite climatology of aerosol properties. 3. MODIS versus MISR versus AERONET", Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 901–909, doi:10.1016/j.jqsrt.2010.11.001, 2011.

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

Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V, Paradise, S. R., Hansen, E. G., and Remer, L. A.: MISR Aerosol Product Attributes and Statistical Comparisons With MODIS, IEEE Transactions on Geoscience and Remote Sensing, 47, 4095–4114, doi:10.1109/TGRS.2009.2023115, 2009.

Kittaka, C., Winker, D. M., Vaughan, M. a., Omar, A., and Remer, L. a.: Intercomparison of column aerosol optical depths from CALIPSO and MODIS-Aqua, Atmospheric Measurement Techniques, 4, 131–141, doi:10.5194/amt-4-131-2011, 2011.

Kleidman, R. G., O'Neill, N. T., Remer, L. a., Kaufman, Y. J., Eck, T. F., Tanré, D., Dubovik, O., and Holben, B. N.: 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, doi:10.1029/2005JD005760, 2005.

Kokhanovsky, A. A., Breon, F.-M., Cacciari, A., Carboni, E., Diner, D., Di Nicolantonio, W., Grainger, R. G., Grey, W. M. F., Höller, R., Lee, K.-H., Li, Z., North, P. R. J., Sayer, A. M., Thomas, G. E., and Von Hoyningen-Huene, W.: Aerosol remote sensing over land: A

comparison of satellite retrievals using different algorithms and instruments, Atmospheric Research, 85, 372–394, doi:10.1016/j.atmosres.2007.02.008, 2007.

Kokhanovsky, A. A., Deuzé, J. L., Diner, D. J., Dubovik, O., Ducos, F., Emde, C., Garay, M. J., Grainger, R. G., Heckel, A., Herman, M., Katsev, I. L., Keller, J., Levy, R., North, P. R. J., Prikhach, A. S., Rozanov, V. V., Sayer, A. M., Ota, Y., Tanré, D., Thomas, G. E., and Zege, E. P.: The inter-comparison of major satellite aerosol retrieval algorithms using simulated intensity and polarization characteristics of reflected light, Atmospheric Measurement Techniques, 3, 909–932, doi:10.5194/amt-3-909-2010, 2010.

Lee, K., Li, Z., and Kim, Y.: Atmospheric Aerosol Monitoring from Satellite Observations : A History of Three Decades, in proceedings of Atmospheric and Biological Environmental Monitoring (Kim, Y.J., Platt, U., Gu, M.B., Iwahashi, H., Eds.), pp. 13–38, doi:10.1007/978-1-4020-9674-7, 2009.

Levy, R. C., Leptoukh, G. G., Kahn, R., Zubko, V., Gopalan, A., and Remer, L. A.: A Critical Look at Deriving Monthly Aerosol Optical Depth From Satellite Data, doi:10.1109/TGRS.2009.2013842, 2009.

Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, Atmospheric Chemistry and Physics, 10, 10399–10420, doi:10.5194/acp-10-10399-2010, 2010.

Li, Z., Zhao, X., Kahn, R., Mishchenko, M., Remer, L., Lee, K.-H., Wang, M., Laszlo, I., Nakajima, T., and Maring, H.: Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: a review and perspective, Annales Geophysicae, 27, 2755–2770, doi:10.5194/angeo-27-2755-2009, 2009.

Martonchik, J. V, Kahn, R. A., and Diner, D. J.: Retrieval of aerosol properties over land using MISR observations, in proceedings of Satellite Aerosol Remote Sensing over Land, edited by A. A. Kokhanovsky and G. Leeuw, pp. 267–293, doi:10.1007/978-3-540-69397-0_9, 2009.

Mishchenko, M. I., Liu, L., Geogdzhayev, I. V., Travis, L. D., Cairns, B., and Lacis, A. A.: Toward unified satellite climatology of aerosol properties., Journal of Quantitative Spectroscopy and Radiative Transfer, 111, 540–552, doi:10.1016/j.jqsrt.2009.11.003, 2010.

Myhre, G., Stordal, F., Johnsrud, M., Diner, D. J., Geogdzhayev, I. V., Haywood, J. M., Holben, B. N., Holzer-Popp, T., Ignatov, A., Kahn, R. A., Kaufman, Y. J., Loeb, N., Martonchik, J. V., Mishchenko, M. I., Nalli, N. R., Remer, L. a., Schroedter-Homscheidt, M., Tanré, D., Torres, O., and Wang, M.: Intercomparison of satellite retrieved aerosol optical depth over ocean during the period September 1997 to December 2000, Atmospheric Chemistry and Physics, 5, 1697–1719, doi:10.5194/acp-5-1697-2005, 2005.

National Institute of Standards and Technology: Engineering Statistics Handbook, 2012.

Omar, A. H., Winker, D. M., Vaughan, M. a., Hu, Y., Trepte, C. R., Ferrare, R. a., Lee, K.-P., Hostetler, C. a., Kittaka, C., Rogers, R. R., Kuehn, R. E., and Liu, Z.: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, Journal of Atmospheric and Oceanic Technology, 26, 1994–2014, doi:10.1175/2009JTECHA1231.1, 2009.

Oo, M. and Holz, R.: Improving the CALIOP aerosol optical depth using combined MODIS-CALIOP observations and CALIOP integrated attenuated total color ratio, Journal of Geophysical Research, 116, D14201, doi:10.1029/2010JD014894, 2011.

O'Neill, N. T., Ignatov, A., Holben, B. N., and Eck, T. F.: The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET sunphotometer data, Geophys. Res. Lett, 27, 3333–3336, doi:10.1029/2000GL011581, 2000.

Petrenko, M., Ichoku, C., and Leptoukh, G.: Multi-sensor Aerosol Products Sampling System (MAPSS), Atmospheric Measurement Techniques, 5, 913–926, doi:10.5194/amt-5-913-2012, 2012.

Redemann, J., Vaughan, M. a., Zhang, Q., Shinozuka, Y., Russell, P. B., Livingston, J. M., Kacenelenbogen, M., and Remer, L. a.: The comparison of MODIS-Aqua (C5) and CALIOP (V2 & V3) aerosol optical depth, Atmospheric Chemistry and Physics Discussions, 11, 22987–23027, doi:10.5194/acpd-11-22987-2011, 2011.

Remer, L. A.: Validation of MODIS aerosol retrieval over ocean, Geophysical Research Letters, 29, 8008, doi:10.1029/2001GL013204, 2002.

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. a., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote, E., and Holben, B. N.: The

MODIS Aerosol Algorithm, Products, and Validation, Journal of the Atmospheric Sciences, 62, 947–973, doi:10.1175/JAS3385.1, 2005.

Remer, L. A., Kleidman, R. G., Levy, R. C., Kaufman, Y. J., Tanré, D., Mattoo, S., Martins, J. V., Ichoku, C., Koren, I., Yu, H., and Holben, B. N.: Global aerosol climatology from the MODIS satellite sensors, Journal of Geophysical Research, 113, 1–18, doi:10.1029/2007JD009661, 2008.

Sinnott, R. W.: Virtues of the Haversine, Sky and Telescope, 68, 159, 1984.

Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., and Slutsker, I.: Cloud-Screening and Quality Control Algorithms for the AERONET Database, Remote Sensing of Environment, 73, 337–349, doi:10.1016/S0034-4257(00)00109-7, 2000.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J.: Derivation of aerosol properties from satellite measurements, Journal of Geophysical Research, 103, 17,099–17,110, doi:10.1029/98JD00900, 1998.

Torres, O., Jethva, H., and Bhartia, P. K.: Retrieval of Aerosol Optical Depth above Clouds from OMI Observations: Sensitivity Analysis and Case Studies, Journal of the Atmospheric Sciences, 69, 1037–1053, doi:10.1175/JAS-D-11-0130.1, 2012.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview, Journal of Geophysical Research, 112, 1–14, doi:10.1029/2007JD008809, 2007.

Winker, D. M., Hunt, W. H., and McGill, M. J.: Initial performance assessment of CALIOP, Geophysical Research Letters, 34, 1–5, doi:10.1029/2007GL030135, 2007.

Yu, H., Kaufman, Y. J., Chin, M., Feingold, G., Remer, L. a., Anderson, T. L., Balkanski, Y., Bellouin, N., Boucher, O., Christopher, S., DeCola, P., Kahn, R., Koch, D., Loeb, N., Reddy, M. S., Schulz, M., Takemura, T., and Zhou, M.: A review of measurement-based assessments of the aerosol direct radiative effect and forcing, Atmospheric Chemistry and Physics, 6, 613–666, doi:10.5194/acp-6-613-2006, 2006.

Yu, H., Zhang, Y., Chin, M., Liu, Z., Omar, A., Remer, L. A., Yang, Y., Yuan, T., and Zhang, J.: An integrated analysis of aerosol above clouds from A-Train multi-sensor measurements, Remote Sensing of Environment, 121, 125–131, doi:10.1016/j.rse.2012.01.011, 2012.

1 Zhang, J., Campbell, J. R., Reid, J. S., Westphal, D. L., Baker, N. L., Campbell, W. F., and
2 Hyer, E. J.: Evaluating the impact of assimilating CALIOP-derived aerosol extinction profiles
3 on a global mass transport model, Geophysical Research Letters, 38, 1–6,
4 doi:10.1029/2011GL047737, 2011.

5 Zhang, J. and Reid, J. S.: A decadal regional and global trend analysis of the aerosol optical
6 depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products,
7 Atmospheric Chemistry and Physics, 10, 18879–18917, doi:10.5194/acp-10-10949-2010,
8 2010.

9

Table 1. Ground-based and spaceborne atmospheric aerosol products analyzed in the study. In the product designation titles, 'O' at the end of the title of a product signifies ocean retrievals, 'L' – land retrievals, 'DT' – land retrievals using the MODIS Dark Target algorithm, and 'DB' – land retrievals using the MODIS Deep Blue algorithm. The AERONET AOD retrievals were interpolated or extrapolated to the studied wavelengths of the spaceborne sensors. The indicated local equatorial crossing times are based on the original orbital designs, and can change during the lifetimes of the satellites. SeaWiFS mission has ended in December 2010.

Sensor	Platform	Product	Designation in the study	Study Wavelength	Spatial Resolution (km x km)	Equator crossing time	Launch date
AERONET	N/A	AOT	AERONET	Varies	N/A	N/A	Varies
MODIS	Terra	MOD04	TMODIS DT	550nm	10x10	10:30 am	Jan'00
			TMODIS DB				
			TMODIS O				
	Aqua	MYD04	AMODIS DT			1:30 pm	Jul'02
			AMODIS DB				
			AMODIS O				
MISR	Terra	MIL2ASAE	MISR	558nm	17.6x17.6	10:30 am	Jan'00
OMI	Aura	OMAERUV	OMI	388nm	13.7x23.7	1:38 pm	Oct'04
POLDER	PARASOL	P3L2TLGC	POLDER3 L	865nm	19x19	1:30 pm	Mar'05
		P3L2TOGC	POLDER3 O	670nm			
CALIOP	CALIPSO	05kmALay	CALIOP	532nm	5x0.1	1:32 pm	Jun'06
SeaWiFS	SeaStar	SWDB	SeaWiFS L	550nm	13.5x13.5	12:00 pm	Jan'98
			SeaWiFS O				

1   Table 2. Studied aerosol data sets, the matching data quality (QA) data sets, and the

2   corresponding QA data screening criteria. Where provided, numbers in parenthesis in the

3   middle column indicate the base-1 layer index, base-0 bit number, and number of bits

4   extracted from this QA data set. For MODIS, MISR, OMI, and SeaWiFS the QA values are

5   integer numbers between 0 and 3, where for MODIS and SeaWiFS larger numbers indicate a

6   better retrieval quality and for OMI and MISR the opposite is true. For POLDER, QA is real

7   number between 0 (worst) and 1 (best). For CALIOP, the QA condition is applied to all layers

8   found in a column; the whole column is rejected if at least one layer fails the test. The listed

9   extinction QC values indicate retrievals that are unconstrained, constrained, have a reduced

10   lidar ratio, or detected an opaque aerosol layer. CAD score and layer type and subtype flags

11   indicate retrievals that classified a layer with a high confidence as containing aerosol and

12   were able to determine the aerosol type. IAB condition is set to prevent the retrieval anomaly

13   of overcorrecting the attenuation of overlaying layers (Kittaka et al., 2011).

AOD data set	QA data set (layer / starting bit /number of bits)	QA condition
MODIS		
Corrected_Optical_Depth_Land	Quality_Assurance_Land (1/1/3)	QA=3
Deep_Blue_Aerosol_Optical _Depth_550_Land	Quality_Assurance_Land (5/1/2)	QA=3
Effective_Optical_Depth _Average_Ocean	Quality_Assurance_Ocean (1/5/3)	QA=[1,2,3]
MISR		
RegBestEstimateSpectralOptDepth	RegBestEstimateQA	QA=[0,1]
OMI		
FinalAerosolOpticalDepth	FinalAlgorithmFlags	QA=0
SeaWiFS		
aerosol_optical_thickness_550_land	aerosol_optical_thickness _confidence_flag_land	QA=3
aerosol_optical_thickness_550_ocean	aerosol_optical_thickness _confidence_flag_ocean	QA=[2,3]
CALIOP		
Column_Optical_Depth_Aerosols_532	ExtinctionQC_[532 and 1064]	QA=[0,1,2,16,18]
	CAD_Score	-100≤CAD<-20
	Feature_Classification_Flags (1/0/3)	Layer type=3
	Feature_Classification_Flags (1/9/3)	Layer subtype>0
	Integrated_Attenuated _Backscatter_[532 and 1064]	IAB<=0.01
POLDER		
Aerosol optical thickness at 865nm corresponding to the polarized particles	Quality index for the inversion	QA≥0.7
Aerosol Optical Thickness at 670 nm	Quality index for the inversion	QA≥0.7

1 Table 3. Statistics of the studied aerosol data sets based on all AERONET stations during the

2 period of 2006-06-07 and 2010-12-11. 'Ntot' indicates the total number of the collocated

3 Spaceborne AOD - AERONET AOD data points, while 'Nfilt' indicates the number of data

4 points after filtering (screening) the spaceborne data by QA as described in Section 4 and

5 Table 2. 'Nout' is the total number of the possible data outliers determined as explained in

6 Section 5. The last 8 columns present the statistics on the collocated data based on regression

7 fits also plotted in Fig. 6.

Dataset	Nfilt	Nfilt/Ntot (%)	Nout	Nfilt/Nout (%)	Complete data				Outliers removed			
					R^2	RMSE	Slope	Intercept	R^2	RMSE	Slope	Intercept
All seasons												
TMODIS DT	56803	72.9	1711	3.0	0.79	0.11	0.96	0.01	0.83	0.08	0.95	0.00
TMODIS DB	4431	32.2	274	6.2	0.63	0.23	0.87	0.06	0.73	0.14	0.88	0.04
TMODIS O	17243	99.7	703	4.1	0.82	0.08	1.02	0.03	0.88	0.05	0.96	0.03
AMODIS DT	48555	68.6	1880	3.9	0.78	0.11	0.97	0.01	0.83	0.08	0.97	0.01
AMODIS DB	13544	33.9	973	7.2	0.63	0.23	0.88	0.06	0.77	0.14	0.9	0.03
AMODIS O	17790	99.6	844	4.7	0.8	0.08	0.94	0.03	0.87	0.05	0.93	0.03
MISR	16561	99.9	959	5.8	0.73	0.13	0.61	0.07	0.84	0.06	0.85	0.03
OMI	52498	95.6	2009	3.8	0.4	0.25	0.72	0.17	0.51	0.19	0.76	0.14
CALIOP	1885	92.4	158	8.4	0.34	0.23	0.49	0.07	0.66	0.11	0.78	0.00
POLDER3 L	31874	86.4	3837	12	0.27	0.21	0.17	0.02	0.57	0.08	0.5	-0.01
POLDER3 O	5400	59.8	518	9.6	0.39	0.16	0.25	0.03	0.6	0.07	0.57	0.00
SeaWiFS L	18305	43.3	1038	5.7	0.75	0.13	0.74	0.04	0.82	0.08	0.88	0.01
SeaWiFS O	13333	82.6	926	6.9	0.65	0.12	0.99	0.03	0.81	0.06	0.92	0.02
Fall												
TMODIS DT	15586	72.2	466	3.0	0.83	0.11	1.02	-0.01	0.86	0.07	0.94	0.00
TMODIS DB	1422	29.6	60	4.2	0.48	0.19	0.91	0.03	0.58	0.13	0.86	0.03
TMODIS O	5120	99.6	184	3.6	0.79	0.07	0.91	0.03	0.86	0.05	0.95	0.03
AMODIS DT	12659	65.9	494	3.9	0.82	0.1	1.03	-0.01	0.86	0.07	0.96	0.00
AMODIS DB	3927	32.4	258	6.6	0.59	0.2	0.97	0.02	0.74	0.12	0.9	0.01
AMODIS O	5135	99.6	271	5.3	0.74	0.07	0.79	0.04	0.86	0.04	0.9	0.02
MISR	4466	99.8	286	6.4	0.7	0.14	0.54	0.07	0.84	0.05	0.83	0.03
OMI	13909	94.1	661	4.8	0.33	0.22	0.65	0.15	0.48	0.16	0.72	0.11
CALIOP	529	92.6	44	8.3	0.32	0.24	0.51	0.08	0.66	0.10	0.79	0.00
POLDER3 L	9926	86.3	1050	10.6	0.43	0.14	0.33	0.00	0.67	0.07	0.62	-0.02
POLDER3 O	1393	56.8	99	7.1	0.48	0.12	0.31	0.02	0.6	0.07	0.57	-0.01
SeaWiFS L	5795	47.5	306	5.3	0.81	0.1	0.79	0.02	0.85	0.07	0.89	0.01
SeaWiFS O	3665	82.4	277	7.6	0.61	0.12	0.94	0.03	0.81	0.06	0.89	0.02

8

9

Table 3 (continued)

Dataset	Nfilt	Nfilt/Ntot (%)	Nout	Nfilt/Nout (%)	Complete data				Outliers removed			
					R^2	RMSE	Slope	Intercept	R^2	RMSE	Slope	Intercept
Winter												
TMODIS DT	6229	64.4	253	4.1	0.79	0.11	0.76	0.03	0.84	0.08	0.86	0.02
TMODIS DB	826	37.0	82	9.9	0.65	0.23	0.9	0.08	0.81	0.13	0.9	0.05
TMODIS O	3297	99.8	140	4.2	0.72	0.09	0.98	0.03	0.87	0.05	0.93	0.03
AMODIS DT	4945	58.6	238	4.8	0.78	0.13	0.77	0.03	0.83	0.09	0.88	0.02
AMODIS DB	3052	39.9	278	9.1	0.62	0.21	0.84	0.06	0.81	0.12	0.91	0.02
AMODIS O	3319	99.5	183	5.5	0.74	0.08	0.84	0.03	0.84	0.04	0.92	0.02
MISR	2701	99.9	179	6.6	0.75	0.11	0.54	0.06	0.82	0.05	0.77	0.04
OMI	8569	93.9	279	3.3	0.44	0.24	0.63	0.17	0.53	0.19	0.69	0.14
CALIOP	313	91.3	29	9.3	0.43	0.25	0.37	0.09	0.71	0.10	0.72	0.02
POLDER3 L	4954	87.1	488	9.9	0.35	0.21	0.15	0.02	0.56	0.08	0.4	0.00
POLDER3 O	792	57.4	45	5.7	0.54	0.11	0.39	0.01	0.59	0.08	0.54	-0.01
SeaWiFS L	3063	46.1	163	5.3	0.69	0.15	0.65	0.05	0.76	0.1	0.8	0.03
SeaWiFS O	1917	81.6	137	7.1	0.57	0.12	0.89	0.05	0.75	0.07	0.84	0.03
Spring												
TMODIS DT	12990	74.0	505	3.9	0.79	0.11	0.92	0.03	0.85	0.08	0.96	0.01
TMODIS DB	746	34.4	33	4.4	0.65	0.3	0.78	0.12	0.69	0.21	0.84	0.07
TMODIS O	3464	99.7	130	3.8	0.85	0.09	1.05	0.03	0.88	0.06	0.97	0.04
AMODIS DT	11266	70.7	511	4.5	0.78	0.12	0.96	0.03	0.84	0.08	0.97	0.02
AMODIS DB	3079	35.6	163	5.3	0.63	0.27	0.83	0.09	0.73	0.19	0.86	0.06
AMODIS O	3625	99.8	160	4.4	0.81	0.09	0.99	0.03	0.87	0.06	0.92	0.04
MISR	4016	99.9	260	6.5	0.76	0.13	0.64	0.08	0.86	0.06	0.86	0.04
OMI	12170	96.5	446	3.7	0.45	0.29	0.77	0.18	0.55	0.22	0.79	0.15
CALIOP	439	93.2	41	9.3	0.37	0.25	0.56	0.05	0.71	0.12	0.81	-0.02
POLDER3 L	7809	87.9	1170	15	0.25	0.27	0.11	0.02	0.52	0.10	0.4	-0.01
POLDER3 O	1443	61.7	162	11.2	0.38	0.18	0.24	0.04	0.56	0.09	0.52	0.00
SeaWiFS L	4345	45.4	306	7.0	0.78	0.15	0.76	0.05	0.87	0.09	0.91	0.02
SeaWiFS O	3038	82.2	221	7.3	0.68	0.15	1.05	0.02	0.83	0.07	0.95	0.02
Summer												
TMODIS DT	21998	75.6	519	2.4	0.76	0.1	1.01	0.00	0.78	0.08	0.98	0.00
TMODIS DB	1437	31.5	91	6.3	0.65	0.24	0.91	0.04	0.77	0.14	0.90	0.02
TMODIS O	5362	99.6	247	4.6	0.84	0.09	1.06	0.02	0.89	0.06	0.98	0.03
AMODIS DT	19685	72.3	551	2.8	0.76	0.1	1.05	0.01	0.79	0.08	1.00	0.01
AMODIS DB	3486	30.1	228	6.5	0.62	0.24	0.90	0.06	0.74	0.15	0.91	0.04
AMODIS O	5711	99.5	276	4.8	0.84	0.08	1.00	0.02	0.88	0.05	0.95	0.02
MISR	5378	99.9	222	4.1	0.73	0.13	0.64	0.07	0.83	0.06	0.86	0.03
OMI	17850	97.1	542	3.0	0.35	0.25	0.73	0.18	0.45	0.2	0.76	0.15
CALIOP	604	92.1	40	6.6	0.33	0.18	0.52	0.05	0.6	0.11	0.76	0.00
POLDER3 L	9185	84.9	1125	12.2	0.22	0.19	0.15	0.02	0.53	0.07	0.52	-0.01
POLDER3 O	1772	62.0	192	10.8	0.32	0.17	0.20	0.04	0.6	0.07	0.59	0.00
SeaWiFS L	5102	36.9	243	4.8	0.67	0.13	0.70	0.04	0.75	0.08	0.85	0.01
SeaWiFS O	4713	83.4	286	6.1	0.68	0.11	1.00	0.03	0.82	0.06	0.93	0.02

Table 4. Linear fit correlation coefficient (R2) between the collocated spaceborne and ground-based observations of AOD estimated at the stations that coincide with different IGBP land cover types. Empty cells indicate no collocated data available from a specific sensor over a specific land cover type. No AERONET stations are available at the areas occupied by Deciduous needleleaf forest. The statistics were calculated based on the data that was pre-filtered by QA and screened of outliers as described in Sections 4 and Section 5. A graphical representation of this table is in Figure 13.

	TMODIS DT	TMODIS DB	TMODIS O	AMODIS DT	AMODIS DB	AMODIS O	MISR	OMI	CALIOP	POLDER3 L	POLDER3 O	SeaWiFS L	SeaWiFS O
Water			0.74			0.80	0.72	0.39	0.59		0.55		0.68
Evergreen needleleaf forest	0.78			0.79	0.71		0.74	0.33	0.59	0.55		0.66	
Evergreen broadleaf forest	0.85	1.00		0.85			0.89	0.53	0.17	0.70		0.94	
Deciduous broadleaf forest	0.84			0.87			0.84	0.55	0.27	0.70		0.84	
Mixed forests	0.78			0.82			0.70	0.42	0.61	0.65		0.69	
Closed shrubland	0.51	0.74		0.63	0.85		0.81	0.50	0.88	0.33		0.65	
Open shrublands	0.67	0.52		0.62	0.65		0.64	0.31	0.68	0.32		0.54	
Woody savannas	0.83	0.91		0.86	0.86		0.63	0.66	0.34	0.54		0.79	
Savannas	0.73	0.69		0.80	0.67		0.75	0.53	0.59	0.67		0.78	
Grasslands	0.56	0.73		0.67	0.43		0.70	0.48	0.40	0.44		0.55	
Permanent wetlands	0.74			0.76			0.77	0.31	0.54	0.39		0.02	
Croplands	0.78	0.72		0.78	0.62		0.80	0.47	0.56	0.49		0.68	
Urban and built-up	0.70	0.64		0.70	0.59		0.76	0.43	0.51	0.44		0.62	
Cropland / natural veget. mosaic	0.77			0.79	0.49		0.83	0.54	0.46	0.56		0.72	
Snow and ice	0.26			0.27			0.83		0.03	0.53			
Barren or sparsely vegetated	0.60	0.57		0.62	0.34		0.78	0.29	0.58	0.33		0.36	

Table 5. Root mean square error (RMSE) between the collocated spaceborne and ground-based observations of AOD estimated at the stations that coincide with different IGBP land cover types. Empty cells indicate no collocated data available from a specific sensor over a specific land cover type. No AERONET stations are available at the areas occupied by Deciduous needleleaf forest. The statistics were calculated based on the data that was pre-filtered by QA and screened of outliers as described in Sections 4 and Section 5. A graphical representation of this table is in Figure 14.

	TMODIS DT	TMODIS DB	TMODIS O	AMODIS DT	AMODIS DB	AMODIS O	MISR	OMI	CALIOP	POLDER3 L	POLDER3 O	SeaWiFS L	SeaWiFS O
Water			0.06			0.05	0.07	0.17	0.12		0.08		0.09
Evergreen needleleaf forest	0.07	0.09		0.06	0.30		0.06	0.18	0.08	0.06		0.05	
Evergreen broadleaf forest	0.10	0.80		0.09	0.30		0.11	0.25	0.35	0.07		0.23	
Deciduous broadleaf forest	0.06			0.06			0.04	0.12	0.11	0.05		0.05	
Mixed forests	0.08			0.07			0.04	0.14	0.12	0.06		0.05	
Closed shrubland	0.09	0.05		0.08	0.05		0.06	0.18	0.04	0.09		0.06	
Open shrublands	0.10	0.10		0.09	0.14		0.08	0.23	0.11	0.13		0.08	
Woody savannas	0.08	0.27		0.09	0.23		0.09	0.24	0.18	0.15		0.08	
Savannas	0.11	0.17		0.10	0.17		0.08	0.27	0.12	0.12		0.14	
Grasslands	0.10	0.18		0.09	0.15		0.05	0.20	0.12	0.10		0.08	
Permanent wetlands	0.07			0.06			0.06	0.17	0.11	0.09		0.13	
Croplands	0.08	0.16		0.09	0.22		0.08	0.19	0.16	0.11		0.09	
Urban and built-up	0.09	0.13		0.09	0.16		0.07	0.21	0.13	0.10		0.10	
Cropland / natural veget. mosaic	0.07			0.08	0.19		0.10	0.14	0.20	0.09		0.11	
Snow and ice	0.11			0.14			0.02		0.10	0.02			
Barren or sparsely vegetated	0.11	0.16		0.12	0.11		0.06	0.39	0.11	0.12		0.11	

32
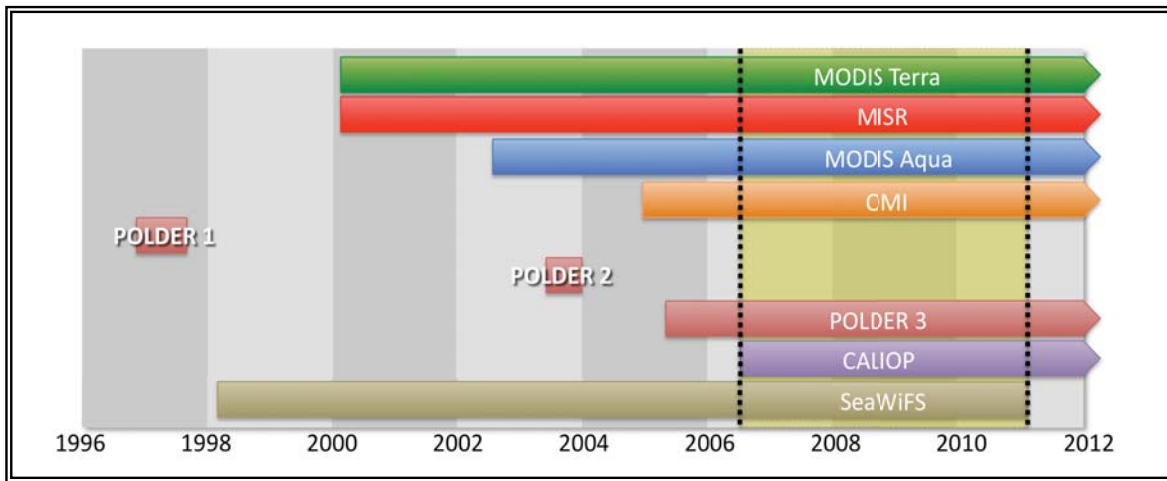
Figure 1. Periods of operation of major past and current aerosol-measuring satellite sensors. The pair of dotted vertical lines marks the 'golden' period (between the start of CALIOP in July 2006 and the end of SeaWiFS in December 2010) when as many as seven of these sensors were measuring aerosols concurrently. The golden period was used as the base for the studies reported in the rest of this paper.
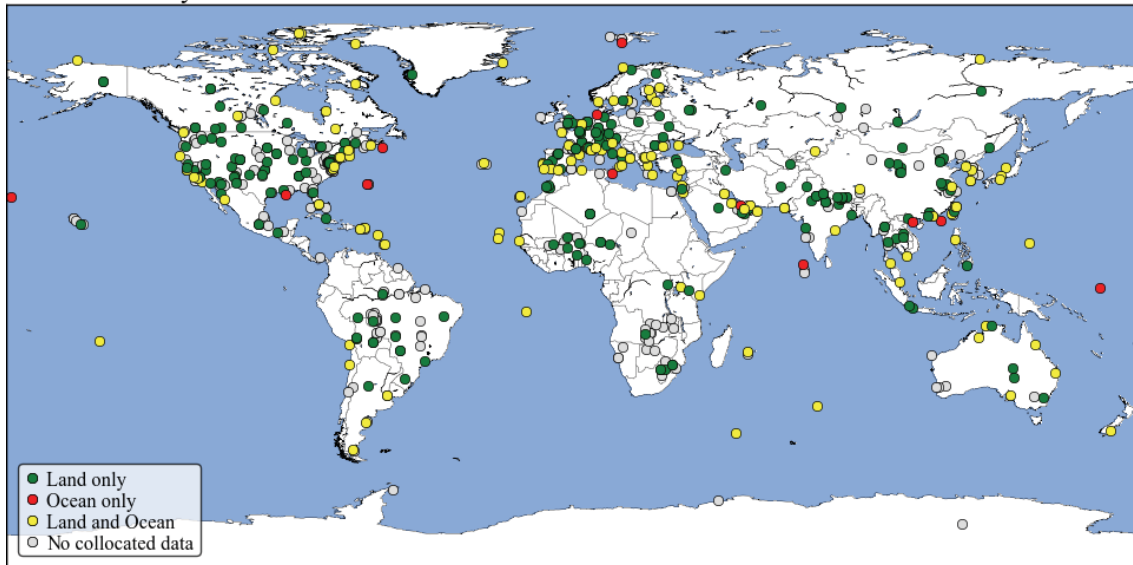
Figure 2. Distribution of AERONET stations used in the study. Green, red, and yellow colors indicate stations that can be classified as land only (226 sites), ocean only (12 sites), or both land and ocean (142 sites), respectively. The classification was established based on data availability in separate over-land and over-ocean datasets in MODIS, SeaWiFS, and POLDER aerosol products. Gray color indicates stations that do not have any collocated data for the studied period of time.
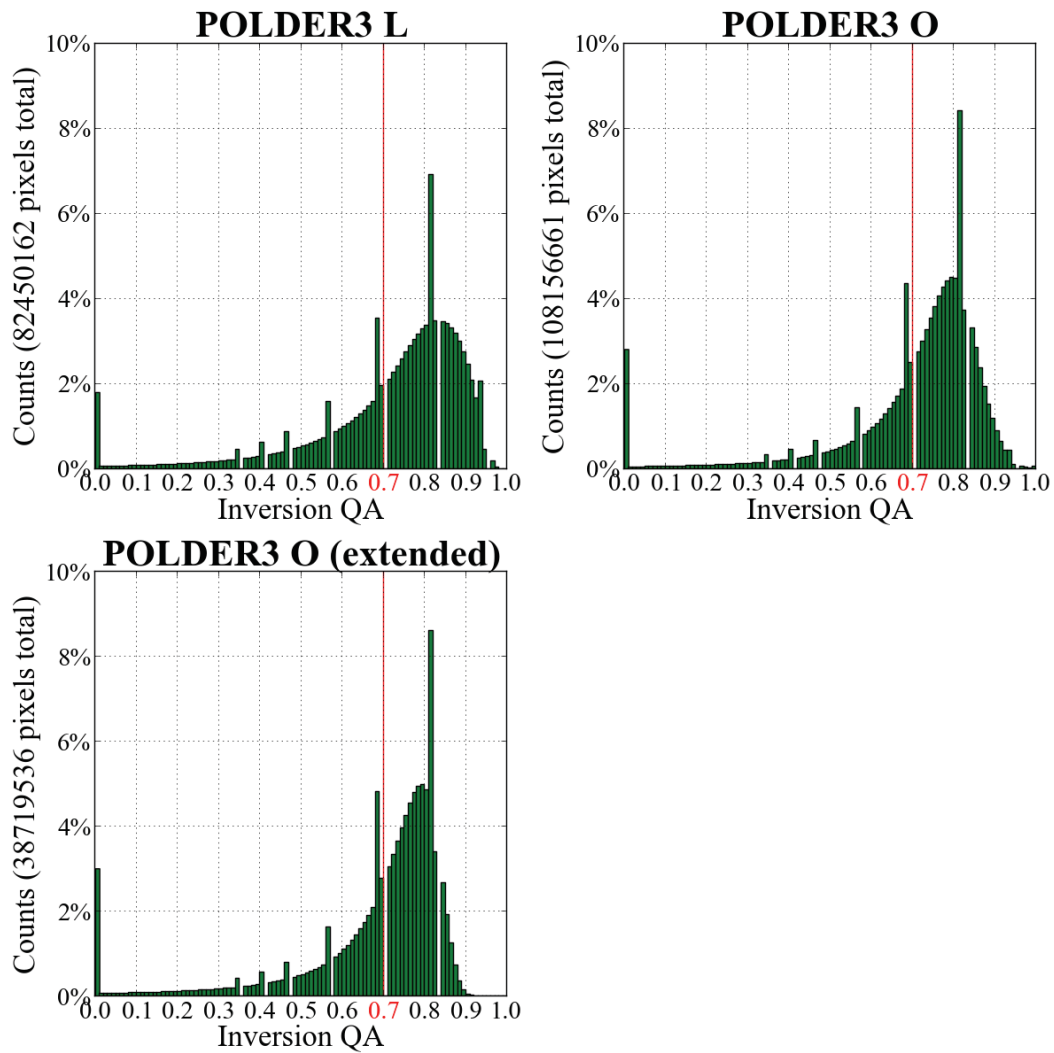
Figure 3. The total quantities of POLDER3 Land ('L') and Ocean ('O') pixels based on different values of 'Quality of Inversion' flag during the analyzed period of July 2006 - December 2010. The 'POLDER3 O (extended)' histogram is based on those pixels in Ocean retrievals, where the retrieval algorithm considered the sensor viewing geometry conditions to be especially 'favorable' and produced a set of additional aerosol parameters, such as spherical large-mode AOD, Refractive Index of fine mode, and others. The quality flag values are binned into 0.01 intervals and the red line indicate the 0.7 QA threshold used in this study.

1

2
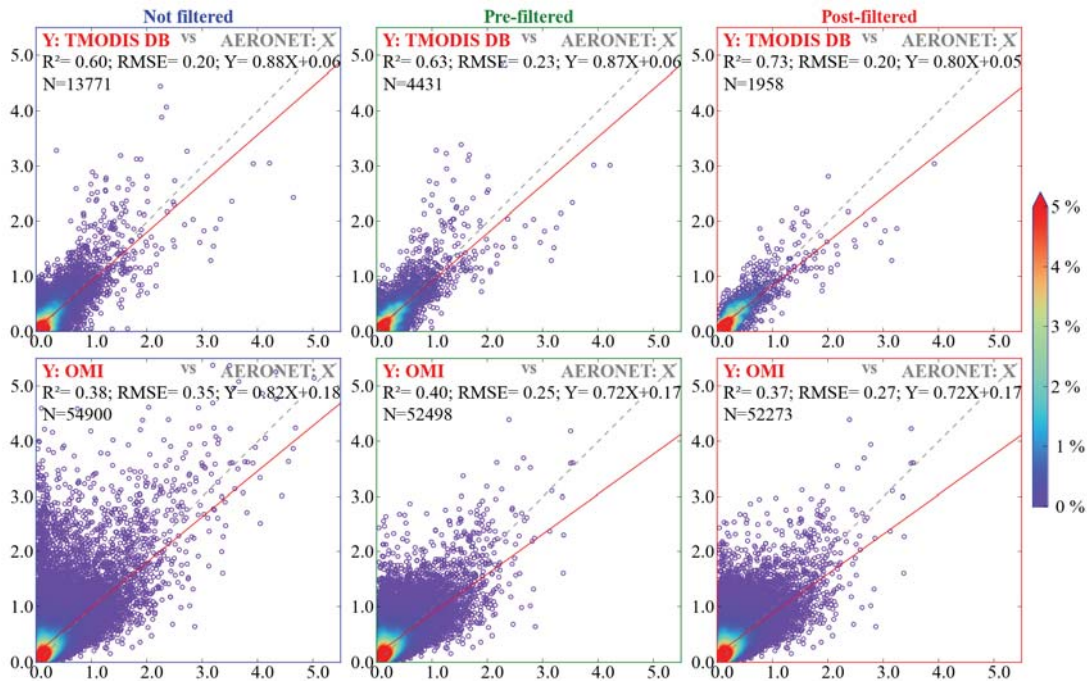
3   Figure 4. Effects of two different data QA filtering schemes on the accuracy of the global

4   collocated spaceborne AOD, as discussed in Section 4. AERONET AOD data are shown on

5   the X-axes, while AOD measured by spaceborne sensors are on the Y-axes. Color of each

6   data point indicates the percentage of all data points that fall within 0.05 AOD of this point (in

7   Cartesian coordinates). Left column displays the original unfiltered data with all QA values.

8   Middle column displays the data *pre-filtered* by QA, where individual pixels in each data

9   sample were filtered based on their QA values *before* calculating the mean value of the

10   sample. Right column shows the data *post-filtered* by QA, where the mean of each sample

11   was calculated based on all pixels in the sample; *after* this, the whole sample was rejected if at

12   least half of the pixels in the sample had QA values below the specified threshold. Note that

13   OMI data have better properties when pre-filtered, while Terra MODIS – Deep Blue data are

14   in a better agreement with AERONET AOD when pos-filtered.

15

Figure 5. Impact of QA screening on the statistical properties of AOD retrieved by different
sensors over Djougou, Benin. The top part of the figure shows scatter plots of 2 years of data
that is unfiltered (left) or pre-filtered (right) by QA. It can be observed that while filtering
improved the properties of certain datasets, it degraded the properties of the others,
particularly Aqua MODIS Deep Blue and OMI. This effect can be partially explained by
observing that the retrieval algorithms can mistakenly assign bad QA to pixels with good
retrievals, as demonstrated in a high-AOD event in the bottom part of the picture, e.g., see
AMODIS DB on Dec. 11 and Dec 19.
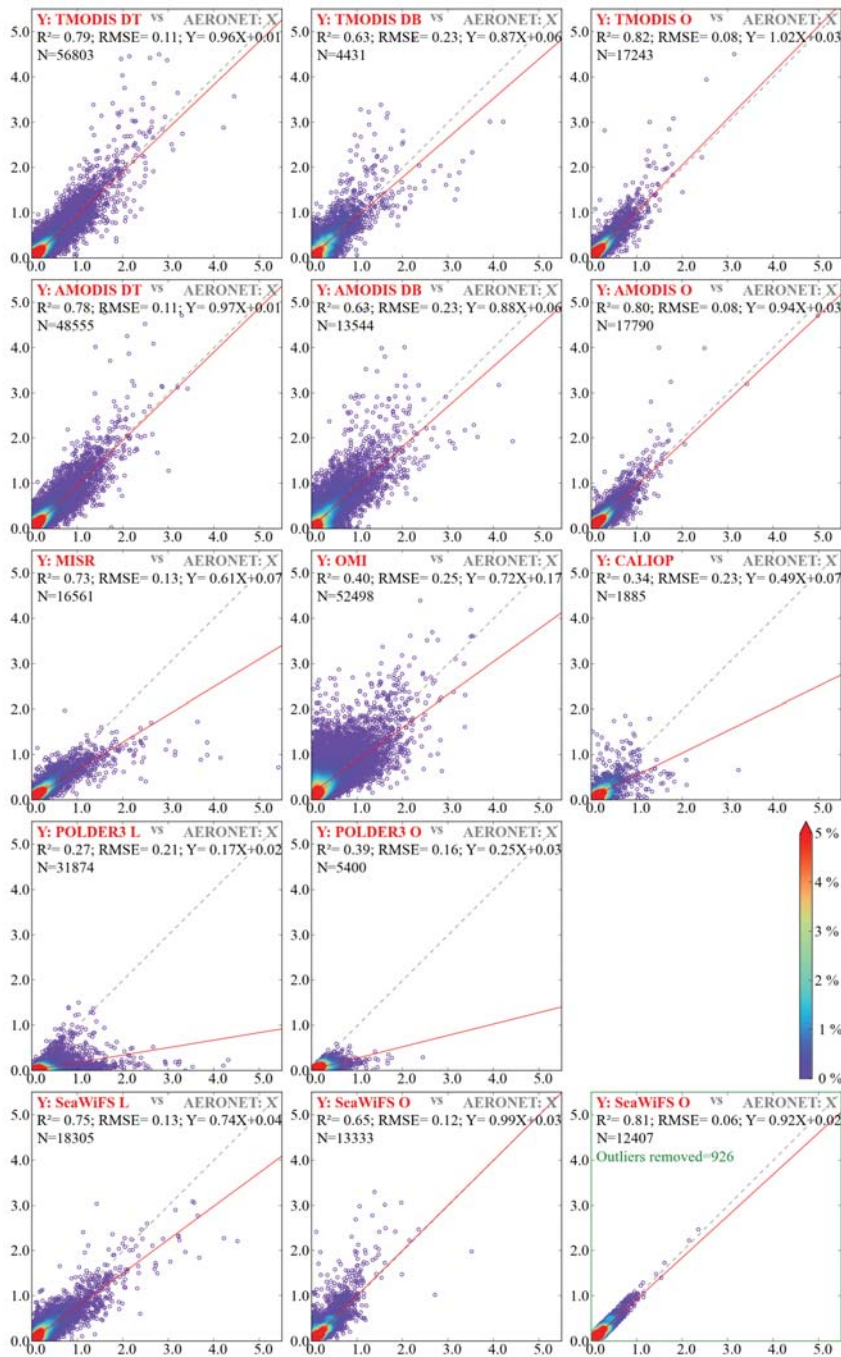
Figure 6. Regression fits of AERONET AOD (x-axes) to AOD measured by spaceborne sensors (y-axes). Satellite data were pre-screened by QA as explained in Section 4. The color of each data point indicates the percentage of all data points on the plot that fall within 0.05 AOD of this point (in Cartesian coordinates). Scatter plot in the green frame demonstrates the results of the possible data outlier detection and removal procedure described in Section 5.
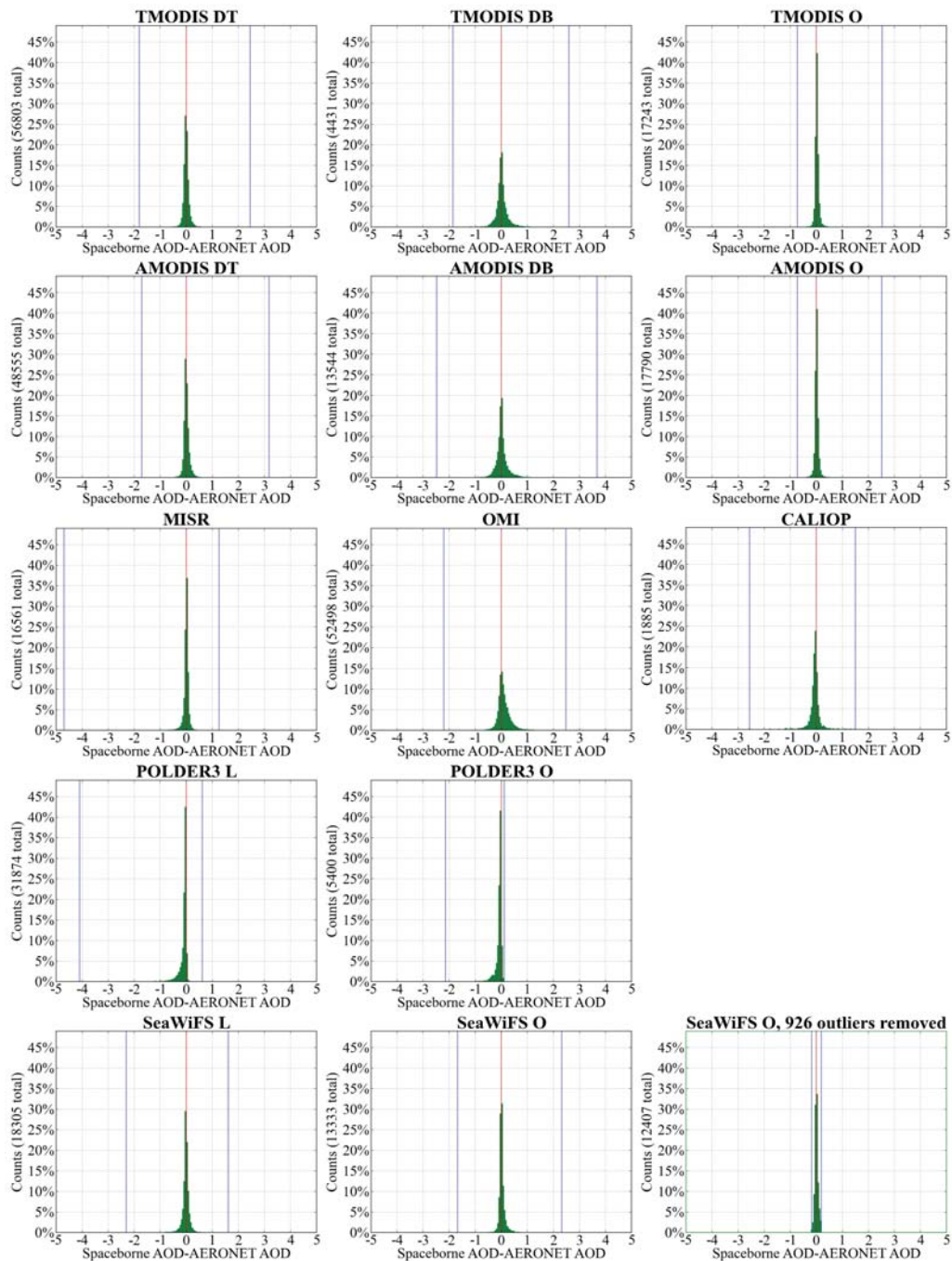
Figure 7. Distribution of the difference (residuals) between Spaceborne AOD and AERONET AOD. Satellite data were pre-screened by QA as explained in Section 4. In each histogram, the data are split into equal-length bins of 0.05 AOD. The red vertical line indicates the residual of 0 AOD, while the blue lines mark minimum and maximum residuals of each distribution. Histogram in the green frame demonstrates the results of the possible data outlier detection and removal procedure described in Section 5.
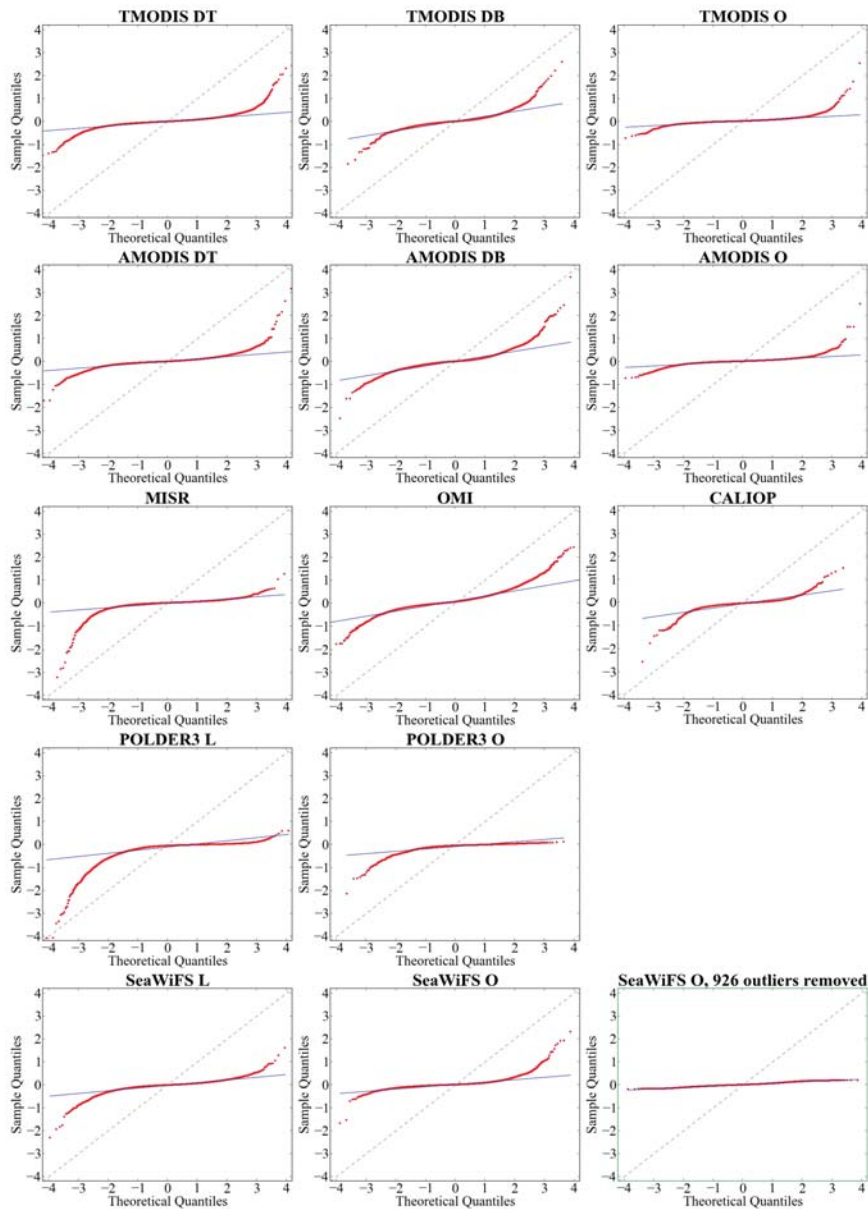
Figure 8. Normality of the difference between Spaceborne AOD and AERONET AOD. In each plot, points closely following the blue fitted line indicate the data that are approximately normally distributed. Curvatures around the center of the straight line represent the departure from the normality and indicate a presence of possible outliers, particularly at the tails of the distributions. The difference in the slope and offset of the fitted blue line from the gray 1:1 line indicates a deviation from the standard location (i.e., mean=0) and scale (i.e., standard deviation=1) of the normal distribution. Satellite data were pre-screened by QA as explained in Section 4. Plot in the green frame demonstrates the results of the possible data outlier detection and removal procedure described in Section 5.

40
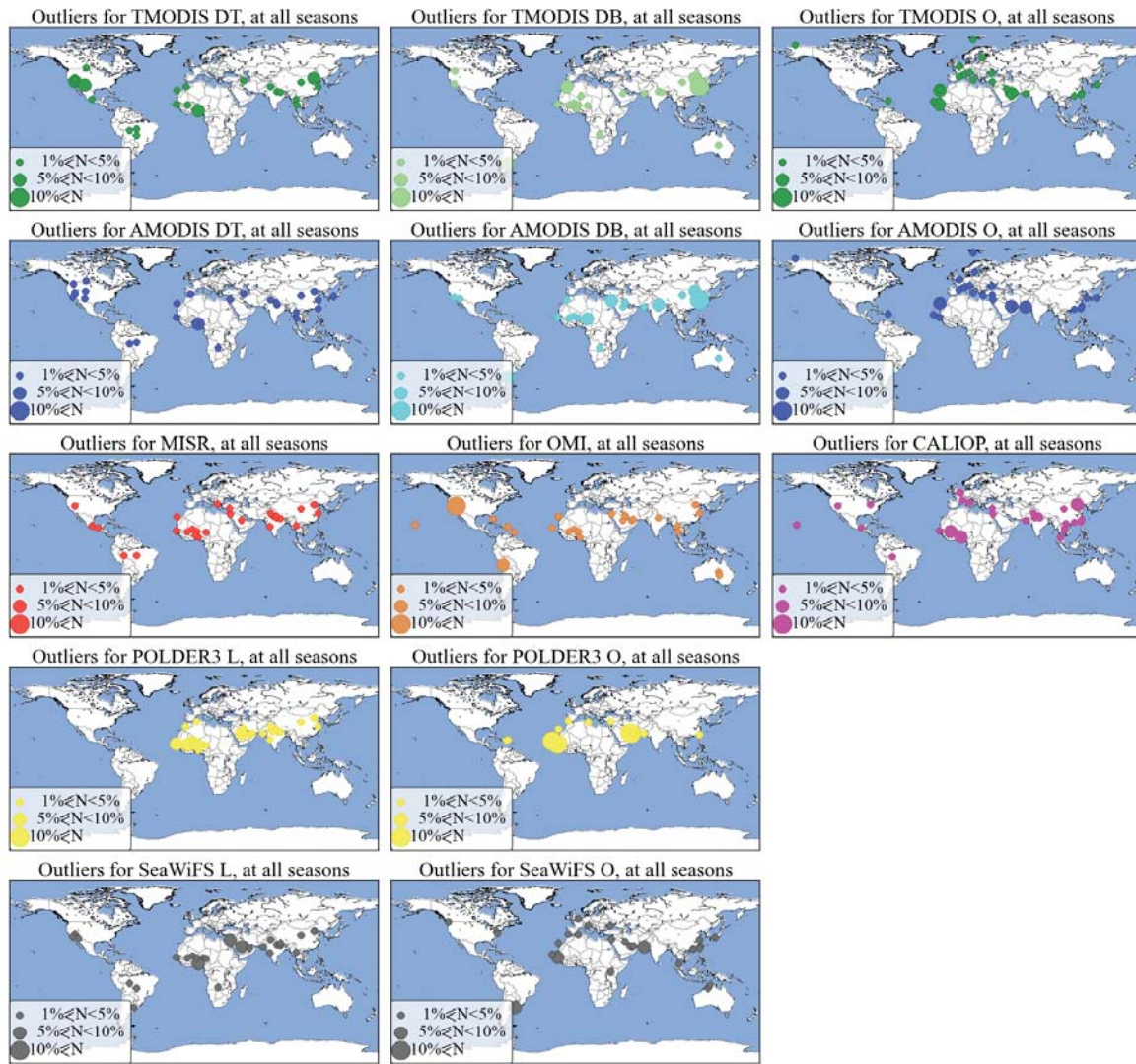
Figure 9. Distribution of the possible data outliers for the studied spaceborne aerosol data sets. Displayed values are percentages from all outliers detected for each of the data sets as listed in the 4th column of Table 3. Stations with less than 1% from the total number of outliers are not shown. The statistical technique for detection and removal of the possible data outliers is described in Section 5.
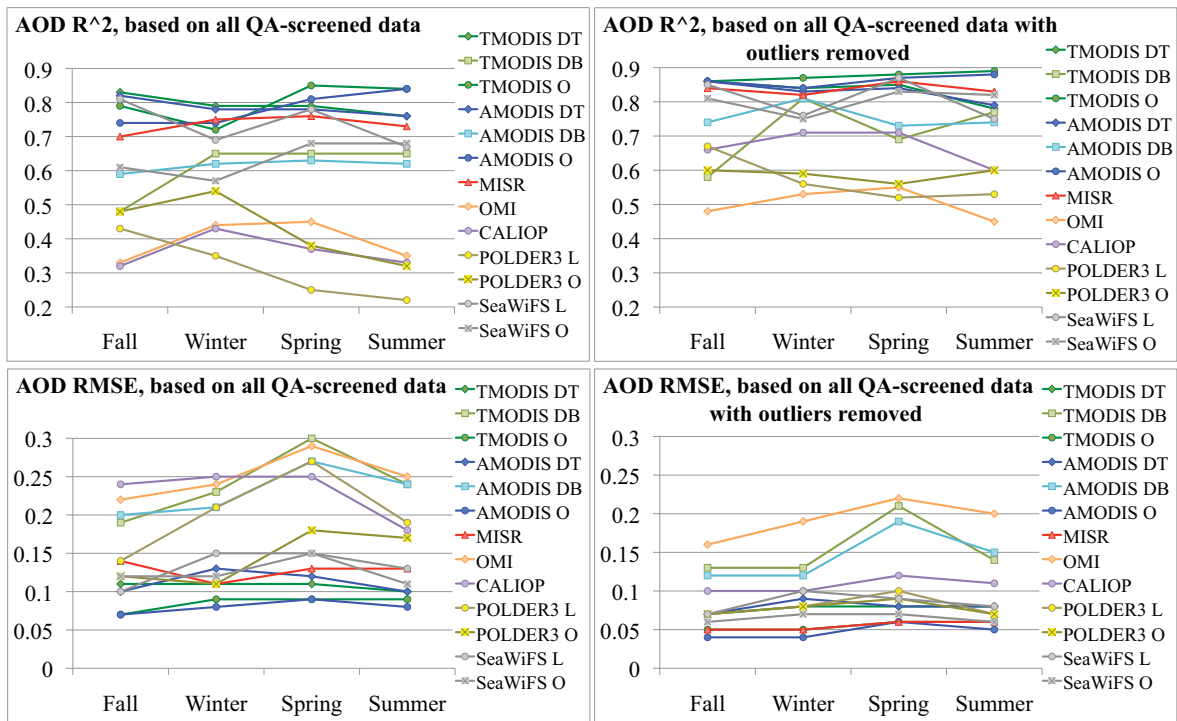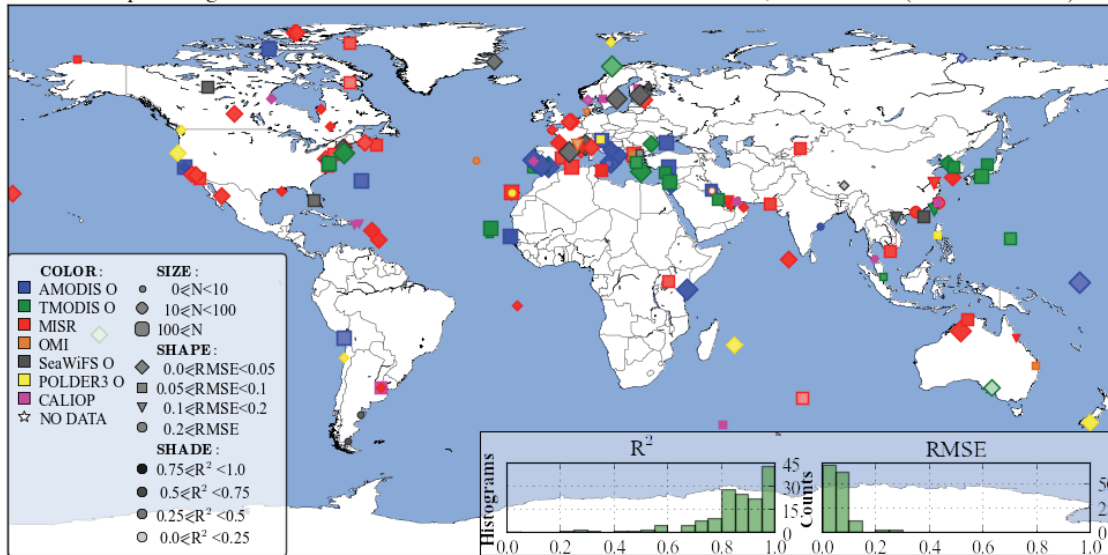
Figure 10. Seasonal dependence of squared linear fit correlation coefficient (R^2) and root mean square error (RMSE) statistics between the collocated spaceborne and ground-based (AERONET) observations of AOD, based on the data in Table 3.
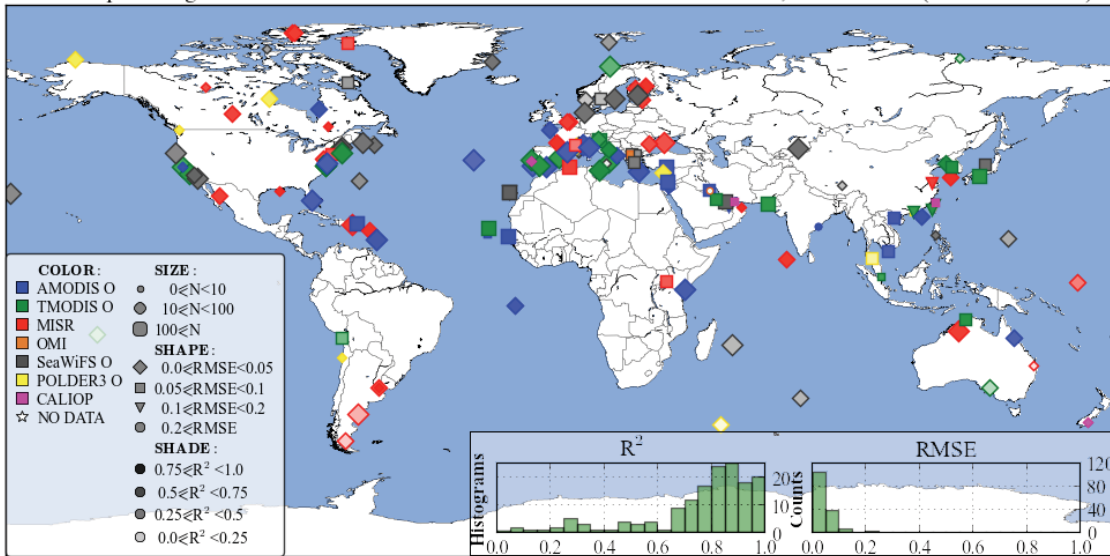
Figure 11. Spaceborne datasets with the best correlation (R^2) of the retrieved AOD to the AOD measured by inland (top) and coastal or island-based (bottom) AERONET sites. The intensity of marker shading indicates the degree of correlation. Marker shape indicates the range of root mean square error (RMSE) associated with the displayed best R^2. Finally, marker size corresponds to the number of collocated data points used to compute the displayed statistics. Histograms in the bottom insets highlight the distribution of these statistics over all sites based on bins of 0.05 AOD. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sections 4 and Section 5.

Sensors providing the best RMSE of AOD over land at 368 AERONET stations, at all seasons (outliers removed)
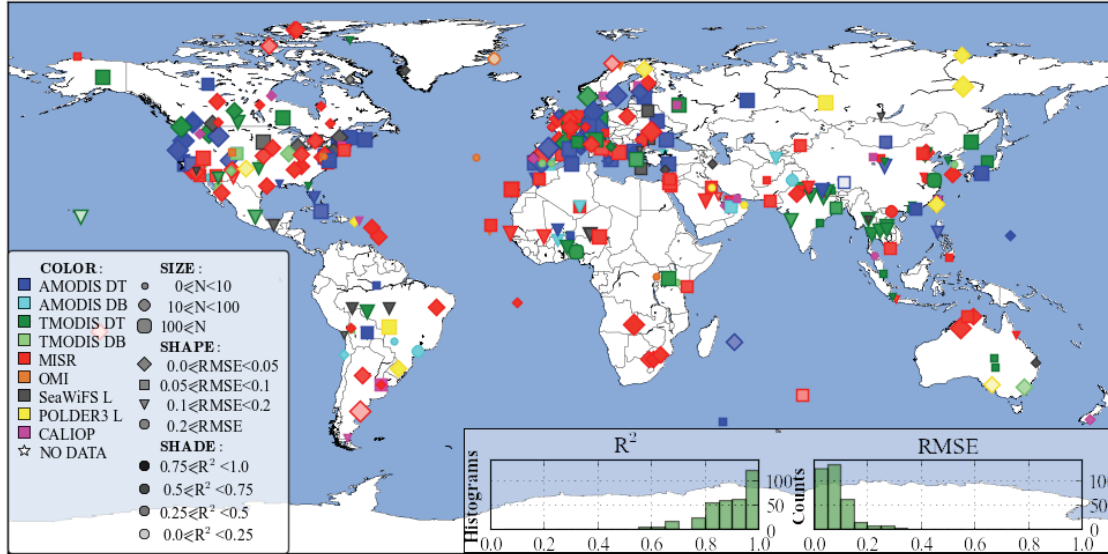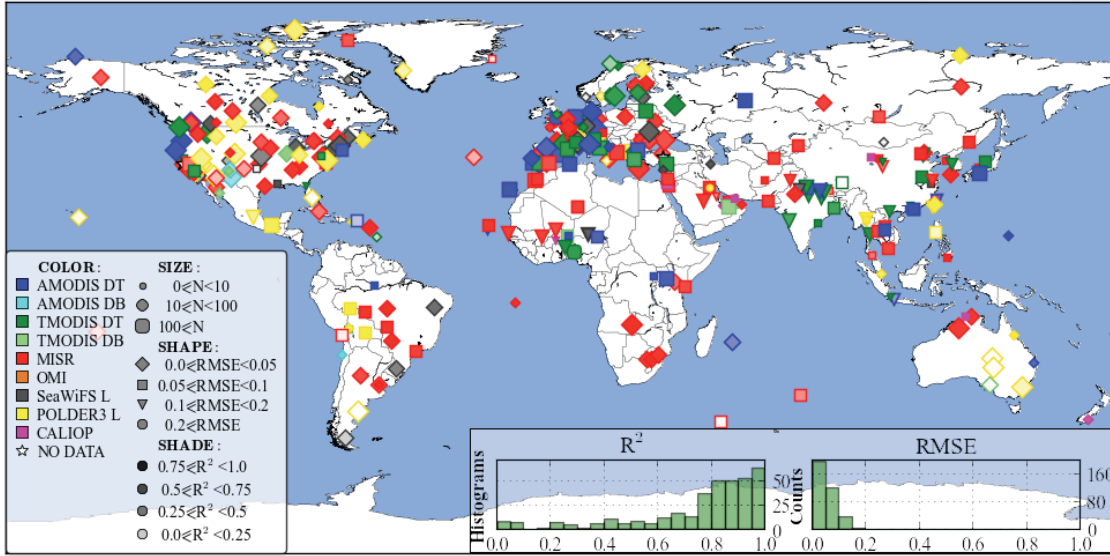
Sensors providing the best RMSE of AOD over ocean at 154 AERONET stations, at all seasons (outliers removed)

Figure 12. Spaceborne datasets with the best root mean square error (RMSE) of the retrieved AOD to the AOD measured by inland (top) and coastal or island-based (bottom) AERONET sites. The symbols used are the same as the symbols in Figure 7. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sections 4 and Section 5.
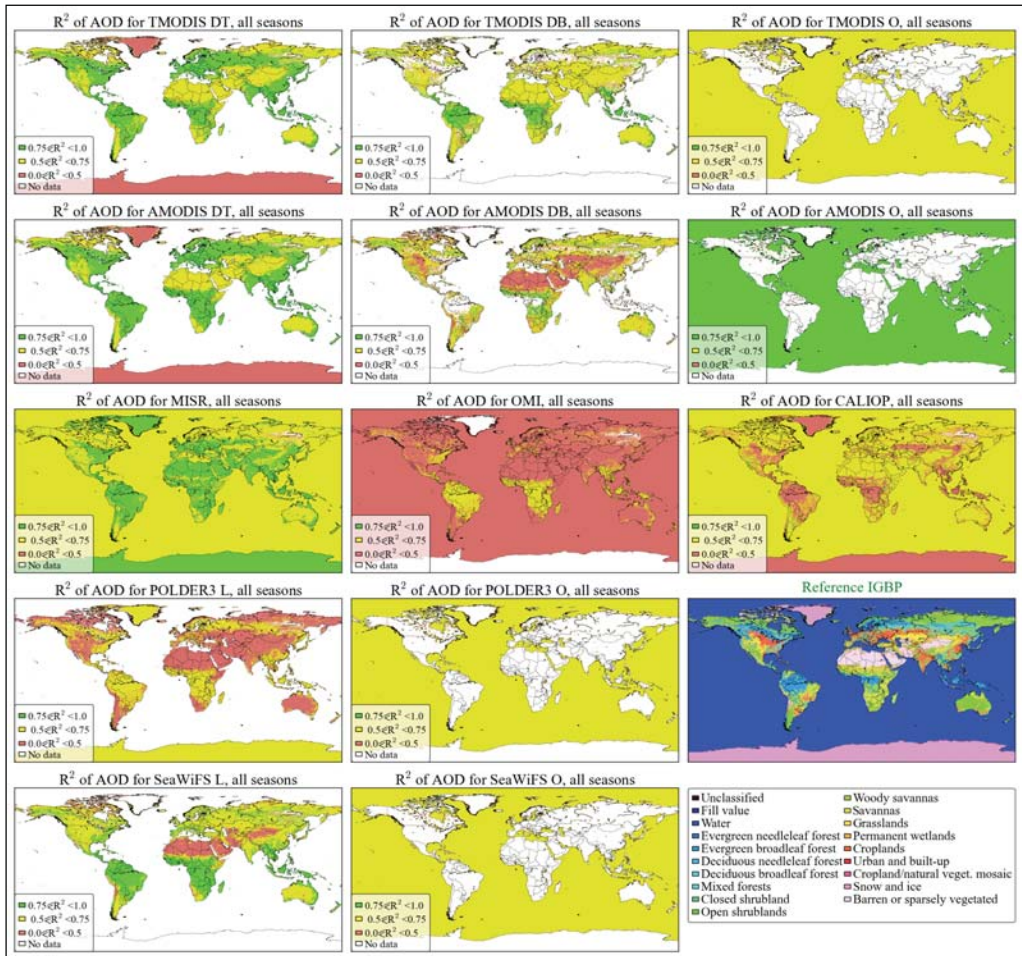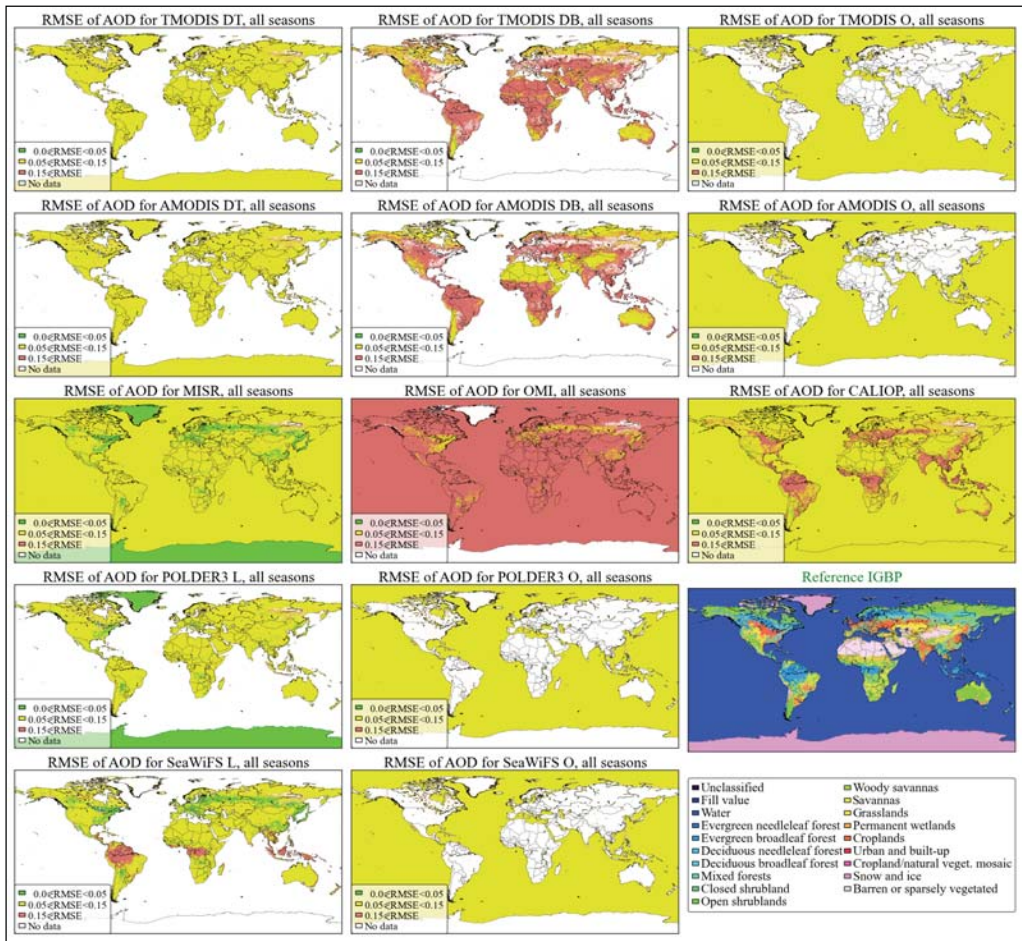
Figure 13. Land cover type dependence of squared linear fit correlation coefficient (R2) between the collocated spaceborne and ground-based (AERONET) observations of AOD. Areas corresponding to each IGBP land cover type (bottom right inset) are colored based on the average of the data from those AERONET sites that reside in these areas. The statistics were calculated based on data that were pre-filtered by QA and screened of outliers as described in Sections 4 and Section 5.

1

Figure 14. Land cover type dependence of root mean square error (RMSE) between the collocated spaceborne and ground-based (AERONET) observations of AOD. Areas corresponding to each IGBP land cover type (bottom right inset) are colored based on the average of the data from those AERONET sites that reside in these areas. The statistics were calculated based on the data that were pre-filtered by QA and screened of outliers as described in Sections 4 and Section 5.