

Validation and uncertainty estimates for MODIS Collection 6 “Deep Blue” aerosol data

A. M. Sayer,^{1,2} N. C. Hsu,² C. Bettenhausen,^{2,3} and M.-J. Jeong⁴

Received 27 March 2013; revised 20 June 2013; accepted 24 June 2013; published 25 July 2013.

[1] The “Deep Blue” aerosol optical depth (AOD) retrieval algorithm was introduced in Collection 5 of the Moderate Resolution Imaging Spectroradiometer (MODIS) product suite, and complemented the existing “Dark Target” land and ocean algorithms by retrieving AOD over bright arid land surfaces, such as deserts. The forthcoming Collection 6 of MODIS products will include a “second generation” Deep Blue algorithm, expanding coverage to all cloud-free and snow-free land surfaces. The Deep Blue dataset will also provide an estimate of the absolute uncertainty on AOD at 550 nm for each retrieval. This study describes the validation of Deep Blue Collection 6 AOD at 550 nm (τ_M) from MODIS Aqua against Aerosol Robotic Network (AERONET) data from 60 sites to quantify these uncertainties. The highest quality (denoted quality assurance flag value 3) data are shown to have an absolute uncertainty of approximately $(0.086+0.56\tau_M)/AMF$, where AMF is the geometric air mass factor. For a typical AMF of 2.8, this is approximately $0.03+0.20\tau_M$, comparable in quality to other satellite AOD datasets. Regional variability of retrieval performance and comparisons against Collection 5 results are also discussed.

Citation: Sayer, A. M., N. C. Hsu, C. Bettenhausen, and M.-J. Jeong (2013), Validation and uncertainty estimates for MODIS Collection 6 “Deep Blue” aerosol data, *J. Geophys. Res. Atmos.*, 118, 7864–7872, doi:10.1002/jgrd.50600.

1. Introduction

[2] Introduced in Collection 5 (C5) of the Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol dataset, the “Deep Blue” algorithm [Hsu *et al.*, 2004, 2006] provided operational retrievals of aerosol optical depth (AOD) over bright arid land surfaces (such as deserts) from MODIS for the first time. Previously, such coverage had been absent from the MODIS aerosol product because the assumptions made in the “Dark Target” algorithm [Levy *et al.*, 2007], in use over vegetated land surfaces, did not hold for such bright surfaces. Deep Blue data are included alongside the Dark Target retrievals within the MxD04 (level 2, swath-level) and MxD08 (level 3, aggregated spatiotemporally) products, where x is O for MODIS Terra and Y for MODIS Aqua.

[3] The combination of two instruments in flight (aboard the Terra satellite from 2000 onward and Aqua from 2002 onward), and a wide swath giving daily near-global observations, make the MODIS sensors an attractive choice for such an aerosol dataset. Deep Blue data have since been well-used for various applications, such as identification of dust

sources [Ginoux *et al.*, 2012; Karimi *et al.*, 2012; Schepanski *et al.*, 2012], evaluation/development of chemistry transport models [Draxler *et al.*, 2010; Laurent *et al.*, 2010; Wang *et al.*, 2012], and comparison with other satellite-based dust aerosol datasets [DeSouza-Machado *et al.*, 2010; Carboni *et al.*, 2012]. Although other instruments on polar-orbiting satellites, notably the Multiangle Imaging Spectroradiometer (MISR) [Martonchik *et al.*, 1998] and Along-Track Scanning Radiometer (ATSR) [North, 2002; Sayer *et al.*, 2012a] series, are also able to retrieve AOD over deserts, they suffer from a more limited spatial coverage. The Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensors above Africa have also been used to retrieve AOD over deserts [Brindley and Russell, 2009; Wagner *et al.*, 2010; Banks and Brindley, 2013], giving high temporal resolution over the Earth disk viewed, although their geostationary platforms mean that global coverage cannot be achieved.

[4] The next version of the MODIS product suite (Collection 6, C6) is expected to be released in late 2013. Aerosol data will be available from the MODIS Level 1 and Atmosphere Archive and Distribution System at <http://ladsweb.nascom.nasa.gov/>. The MODIS aerosol products will include a “second generation” Deep Blue algorithm (N. C. Hsu *et al.*, Enhanced Deep Blue aerosol retrieval algorithm: The 2nd generation, submitted to *J. Geophys. Res.*, 2013, hereinafter referred to as Hsu *et al.*, submitted manuscript, 2013), building on experiences with the C5 data; many of these updates were developed in the recent implementation of Deep Blue to Sea-viewing Wide Field-of-view Sensor (SeaWiFS) data [Sayer *et al.*, 2012b]. One important feature of C6 is that Deep Blue will provide coverage over

¹Goddard Earth Sciences Technology and Research, Universities Space Research Association, Greenbelt, Maryland, USA.

²NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

³Science Systems and Applications, Inc., Lanham, Maryland, USA.

⁴Gangneung-Wonju National University, Gangneung City, Korea.

Corresponding author: A. M. Sayer, Climate and Radiation Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. (andrew.sayer@nasa.gov)

©2013. American Geophysical Union. All Rights Reserved.
2169-897X/13/10.1002/jgrd.50600

all cloud-free land surfaces except for snow-covered pixels, meaning that, for a significant fraction of the world, both Dark Target and Deep Blue retrievals will be available.

[5] To facilitate data use and better inform users of retrieval quality, the Deep Blue dataset will also provide an estimate of the uncertainty (one-standard-deviation confidence interval) on AOD at 550 nm for each retrieval. The purpose of this study is to act as a description of how these uncertainty estimates have been calculated and provide an indication of the expected regional and global performance of the C6 Deep Blue algorithm. Section 2 discusses the data and matchup methodology used, section 3 provides global and regional validation results, and section 4 provides a summary and discussion of the results in the context of satellite-based AOD records.

2. Datasets Used

2.1. AERONET

[6] The Aerosol Robotic Network (AERONET) of Sun photometers provides a dataset of spectral AOD with low uncertainty (~ 0.01 – 0.02) and high temporal resolution (~ 15 min) under cloud-free conditions, by observing the extinction of direct solar radiation through the aerosol-laden atmosphere [Holben *et al.*, 1998]. Although the wavelengths at which AOD is reported vary for different sites (and can vary for different periods at a given site), almost all AERONET sites provide AOD at 440, 500, 675, and 870 nm at a minimum. AERONET represents a standard resource for the validation and bias-correction of satellite AOD datasets [e.g., Ichoku *et al.*, 2002; Zhang and Reid, 2006; Levy *et al.*, 2010; Kahn *et al.*, 2010; Hyer *et al.*, 2011; Sayer *et al.*, 2012b; Carboni *et al.*, 2012], among other applications, due to its high data quality, consistency of processing standards, wide global range of sites, and free and simple data access. In this study, the current Level 2.0 Version 2 (cloud-screened and quality-assured) [Smirnov *et al.*, 2000] direct-Sun dataset is used, from a total of 60 sites, chosen for their large data records and to provide a representative set of geometric, atmospheric, and surface conditions. The locations of these sites are shown in Figure 1.

2.2. MODIS Deep Blue

[7] Initially, described by Hsu *et al.* [2004], an updated MODIS Deep Blue (hereafter “MODIS”) algorithm description is given by Hsu *et al.* (submitted manuscript, 2013). A brief summary is provided here. The aerosol data are provided at a nominal spatial resolution of $10\text{ km} \times 10\text{ km}$ at the subsatellite point, although the sensor scanning design means that pixels near the edge of the scan are significantly larger than those at the center. The algorithm determines spectral AOD first at nominal $1\text{ km} \times 1\text{ km}$ spatial resolution for clear-sky snow-free pixels using observed top-of-atmosphere (TOA) reflectance at 412, 470, and 650 nm (although over deserts, the 650 nm band is only used in conditions of heavy dust aerosol loading). This is achieved by finding the best match between these observations and precalculated reflectances stored in a lookup table as a function of solar/sensor geometry, AOD, aerosol scattering/absorption properties, and surface reflectance.

[8] The surface reflectance assumed in the retrieval is prescribed dynamically from one of a variety of methods

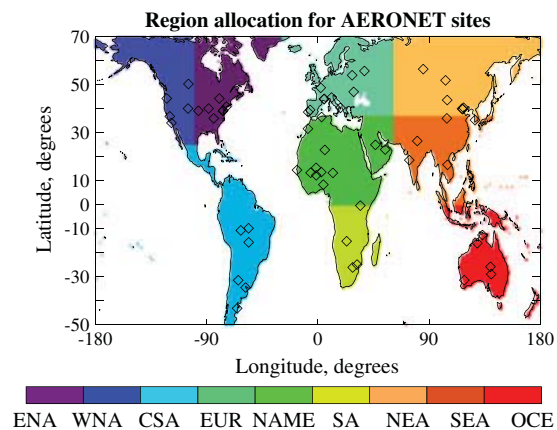


Figure 1. Geographical bounds of regions used in this study for grouping of MODIS and AERONET AOD. Locations of the AERONET sites used are indicated by black diamonds. Region names and associated abbreviations are Eastern North America (ENA), Western North America (WNA), Central/South America (CSA), Eurasia (EUR), North Africa/Middle East (NAME), Southern Africa (SA), Northeast Asia (NEA), Southeast Asia (SEA), and Oceania (OCE).

dependent on location/surface type, normalized difference vegetation index (NDVI), and solar/sensor geometry; one of the main principles of the algorithm is that the reflectance of most surfaces is low at 412 nm (and to a lesser extent at 470 nm), increasing the relative contribution from aerosols to the radiance measured at TOA by the satellite sensor. Aerosol phase function and single scatter albedo (SSA) at each wavelength are prescribed as a function of location and season, unless heavy dust is detected, in which case, a maximum likelihood method is used to pick from a suite of aerosol optical models (and in this case, the SSA is also reported based on this optimal aerosol model).

[9] The individual $1\text{ km} \times 1\text{ km}$ retrievals are then averaged to the $10\text{ km} \times 10\text{ km}$ “retrieval product pixel” (hereafter “retrieval”) scale. Each such retrieval has an associated quality assurance (QA) flag, determined by tests on the $1\text{ km} \times 1\text{ km}$ -pixel AOD within each $10\text{ km} \times 10\text{ km}$ retrieval, with $QA = 3$ indicating the retrievals of highest confidence and $QA = 2/QA = 1$ progressively lower confidence (due to factors such as possibility of cloud contamination or heterogeneity of scene).

[10] Although there are two MODIS sensors in orbit, this analysis concerns only that mounted on the Aqua satellite; as at the present time, radiometric calibration updates for Terra are not finalized. As the sensors are near-identical, and the algorithms applied identical, Deep Blue performance is expected to be very similar for the two sensors. The C6 products for Aqua are anticipated to be released prior to those for Terra, and an analysis of the type described herein will be performed to determine the uncertainty on Terra C6 Deep Blue data, which will also be provided in the MOD04 products.

2.3. Matchup Methodology

[11] As AERONET provides a point measurement repeated frequently in time, while satellites provide a

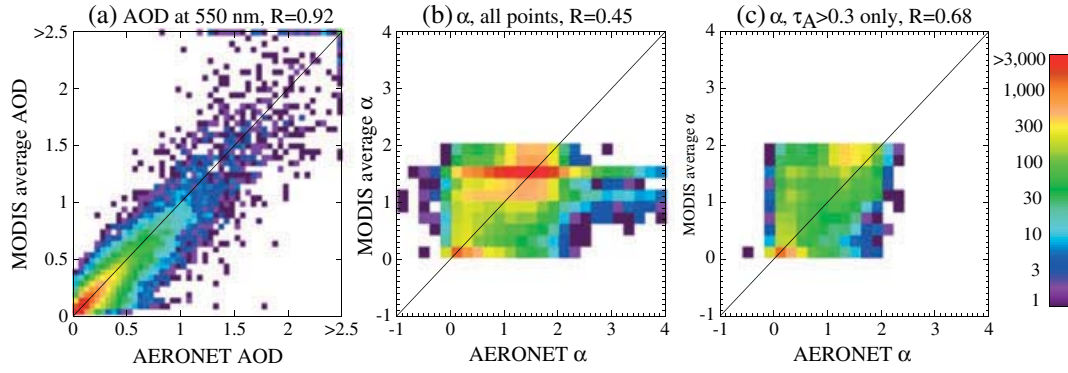


Figure 2. Scatter density histograms comparing AERONET and MODIS (a) AOD at 550 nm, (b) α , and (c) α for only those points where $\tau_A > 0.3$. Only QA = 3 MODIS data are used.

snapshot of a larger region at a single time, validation exercises of this type (e.g., previously cited references) typically proceed by taking a temporal average of AERONET data around the time of the satellite overpass, and a spatial average of the satellite data over the ground site. This mitigates the effect of variability in the underlying aerosol field. In this case, AERONET data averaged within 30 min of the MODIS overpass are extracted and compared with MODIS data averaged within a 25 km radius of the AERONET site. As AERONET does not make measurements at 550 nm, data are interpolated to 550 nm using the standard Ångström exponent α , defined

$$\alpha = -\frac{\ln(\tau_1/\tau_2)}{\ln(\lambda_1/\lambda_2)}, \quad (1)$$

where τ_1 , τ_2 are the AOD at wavelengths λ_1 , λ_2 . The nearest available pair of bounding wavelengths from AERONET (normally 675 nm and either 440 or 500 nm), together with the appropriate α , are used. Negligible uncertainty is introduced through this spectral interpolation. Hereafter, the spatiotemporally matched MODIS and AERONET AOD data are denoted τ_M and τ_A , respectively. All references to “AOD” indicate 550 nm, unless otherwise indicated, and so wavelength subscripts will be dropped for conciseness.

3. Validation Results

3.1. Global Scatter Plots

[12] Figure 2 compares AERONET and MODIS AOD and α for the resulting matched dataset. Note the AERONET α used is defined using the wavelength pairs 440/675 nm; MODIS provides either 412/470 nm or 470/650 nm (dependent on AOD and the underlying surface) (Hsu et al., submitted manuscript, 2013). The AOD is well-correlated between the two datasets ($R = 0.92$), with increased scatter for high aerosol loadings. Correlation and scatter at other wavelengths reported by AERONET and MODIS (not shown; 470 and 650 nm, together with 412 nm for some sites) are similar. The majority of the data (75.4%) are for low aerosol loadings ($\tau_A < 0.3$).

[13] In contrast, and as expected from the known information content of such sensors and algorithm types [e.g., Wagner and Silva, 2008; Levy et al., 2010; Sayer et al., 2012b], the comparison with α shows significant scatter. This is somewhat improved for those cases where $\tau_A >$

0.3. Note that in low-AOD conditions, the uncertainty on AERONET α itself can be non-negligible [Wagner and Silva, 2008]; Figure 2c shows that the removal of low-AOD cases also removes many of the extreme values of α from the AERONET record. Because of a lack of information on the parameter, and to prevent retrieval of unphysical values, the Deep Blue algorithm permits retrieval of $0 \leq \alpha \leq 1.8$, and, in low-AOD conditions, fixes $\alpha = 1.5$ over vegetated surfaces and limits to $\alpha \leq 1$ over bright surfaces (Hsu et al., submitted manuscript, 2013). This leads to the observed artifacts in Figure 2. Multiangular and, ideally, polarimetric measurements are required to infer aerosol-related quantities other than AOD reliably over land, particularly under low-AOD conditions [Hasekamp and Landgraf, 2007; Kokhanovsky et al., 2010]. For these reasons, α will not be considered further here.

[14] If the time window is shrunk to 15 min, and the spatial window restricted to consider only the MODIS retrieval in whose footprint the AERONET site lies, the data volume decreases by a factor of approximately 2, while other results change only negligibly. This indicates that the bulk of the disagreement between MODIS and AERONET is likely due to factors such as systematic uncertainties in aerosol/surface properties for a given location and time, rather than radiometric noise or true spatiotemporal variability (although these can remain as factors). Thus, to increase the data volume, only the 25 km/30 min comparison is considered hereafter.

3.2. Calculating AOD Uncertainty Estimates

[15] In low-AOD conditions, the TOA reflected radiance observed by sensors such as MODIS arises largely from surface reflectance and (particularly at shorter visible wavelengths) Rayleigh scattering, while the aerosol contribution is small. As AOD increases, so does, generally, the contribution to TOA radiance (although there remain some conditions in which TOA radiance is insensitive to AOD) [e.g., Seidel and Popp, 2012]. For these reasons, quoted absolute uncertainties on satellite AOD products are often of the form $a+b\tau_A$, where a and b are the constants to be determined. The size of a is linked to factors such as uncertainties in modeling surface reflectance, while b is more associated with assumptions about aerosol microphysical properties made in the retrieval (although this is a simplistic first-order separation).

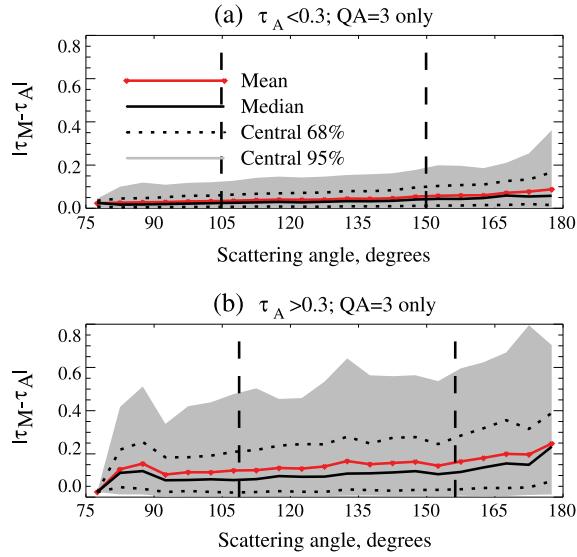


Figure 3. Absolute error on MODIS (Aqua) Deep Blue AOD at 550 nm compared to AERONET, for (a) $\tau_A < 0.3$ and (b) $\tau_A > 0.3$, as a function of scattering angle (180° indicates the backscatter direction), in 5° bins. Within each bin, the red line indicates the mean absolute error, the black line the median, the dotted black the central 68% of the data, and the shaded grey the central 95% of the data. The vertical dashed lines indicate the central 68% of MODIS scattering angles for the matched data. Only QA = 3 MODIS data are used.

[16] The objective is to determine a and b such that the resulting confidence envelope contains one standard deviation (1σ , i.e., $\sim 68\%$) of MODIS/AERONET matchups. If this is taken as a true Gaussian measure of uncertainty, the corresponding 2σ envelope should then contain $\sim 95\%$ of matchups, and the 0.5σ envelope $\sim 38\%$. Although no comprehensive validation exercise was published for C5, experience with the data suggested $0.05 + 0.2\tau_A$ as the typical level of uncertainty.

[17] For the C6 Deep Blue dataset, the approach is subtly different: the uncertainty estimates are defined in terms of τ_M rather than τ_A . This is because only τ_M is available to compute these uncertainty estimates for inclusion in the data product, so the validation exercise should be framed in a self-consistent manner. A “prognostic” (as opposed to “diagnostic”) uncertainty model of this type is more useful for applications such as data assimilation [Shi *et al.*, 2013].

[18] Figures 2 and 3 show that the absolute error in retrieved AOD is generally larger for high-AOD cases than for clean conditions. Figure 3 also reveals the fact that, for both low-AOD and high-AOD conditions, the absolute uncertainty is larger, and its distribution wider, for retrievals closer to backscattering geometries (scattering angles near 180°). Such geometries are generally less favorable for the determination of aerosol loading because the lower atmospheric path length near backscatter means that the atmospheric contribution to TOA radiance is smaller than at more oblique views, and additionally, because the bidirectional reflectance distribution function of many surfaces tends to show a “hotspot” at these geometries [e.g., Wanner *et al.*, 1995], further decreasing the relative

atmospheric contribution to the signal. These near-backscatter geometries are encountered by MODIS more frequently in the tropics than at high latitudes. Note that this angular dependence is also observed in other datasets [e.g., Levy *et al.*, 2010; Hyer *et al.*, 2011]. For these reasons, the absolute expected error (EE) confidence envelope is here defined

$$EE = \frac{a + b\tau_M}{\mu_0 + \mu}, \quad (2)$$

where $\mu_0 + \mu$, the sum of the cosines of the solar and viewing zenith angles, is the geometric air mass factor (AMF). The minimum possible AMF is 2 for a nadir-viewing observation with overhead Sun, and the maximum encountered by MODIS ~ 5.2 ; the mean over the matched MODIS/AERONET data used in this study is ~ 2.8 . Figure 4 shows the relationship between τ_M and the 38th, 68th, and 95th percentiles of absolute error multiplied by AMF, for QA = 3 points, with data grouped into bins by ascending MODIS AOD (in bins of 500 matchups). The 68th percentiles show a strong linear relationship, whose least squares regression equation gives $EE = (0.086 + 0.56\tau_M)/(\mu_0 + \mu)$. The corresponding linear fit equation for QA = 2 points is $(0.10 + 0.60\tau_M)/(\mu_0 + \mu)$, and for QA = 1 points is $(0.083 + 0.83\tau_M)/(\mu_0 + \mu)$. These equations are used to define the uncertainty estimates provided with the C6 Deep Blue data.

[19] Taking twice this envelope results in a line which slightly underestimates the 95th percentile of the error distribution (Figure 4), indicating that this metric does not quite provide a true Gaussian estimate of the uncertainty. However, the deviations from this relationship, particularly for conditions of low and moderate AOD, are mostly small. For the highest MODIS AOD bin, the 95th percentile lies significantly above the expected line; this is due to a small number of cases where a retrieval with either severe cloud contamination or underestimation of surface reflectance near an AERONET site is assigned QA = 3 and the true AOD is low. Taking half the 68th percentile line gives a very close agreement to the 38th percentile, confirming the validity of the relationship for the bulk of the data.

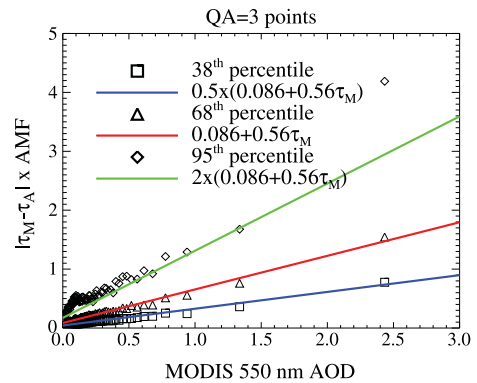


Figure 4. The 38th (squares), 68th (triangles), and 95th (diamonds) percentiles of absolute error on MODIS (Aqua) Deep Blue AOD at 550 nm scaled by geometric AMF, as a function of MODIS AOD. The red line is the least squares linear fit of the 68th percentiles, the blue line half this linear fit, and the green line twice this linear fit. Only QA = 3 MODIS data are used.

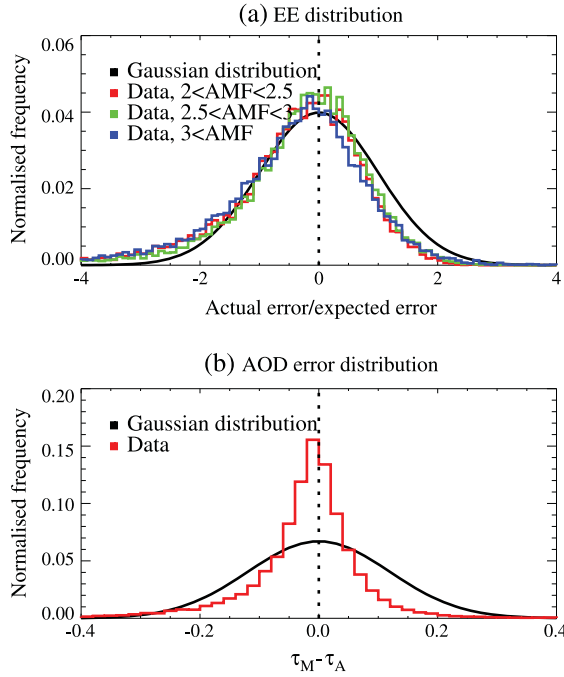


Figure 5. Comparison between observed AOD error distributions and theoretical Gaussian distributions, for (a) the EE, split into different AMF ranges, and (b) the raw AOD error. The dashed vertical line indicates zero error. Only QA = 3 MODIS data are used.

[20] Further, Figure 5 shows histograms of both the ratio of actual retrieval error to EE, i.e., $(\tau_M - \tau_A)/EE$, and the AOD error itself $(\tau_M - \tau_A)$. These are compared to idealized Gaussian distributions, i.e., those with a mean of zero and variance of one in the former case, namely $\mathcal{N}(0, 1)$, or a mean of zero and variance of the AOD error, namely $\mathcal{N}(0, \text{Var}(\tau_M - \tau_A))$, in the latter case. For the matched data here, $\text{Var}(\tau_M - \tau_A) = 0.014$ (i.e., the standard deviation of $\tau_M - \tau_A$ is about 0.12). Bin sizes are 0.1 for the EE histogram and 0.02 for the AOD histogram; in the former case, data are also split into three AMF ranges with approximately equal total populations. The slight negative bias and longer negative than positive tail are observed in both cases, although framing the error in terms of EE results in a much closer correspondence with the theoretical Gaussian distribution, for all three AMF groupings. This further shows the basic validity of equation (2) for this dataset although provides a reminder that this simple formulation is imperfect. Although the histogram shown in terms of $\tau_M - \tau_A$ also appears approximately Gaussian, it is somewhat narrower than the distribution formed by taking $\mathcal{N}(0, \text{Var}(\tau_M - \tau_A))$ would suggest, due to the presence of large-error outliers from high-AOD cases (consistent with Figures 2 and 3).

[21] The aforementioned linear formulations for EE for each QA value show that QA = 3 data have the lowest absolute uncertainties, although results for QA = 2 are numerically similar. It is generally therefore recommended, for most applications, that only QA = 2 and QA = 3 retrievals should be used. Part of the rationale for providing such retrieval-level uncertainty estimates is to aid users in making decisions on data usability for their particular application. QA = 1 uncertainties are larger. Additionally, one of the main

reasons for assignment of QA = 1 is suspicion of residual cloud contamination. The true extent of this may be understated by this type of AERONET comparison, because for a valid matchup, both AERONET and MODIS data must be flagged as cloud-free (so the data are, in a sense, doubly cloud-screened). Therefore, the risk of cloud contamination may be greater “in the wild” than in the subset of matched MODIS/AERONET data. For this reason, only QA ≥ 2 Deep Blue retrievals are propagated through into Level 3 (daily/monthly composites) products in C6, and QA = 1 data are not generally recommended for use.

3.3. Global and Regional Performance

[22] Maps of validation statistics for QA = 3 data are shown in Figure 6, and global and regional statistics (using the regions defined in Sayer *et al.* [2012b] with boundaries shown in Figure 1) are given in Table 1. For reference, these also present the fraction of points falling within the approximate C5 uncertainty of $0.05 + 0.2\tau_A$, to give an additional metric perhaps more familiar to users. These reveal that, on a global basis, the median bias between Deep Blue data and AERONET is small (-0.012 for QA = 3), and as the assigned QA decreases, the bias becomes larger and positive (0.03) and the data become less well-correlated (from $R=0.92$ for QA=3 to $R=0.88$ for QA=1), although the level of correspondence remains high. Biases at individual sites are typically small (magnitude 0.025 or less at most sites) and more likely to be negative than positive.

[23] There is regional variation in retrieval performance, likely due to differing regional uncertainties in assumed surface reflectance, aerosol microphysical properties, and performance of cloud-screening tests. Some regions (notably Eastern North America and Eurasia) have significantly more than 68% of retrievals within the expected confidence envelope, while the eastern Asian and northern African regions have less. The data in these latter regions remain well-correlated, indicating Deep Blue tracks spatiotemporal variability in AOD well, but with a low bias, which perhaps pushes the retrieval outside of the confidence envelope more frequently than expected.

[24] The site of Saada (Morocco) contributes strongly to the comparatively poor performance overall in the North Africa/Middle East region, as it contains a large number of matchups yet shows a negative bias against AERONET. Negative biases could indicate insufficient absorption in aerosol properties assumed in these regions, overestimates of surface reflectance, or perhaps issues associated with subretrieval heterogeneity or over-zealous cloud-flagging of some plumes. This particular site is also in horizontally and vertically heterogeneous terrain. Equally important to note is that not all sites in these regions perform poorly: for example, Banizoumbou and Mongu in Africa both perform in excess of expectations.

[25] Oceania shows the lowest correlation between the two datasets, although here the typical AOD and its variability are low, close to the magnitude of the Deep Blue uncertainty, which will decrease the correlation even if the data often agree within EE.

[26] The variability in performance across the regions indicates potential foci for future algorithmic developments, or development of a more advanced error model. Unfortunately, this was not possible within the timespan available

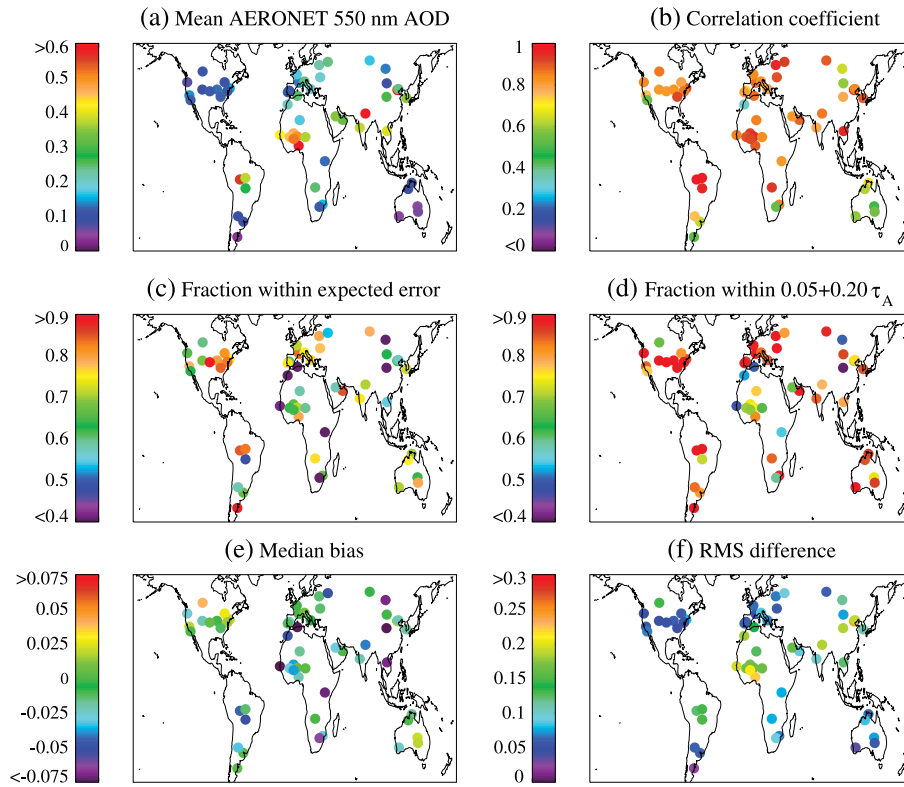


Figure 6. Global variability of validation statistics for AOD at 550 nm. Shown, for each site, are (a) the mean AERONET AOD at 550 nm, (b) Pearson’s linear correlation coefficient between the matched MODIS and AERONET data, (c and d) the fraction of matchups within the EE and $0.05+0.2\tau_A$ of AERONET, respectively, (e) the median (MODIS-AERONET) bias, and (f) the RMS difference against AERONET. Only QA = 3 MODIS data are used.

for the C6 reprocessing, which is an extensive and multidisciplinary effort. It is worth noting that in the poorer-performing regions of Africa and eastern Asia, the C5 guideline uncertainty (68% within $0.05+0.2\tau_A$) is approximately met or exceeded. Some caution is advised when interpreting these results, as some regions contain only a few sites (both due to limitations in processing test data around AERONET sites, and limited spatial coverage of AERONET sites themselves in some regions), and as such results may not be representative of all regimes encountered

over a region. These aspects may be assessed in more detail once the full C6 dataset is available, and through analysis of past, current, and future field campaigns.

3.4. Performance Relative to C5 Data

[27] Comparison of global validation statistics of C5 and C6 is not directly a useful measure of change in algorithm performance, as the spatial coverage of Deep Blue retrievals is significantly larger in C6, because only bright surfaces were processed by the C5 version of the algorithm

Table 1. Global and Regional Statistics of Comparison Between MODIS Deep Blue and AERONET AOD at 550 nm^a

Region	Number of Sites	Number of Matches	R	Median Bias	Fraction Within EE	Fraction Within $0.05+0.2\tau_A$
Global, QA = 3	60	41297	0.92	−0.012	0.68	0.80
Global, QA = 2	60	18200	0.90	0.0049	0.67	0.69
Global, QA = 1	60	38275	0.88	0.030	0.68	0.63
Eastern North America (ENA)	7	4155	0.86	0.0094	0.82	0.89
Western North America (WNA)	5	4101	0.78	−0.0076	0.70	0.84
Central/South America (CSA)	6	3032	0.96	−0.016	0.69	0.85
Eurasia (EUR)	10	8025	0.89	−0.013	0.75	0.89
North Africa/Middle East (NAME)	12	9938	0.90	−0.036	0.58	0.67
Southern Africa (SA)	4	1958	0.90	−0.026	0.66	0.82
Northeast Asia (NEA)	6	4193	0.93	−0.011	0.63	0.75
Southeast Asia (SEA)	5	2422	0.88	−0.057	0.60	0.72
Oceania (OCE)	5	3473	0.54	0.0021	0.72	0.83

^a R is Pearson’s linear correlation coefficient. The bias is the median offset between the datasets, defined such that positive values indicate MODIS AOD is larger. Regional statistics are shown for QA = 3 data only, for regions as shown in Figure 1.

Table 2. Comparative Assessment of C5 and C6 Deep Blue Performance for Selected AERONET Sites^a

QA Value	Number of Matches		<i>R</i>		Fraction Within EE		Fraction Within $0.05+0.2\tau_A$	
	C5	C6	C5	C6	C5	C6	C5	C6
Agoufou, Mali (15.35°N, 1.48°W)								
3	681	934	0.87	0.92	0.58	0.66	0.62	0.74
2	889	409	0.84	0.90	0.62	0.73	0.61	0.69
1	917	679	0.83	0.87	0.68	0.80	0.54	0.69
Banizoumbou, Niger (13.54°N, 2.66°E)								
3	447	1282	0.93	0.94	0.72	0.71	0.76	0.75
2	954	536	0.88	0.91	0.65	0.76	0.62	0.69
1	1369	861	0.87	0.88	0.69	0.80	0.49	0.63
Beijing, China (39.98°N, 116.38°E)								
3	849	1664	0.89	0.94	0.39	0.64	0.56	0.76
2	981	937	0.88	0.91	0.41	0.59	0.51	0.62
1	1252	1382	0.85	0.91	0.43	0.67	0.48	0.63
BSRN BAO Boulder, USA (40.05°N, 105.00°W)								
3	223	937	0.84	0.82	0.61	0.69	0.78	0.91
2	632	469	0.77	0.72	0.63	0.80	0.78	0.86
1	1252	1099	0.64	0.72	0.54	0.76	0.70	0.84
Fresno, USA (36.78°N, 119.77°W)								
3	533	1441	0.79	0.75	0.52	0.78	0.71	0.84
2	1026	509	0.73	0.61	0.65	0.72	0.76	0.66
1	1195	581	0.74	0.54	0.70	0.61	0.76	0.53
Hamim, UAE (22.97°N, 54.30°E)								
3	348	490	0.89	0.88	0.72	0.85	0.83	0.89
2	488	79	0.89	0.90	0.86	0.80	0.87	0.75
1	50	102	0.85	0.87	0.86	0.81	0.72	0.61
Kanpur, India (26.51°N, 80.23°E)								
3	762	989	0.76	0.86	0.41	0.71	0.44	0.78
2	769	530	0.74	0.85	0.43	0.72	0.41	0.70
1	690	677	0.70	0.79	0.46	0.79	0.35	0.66
Mongu, Zambia (15.25°S, 23.15°E)								
3	374	814	0.77	0.93	0.20	0.76	0.24	0.84
2	334	293	0.77	0.93	0.22	0.86	0.23	0.81
1	417	571	0.76	0.88	0.29	0.76	0.33	0.61
Solar Village, Saudi Arabia (24.91°N, 46.40°E)								
3	1514	1843	0.79	0.82	0.56	0.58	0.67	0.68
2	1527	1010	0.68	0.69	0.53	0.58	0.56	0.59
1	385	1070	0.57	0.62	0.54	0.64	0.39	0.50
Tinga Tingana, Australia (28.98°S, 139.99°E)								
3	604	839	0.67	0.56	0.58	0.79	0.74	0.86
2	682	164	0.60	0.49	0.60	0.58	0.64	0.59
1	625	314	0.54	0.40	0.54	0.46	0.55	0.46

^a*R* is Pearson's linear correlation coefficient.

[Hsu *et al.*, 2004, 2006]. For this reason, performance is compared using a selection of AERONET sites with large data volumes in both the C5 and C6 datasets, covering a range of aerosol/surface conditions. Statistics of the comparison at these sites are presented in Table 2; C5 data were collocated with AERONET in the same manner to C6 for this comparison. Strictly, the C5 products used here belong to Collection 5.1, which was an intermediate reprocessing of C5 with minor updates and bugfixes, although the term C5 is used for simplicity.

[28] At Agoufou, Beijing, Kanpur, Mongu, and Solar Village, performance for C6 is better for all statistics than for C5. The data volume is also significantly increased for Beijing and Mongu, which, along with Kanpur, show very strong improvements in retrieval performance. At Banizoumbou, Boulder, Fresno, Hamim, and Tinga Tingana, C6 performance is similar to C5 for some statistics and better for others; additionally, at the first three of these sites, the data volume is increased by a factor of 2.6–4.2. For those cases where a statistical metric is poorer for C6, the

difference is typically small, and the metric is often the correlation coefficient. The typical AERONET AOD at 550 nm is ~ 0.15 or less for Boulder and Fresno, and < 0.1 for Tinga Tingana, and variability at these sites is similar to the Deep Blue uncertainty. Combined with the increase in data volume, this small decrease of correlation for QA = 3 points at these sites is not thought to be a major issue.

[29] If only AERONET data points where both C5 and C6 provide QA = 3 retrievals are considered, the relative performance is similar. This shows that the overall improvement in C6 as compared to C5 is not due only to those additional AERONET data points for which C6 provides a QA = 3 retrieval but is also present for those situations in which the C5 algorithm retrieved.

[30] The common thread among these sites is that the C6 version of Deep Blue provides a greater number of comparisons with QA = 3 data than C5 and that the correspondence with AERONET (in terms of fraction of points matching within some uncertainty confidence envelope) is approximately the same or better. This shows that the C6 algorithm provides an improvement over C5 both in terms of data volume and retrieval quality. These improvements are observed for all QA flag values, suggesting they are the result of algorithmic and calibration updates rather than solely more intelligent QA assignment. Statistics for QA = 3 retrievals generally show closer agreement with AERONET than for QA = 2 or QA = 1, for both C5 and C6, confirming that the QA flags provide information relevant to the performance of the retrievals.

3.5. Effect of Radiometric Calibration Updates

[31] As well as algorithmic changes, the C6 Deep Blue product incorporates improvements to MODIS' radiometric calibration since C5. The efforts of the MODIS Characterization Support Team have been focused toward maintaining a high-quality and temporally stable calibration of the MODIS sensors. Here, the effects of calibration updates on the Deep Blue dataset are examined by taking C5 and C6 calibrated radiances (so-called Level 1b data) and applying the C6 algorithm identically to them both. Comparing the validation statistics against AERONET from the years 2002 to 2012 then provides insight into how these changes propagate into the Deep Blue dataset.

[32] This exercise is restricted to a handful of AERONET sites with a long time series (Alta Floresta, Banizoumbou, Beijing, GSFC, MD Science Center, Minsk, Moldova, Mongu, Moscow, and Solar Village); the validation protocol is as described previously, with the added constraint that only matchups where both C5 and C6 calibration coefficients provide a valid QA = 3 matchup are considered. These constraints help to ensure that the expected performance through the mission is fairly homogeneous (as the same AERONET sites are used for the period) and that the resulting matchups are directly comparable (as each AERONET data point contains Deep Blue data with each of the C5 and C6 radiometric calibrations applied). One important point to note is that AERONET level 2 processing runs typically several months to a year or two behind real-time, dependent on the maintenance schedules for the individual site in question, so the 2011 and 2012 records are incomplete at present. Additionally, 2002 is less well-sampled because it was an incomplete year for both Aqua and some of the AERONET sites.

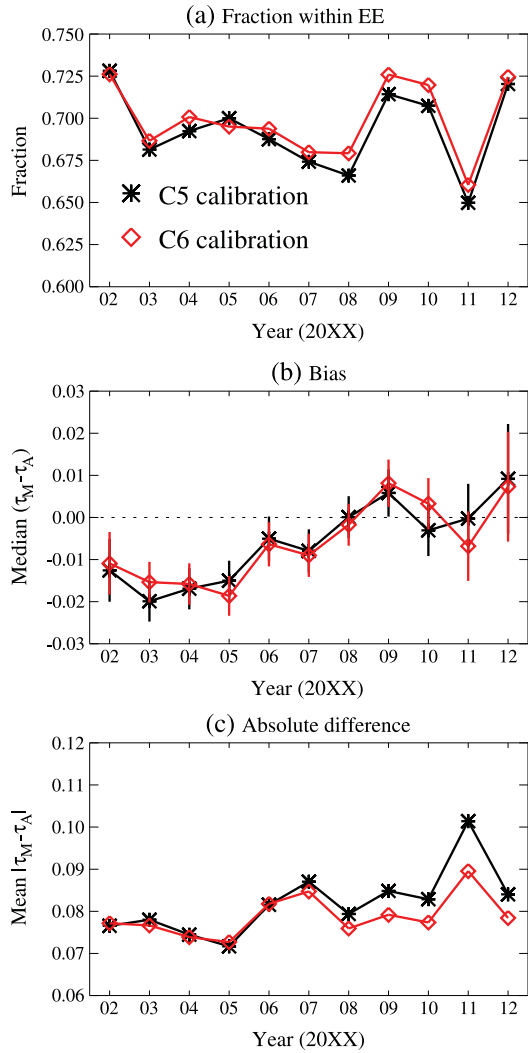


Figure 7. Temporal variability (2002–2012) of validation statistics for AOD at 550 nm, for long-term AERONET sites. In each panel, black asterisks indicate C5 and red diamonds C6 radiometric calibration coefficients. (a) The fraction of matchups in agreement within the expected error, (b) the median bias (vertical lines indicate the standard error on the median), and (c) the mean absolute MODIS–AERONET difference. Only QA = 3 MODIS data are used.

[33] Statistics of the resulting comparison (9316 points total) are presented in Figure 7. The differences between the two sets of retrievals are not large (generally, the C6 calibration gives <1% more matching AERONET within EE, and a smaller mean absolute AOD difference by <0.01) and manifest most strongly in recent years (2008 onward). Interestingly, both sets show a tendency for negative bias in the earlier years of the Aqua mission and positive bias in the later (with the change in AOD bias through the decade ~ 0.03), although the standard errors on the median are large relative to the observed biases, limiting the strength of conclusions which should be drawn. The similar trend in bias observed between both calibrations suggests it may be related to either changes to aerosol composition or surface reflectance which are not captured by the retrieval, or could be a coincidental factor related to different distributions of

AOD in different years. The main result from this analysis is to show that using the C6 calibration provides improved AOD retrievals compared to the C5 calibration, although the effect of algorithmic changes, investigated in section 3.4, is larger. The effect of calibration changes on Terra is expected to be larger than for Aqua.

4. Summary and Conclusions

[34] The Deep Blue algorithm has undergone several refinements for the upcoming Collection 6 MODIS dataset, although the general principles of the algorithm [Hsu *et al.*, 2004, 2006] remain the same. The main update is the extension of coverage to include vegetated surfaces, but there have additionally been modifications to radiometric calibration, aerosol/surface models, cloud screening, and quality flags.

[35] Validation against AERONET suggests that, on a global basis, the EE confidence envelope on Deep Blue C6 AOD at 550 nm from MODIS Aqua can be stated as $\pm(0.086 + 0.56\tau_M)/(\mu_0 + \mu)$ for retrievals with the highest quality flag (QA = 3). With a typical AMF of 2.8, this reduces to $\pm(0.03 + 0.2\tau_M)$, which is comparable to stated uncertainties on other well-used state-of-the-art datasets such as MODIS C5/C6 Dark Target ($0.05 + 0.15\tau_A$) [Levy *et al.*, 2010, 2013] or MISR (greater of 0.05 or $0.2\tau_A$) [Kahn *et al.*, 2010] aerosol products under typical aerosol loadings (midvisible AOD ~ 0.05 – 0.5). It is an improvement on the previous application of Deep Blue in MODIS C5 and to SeaWiFS (uncertainty $0.05 + 0.2\tau_A$) [Sayer *et al.*, 2012b]. Retrievals assigned QA = 2 have only slightly larger uncertainty and so should also be suitable for use in most applications. Comparisons to C5 Deep Blue data at select sites reveal an increase both in the data volume of retrievals with the highest quality assurance flag, and also the level of agreement with AERONET data. In general, performance appears better over vegetated regions and worse over more complicated environments (in terms of heterogeneous surface cover, cloudiness, and aerosol composition) as found in Africa and eastern Asia—although there is a dearth of validation data for parts of these regions. In these latter regions, it appears that the aforementioned C5 uncertainty estimate may be an appropriate metric. Looking to the future, additional validation activities, field campaigns, and intercomparison with other satellite aerosol datasets will be useful to guide further refinement to the Deep Blue algorithm and improve understanding of retrieval performance in different conditions.

[36] The uncertainty estimates derived herein are provided on an individual-retrieval basis in the Collection 6 Deep Blue dataset, in a data field named `Deep_Blue_Aerosol_Optical_Depth_550_Land_Estimated_Uncertainty`. Uncertainty estimates for MODIS Terra are expected to be similar, will be determined using the same method once the required source data are available, and provided within the corresponding Terra data products.

[37] The C6 Deep Blue record will continue to fill gaps in the Dark Target MODIS AOD dataset over bright reflecting surfaces such as deserts, providing quantitatively useful AOD data in these regions. It will now also provide a second, complementary dataset in those land regions where Dark Target does retrieve AOD. The only remaining

“holes” over cloud-free surfaces in the MxD04/MxD08 products are therefore snow-covered areas, in which neither algorithm performs a retrieval. The current application to SeaWiFS and the MODIS sensors aboard the Terra and Aqua platforms provides a continuous multisensor Deep Blue dataset from late 1997 onward. The potential future application to measurements from the Visible and Infrared Imaging Radiometer Suite (VIIRS) aboard the Suomi-National Polar-orbiting Partnership (NPP) satellite (launched late 2011), and similar successor instruments aboard satellites in the United States Joint Polar Satellite System, would eventually provide a multidecadal record of AOD, processed with a single algorithm, with important applications in monitoring global change.

[38] **Acknowledgments.** The authors gratefully acknowledge the AERONET team and site PIs (N. Al-Abbadi, P. Artaxo, F. Baret, C. J. Bruegge, M. Caughey, A. Chaikovsky, B. Chatenet, H.-B. Chen, N. Chubarova, S. Corradini, E. Cuevas-Agullo, P. C. S. Devara, B. Duchemin, M. L. C. Fernandez, P. Goloub, B. Gross, B. N. Holben, J. Huang, Y. J. Kim, K.-N. Liou, B. McArthur, D. Meyer, R. Mitchell, J.-P. Morel, B. Mougnot, N. T. O'Neill, M. Panchenko, M. R. Perrone, S. Piketh, R. T. Pinker, S. Pugnaghi, J. L. Rajot, A. M. Silva, R. P. Singh, D. Tanré, S. N. Tripathi, J. R. Vande Castle, P. Wang, X. Xia, and W. Zhang) for the creation and stewardship of the Sun photometer data records. The MODIS Characterization Support Team team are thanked for their extensive efforts in maintaining the high radiometric quality of MODIS data. We thank two anonymous reviewers for their insightful comments and suggestions.

References

- Banks, J. R., and H. E. Brindley (2013), Evaluation of MSG-EVIRI mineral dust retrieval products over North Africa and the Middle East, *Remote Sens. Environ.*, **128**, 58–73, doi:10.1016/j.rse.2012.07.017.
- Brindley, H. E., and J. E. Russell (2009), An assessment of Saharan dust loading and the corresponding cloud-free longwave direct radiative effect from geostationary satellite observations, *J. Geophys. Res.*, **114**, D23201, doi:10.1029/2008JD011635.
- Carboni, E., et al. (2012), Intercomparison of desert dust optical depth from satellite measurements, *Atmos. Meas. Tech.*, **5**, 1973–2002, doi:10.5194/amt-5-1973-2012.
- DeSouza-Machado, S. G., et al. (2010), Infrared retrievals of dust using AIRS: Comparisons of optical depths and heights derived for a North African dust storm to other collocated EOS A-Train and surface observations, *J. Geophys. Res.*, **115**, D15201, doi:10.1029/2009JD012842.
- Draxler, R. R., P. Ginoux, and A. F. Stein (2010), An empirically derived emission algorithm for wind-blown dust, *J. Geophys. Res.*, **115**, D16212, doi:10.1029/2009JD013167.
- Ginoux, P., J. M. Prospero, T. E. Gill, N. C. Hsu, and M. Zhao (2012), Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products, *Rev. Geophys.*, **50**, RG3005, doi:10.1029/2012RG000388.
- Hasekamp, O. P., and J. Landgraf (2007), Retrieval of aerosol properties over land surfaces: Capabilities of multi-viewing-angle intensity and polarization measurements, *Appl. Opt.*, **46**(16), 3332–3344, doi:10.1364/AO.46.003332.
- Holben, B. N., et al. (1998), 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.
- Hsu, N. C., S.-C. Tsay, M. D. King, and J. R. Herman (2004), Aerosol properties over bright-reflecting source regions, *IEEE Trans. Geosci. Remote Sens.*, **42**(3), 557–569, doi:10.1109/TGRS.2004.824067.
- Hsu, N. C., S.-C. Tsay, M. D. King, and J. R. Herman (2006), Deep Blue retrievals of Asian aerosol properties during ACE-Asia, *IEEE Trans. Geosci. Remote Sens.*, **44**(11), 3180–3195, doi:10.1109/TGRS.2006.879540.
- Hyer, E. H., J. S. Reid, and J. Zhang (2011), An over-land aerosol optical depth dataset 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.
- Ichoku, C., D. A. Chu, S. Mattoo, Y. J. Kaufman, L. A. Remer, D. Tanré, I. Slutsker, and B. N. Holben (2002), A spatio-temporal approach for global validation and analysis of MODIS aerosol products, *Geophys. Res. Lett.*, **29**, 12, doi:10.1029/2001GL013206.
- Kahn, R. A., B. J. Gaitley, M. J. Garay, D. J. Diner, T. F. Eck, A. Smirnov, and B. N. Holben (2010), Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network, *J. Geophys. Res.*, **115**, D23209, doi:10.1029/2010JD014601.
- Karimi, N., A. Moridnejad, S. Golian, J. Samani, D. Karimi, and S. Javadi (2012), Comparison of dust source identification techniques over land in the Middle East region using MODIS data, *Can. J. Remote Sensing*, **38**(5), 586–599, doi:10.5589/m12-048.
- Kokhanovsky, A. A., et al. (2010), The determination of spectral aerosol optical thickness from satellites: An inter-comparison of algorithms using synthetic backscattered solar light characteristics, *Atmos. Meas. Tech.*, **3**, 909–932, doi:10.5194/amt-3-909-2010.
- Laurent, B., I. Tegen, B. Heinold, K. Schepanski, B. Weinzierl, and M. Esselborn (2010), A model study of Saharan dust emissions and distributions during the SAMUM-1 campaign, *J. Geophys. Res.*, **115**, D21210, doi:10.1029/2009JD012995.
- Levy, R. C., L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman (2007), Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance, *J. Geophys. Res.*, **112**, D13211, doi:10.1029/2006JD007811.
- Levy, R. C., L. A. Remer, R. G. Kleidman, S. Mattoo, C. Ichoku, R. Kahn, and T. F. Eck (2010), Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos. Chem. Phys.*, **10**, 103,999–10,420, doi:10.5194/acp-10-10399-2010.
- Levy, R. C., S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, and N. C. Hsu (2013), The Collection 6 MODIS aerosol products over land and ocean, *Atmos. Meas. Tech. Discuss.*, **6**, 159–259, doi:10.5194/amt-d-6-159-2013.
- Martonchik, J. V., D. J. Diner, R. A. Kahn, T. P. Ackerman, M. M. Verstraete, B. Pinty, and H. R. Gordon (1998), Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging, *IEEE Trans. Geosci. Remote Sens.*, **36**, 4, doi:10.1109/36.701027.
- North, P. R. J. (2002), Estimation of aerosol opacity and land surface bidirectional reflectance from ATSR-2 dual-angle imagery: Operational method and validation, *J. Geophys. Res.*, **107**(D12), 4149, doi:10.1029/2000JD000207.
- Sayer, A. M., G. E. Thomas, R. G. Grainger, E. Carboni, C. Poulsen, and R. Siddans (2012a), Use of MODIS-derived surface reflectance data in the ORAC-AATSR aerosol retrieval algorithm: Impact of differences between sensor spectral response functions, *Rem. Sens. Environ.*, **116**, 177.
- Sayer, A. M., N. C. Hsu, C. Bettenhausen, M.-J. Jeong, B. N. Holben, and J. Zhang (2012b), Global and regional evaluation of over-land spectral aerosol optical depth retrievals from SeaWiFS, *Atmos. Meas. Tech.*, **5**, 1761–1778, doi:10.5194/amt-5-1761-2012.
- Schepanski, K., I. Tegen, and A. Macke (2012), Comparison of satellite based observations of Saharan dust source areas, *Remote Sens. Environ.*, **123**, 90–97, doi:10.1016/j.rse.2012.03.019.
- Seidel, F. C., and C. Popp (2012), Critical surface albedo and its implications to aerosol remote sensing, *Atmos. Meas. Tech.*, **5**, 1653–1665, doi:10.5194/amt-5-1653-2012.
- Shi, Y., J. Zhang, J. S. Reid, E. J. Hyer, and N. C. Hsu (2013), Critical evaluation of the MODIS Deep Blue aerosol optical depth product for data assimilation over North Africa, *Atmos. Meas. Tech.*, **6**, 949–969, doi:10.5194/amt-6-949-2013.
- Smirnov, A., B. N. Holben, T. F. Eck, O. Dubovik, and I. Slutsker (2000), Cloud-screening and quality control algorithms for the AERONET database, *Remote Sens. Environ.*, **73**(3), 337–349.
- Wagner, F., and A. M. Silva (2008), Some considerations about Ångström exponent distributions, *Atmos. Chem. Phys.*, **8**, 481–489, doi:10.5194/acp-8-481-2008.
- Wagner, S. C., Y. M. Govaerts, and A. Lattanzio (2010), Joint retrieval of surface reflectance and aerosol optical depth from MSG/SEVIRI observations with an optimal estimation approach: 2. Implementation and evaluation, *J. Geophys. Res.*, **115**, D02204, doi:10.1029/2009JD011780.
- Wang, J., X. Xu, D. K. Henze, J. Zeng, Q. Ji, S.-C. Tsay, and J. Huang (2012), Top-down estimate of dust emissions through integration of MODIS and MISR aerosol retrievals with the GEOS-Chem adjoint model, *Geophys. Res. Lett.*, **39**, L08802, doi:10.1029/2012GL051136.
- Wanner, W., X. Li, and A. H. Strahler (1995), On the derivation of kernels for kernel-driven models of bidirectional reflectance, *J. Geophys. Res.*, **100**, 21,077–21,089, doi:10.1029/95JD02371.
- Zhang, J., and J. S. Reid (2006), MODIS aerosol product analysis for data assimilation: Assessment of over-ocean level 2 aerosol optical thickness retrievals, *J. Geophys. Res.*, **111**, D22207, doi:10.1029/2005JD006898.