Response to "Toward unified satellite climatology of aerosol properties. 3. MODIS versus MISR versus AERONET"

Ralph A. Kahn[1], Michael J. Garay[2], David L. Nelson[2], Robert C. Levy[3], Michael A. Bull[3], David J. Diner[3], John V. Martonchik[3], Earl G. Hansen[3], Lorraine A. Remer[1], and Didier Tanre'[5]

[1]Laboratory for Atmospheres, NASA Goddard Space Flight Center
Greenbelt, MD 20771

[2]Raytheon Company
299 N. Euclid Ave., Suite 500
Pasadena, CA 91101

[3]Science Systems and Applications, Inc., Lanham, MD 20706

[4]Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena CA 91109

[5]Universite' des Sciences et Technologies de Lille
Villeneuve d'Ascq., France

Abstract

A recent paper by *Mishchenko et al.* compares near-coincident MISR, MODIS, and AERONET aerosol optical depth (AOD) products, and reports much poorer agreement than that obtained by the instrument teams and others. We trace the reasons for the discrepancies primarily to differences in (1) the treatment of outliers, (2) the application of absolute vs. relative criteria for testing agreement, and (3) the ways in which seasonally varying spatial distributions of coincident retrievals are taken into account.

1. Introduction

During the past decade, the research community has found many productive ways to use the Multi-angle Imaging SpectroRadiometer (MISR) and MODerate resolution Imaging Spectroradiometer (MODIS) aerosol products. MISR and MODIS, both of which fly aboard the NASA Earth Observing System's Terra spacecraft, represent a significant advance over the previous generation of space-based aerosol instruments. Relatively high spatial resolution imaging, calibration accuracy and radiometric stability, along with an increased number of spectral bands for MODIS and the combination of spectral bands and multiple view angles from MISR, have led to more robust aerosol optical depth (AOD) retrievals over both water and land, with less restrictive algorithmic assumptions [1-5]. In addition, the MODIS algorithm derives coarse vs. fine-mode ratio over water, whereas MISR can distinguish about a dozen aerosol air mass types, under favorable retrieval conditions, based on particle size, shape, and single-scattering albedo constraints. And unlike most

remote-sensing algorithms that assume aerosol properties based on seasonally and/or geographically fixed prescriptions, MISR AOD retrievals are performed self-consistently, using retrieved aerosol types.

Critical to the application of these satellite products is validation, which entails establishing uncertainties, assessing strengths and limitations, and reporting overall data quality. The MISR and MODIS instrument teams, as well as many other research groups, have performed such evaluations using a variety of methods, ranging from theoretical sensitivity studies to comparisons with other satellite and suborbital data sets [e.g., 2, 6-23]. Over the past ten years, the exchange of data, information, and ideas represented by this body of work has yielded better characterization of existing products, as well as algorithm upgrades.

Although operating from the same Terra satellite, MODIS and MISR offer different views of aerosols, clouds, and surface radiation. Even when viewing nearly coincident target scenes, these instruments are sensitive to different properties of the scene, due to differences in their observational approaches. Aside from global retrieval coverage differences [e.g., 12], MISR and MODIS do not observe, or perform aerosol retrievals on, top-of-atmosphere radiances from exactly the same locations within their 17.6 km and 10 km retrieval regions, respectively. Where variability occurs on spatial scales of ~10 km or less, this leads to sampling differences even for "collocated" observations [11]. MISR and MODIS data are often complementary, so differences do not always represent algorithm issues; sometimes they offer a deeper understanding of realistic natural scenes. For example, case studies that take detailed account of instrument sampling have shown where MISR and MODIS–retrieved AOD differences are due to aerosol spatial variability [9, 11].

Broad statistical comparisons among MISR, MODIS, AERONET surface-based, and aircraft aerosol measurements under a wide range of environmental conditions have been complemented by representative case-study analyses by many different groups, often aimed at identifying the underlying causes of discrepancies [e.g., 2, 6, 8-10, 12, 14-16, 18-21]. Explanations for the most prominent differences between the MISR Standard aerosol products and validation data are given in several of the papers cited above, making use of other studies where individual effects for dust [13], spherical absorbing and non-absorbing particles [7], thin cirrus [17], and algorithmic issues [11] are explored for representative cases with the MISR Research aerosol retrieval algorithm. Detailed analysis of MODIS issues is presented in [2, 4, 14, 20, 21, 24-27].

In their recent paper, *Mishchenko et al.* [28] evaluate the MISR and MODIS AOD products in light of near-coincident AERONET observations, and report that their results are much less favorable than those presented in the studies listed above. Given this disparity, the current response addresses the following question: How do *Mishchenko et al.* [28] reach conclusions that seem to contradict the published work of others?

Analysis shows that the largest contribution to the disparate results comes from differences in the treatment of *outliers*; other factors are involved as well, such as the seasonally varying *spatial distribution of coincident MISR-MODIS AOD retrievals*, and differences in the application of *absolute vs. relative criteria* for testing agreement. We review the implications of these below, and clarify several other technical points.

*****

2. ***The Influence of Outliers.*** The treatment of outliers in *Mishchenko et al.* [28] has a major effect on the overall magnitude of the errors they report. In [28, Figure 1c], a normalized standard deviation metric (RSTD) is used to assess the MISR and MODIS time series, separately for ocean and land. The RSTD, in this case a measure of MISR-MODIS relative discrepancy, is calculated from the root-sum-square of the MISR-MODIS AOD difference. *This metric is especially sensitive to the largest outliers*, which can be caused by retrieval error or by sampling differences, as mentioned above. In addition, although the sampling differences tend to produce outliers of both signs, AOD retrieval errors are often systematic rather than random [10, 11], and most are localized geographically within regions where certain algorithmic assumptions are violated [12, Figure 6; 29, Section 3.4]. As a statistical tool, the RSTD is most appropriate for assessing random error about an expected value; interpretation of RSTD values under other circumstances is not always straightforward.

Despite generally good agreement between MISR and MODIS in the AOD time-series [28, Figure 1a], high percentage-differences appear in the RSTD plots [28, Figure 1c], averaging around 40% over ocean and near 90% over land. The high RSTD values are dominated by a small fraction of cases having large AOD differences, which are then squared in the summation of the RSTD statistic of [28]:

$$RSTD = \frac{1}{Y} \sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{(N-1)}} \tag{1}$$

where x_i is the MISR-MODIS AOD deviation, the over-bar indicates the monthly mean, N is the total number of coincident measurements for the month, and Y is half the sum of the MISR and MODIS monthly mean AOD values. We demonstrate the contributions of outliers to the RSTD over land and ocean in two ways. First, we compare the RSTD values with a metric that replaces the squared deviation in the numerator of the RSTD with the average (mean) absolute deviation (AAD), and remove the square root, as follows:

$$AAD = \frac{1}{Y} \sum_{i=1}^{N} \frac{ABS[x_i - \bar{x}]}{N} \tag{2}$$

The AAD still weights larger outliers more heavily, but not to the same degree as the root-sum-square of the RSTD, so the difference between the AAD and RSTD statistics gives an indication of the influence of outliers [30]. Using as examples the monthly, global, coincident MISR and MODIS data for January and July 2006, the RSTD amounts to 32% and 37% over ocean, and 85% and 110% over land, respectively. The corresponding AAD values drop to 21% and 23% over ocean, and 62% and 61% over land (Table 1). The deviations are much higher over land, where large outliers are more probable, due to generally higher AOD values and greater spatial variability near sources, as well as greater differences in the way the MISR and MODIS retrievals treat the surface boundary condition.

RSTD is actually sensitive to the *distribution* of deviation magnitudes and in general, increases as the square root of the *number of points* in the sum [31], whereas it is often mistakenly interpreted as representing only the average error. So, for example, if all the deviations in a data set are equal, the RSTD equals the AAD, but as the distribution of

deviations becomes more uneven, RSTD increases, and it is maximal when all the deviation is concentrated in a single point. If the distribution of deviations were simply Gaussian, the RSTD would be higher than the AAD by a factor 0.77; the higher RSTD values shown in Table 1 are considerably greater than that over land, where outliers make especially large contributions to the RSTD.

A second view of the RSTD situation is given in Figure 1. Here we have arranged the absolute deviations in rank order, for ocean and land separately, and calculated the RSTD for progressively larger fractions of the total population. January, March, and July 2006 data are displayed. Note that when the last few percent of the populations are included, the RSTD values jump considerably – by as much as 40 percent over land. For about 80% or more of the ocean and land populations, the RSTD values are below 20% and 60% respectively, which correspond roughly to the AAD values reported in Table 1.

MISR-MODIS AOD comparisons can also be made using a metric that counts each event equally, such as confidence envelopes, thereby limiting the influence of outliers, rather than weighting each case by the magnitude of its deviation. To place the RSTD in further context, we calculated RSTDs comparing the unperturbed MISR AOD values with MISR AOD values perturbed by an amount comparable to the basic mid-visible AOD measurement uncertainty reported in [10]. Specifically, we generated a normal distribution of random numbers (R) with mean zero and standard deviation one, and scaled it to create perturbed MISR values, as:

$$MISR' = MISR + [0.05 + 0.05 \times R] \qquad (3)$$

We then calculated the RSTD using Equation 1, but comparing MISR with MISR' AODs (rather than MISR with MODIS). The numerator is just the standard deviation of the perturbations, whereas the denominator is half the sum of the monthly mean perturbed and unperturbed AOD values. The results are given in Table 1. (Note that the over-ocean RSTDs are slightly larger than the corresponding over-land values, because the monthly mean over-ocean AODs in the RSTD denominators are smaller). The 0.05 perturbations alone produce RSTD values from over 20% to almost 40%, bracketing the MISR-MODIS over-ocean RSTD values, and providing a baseline against which to assess MISR-MODIS comparisons if the RSTD metric is used.

More generally, the RSTD or AAD by themselves do not provide a clear picture of data product quality:

(1) Aside from weighting events by the magnitude of the deviation, or the deviation squared, interpretation of the AAD or RSTD values is complicated by the absence of an absolute criterion to avoid weighting too heavily observations near the lower limit of measurement sensitivity, as indicated by the calculation related to Equation (3), and discussed further in Section 3 below. The need for absolute in addition to relative criteria can be addressed with envelope statistics, or by calculating statistics from data stratified appropriately by AOD magnitude.

(2) The RSTD and AAD statistics do not distinguish between those differences in near-coincident MISR and MODIS AOD values due to differences in spatial sampling convolved with kilometer-scale AOD variability [11, 12], and differences caused by actual limitations in either or both retrieval algorithms. Sampling differences also have implications for the

way regionally representative values must be constructed from the measurements [e.g., 32].

(3) Importantly, deeper insight into the behavior of the underlying data, especially over land, comes from stratifying the observations according to different environmental conditions, rather than looking only at globally aggregated values [e.g., 10, 12, 29].

(4) *Validation* of either AOD product must rely on direct comparisons with ground-truth, stratified based on observational conditions [e.g., 10, 14, 20, 29] to separate the systematic differences between satellite products that different algorithmic assumptions can produce. Note that retrieved AOD error, when assessed relative to ground truth, is often smaller than the MISR-MODIS AOD deviations, again because of systematic differences in algorithm assumptions [11, 12].

The decision about which metrics best represent the MISR-MODIS relative results depends upon the way the data are to be used. For example, outliers tend to be localized geographically, and for applications where AOD values are averaged over large areas, it can be advantageous to avoid specific regions where known algorithm issues tend to produce outliers [e.g., 12]. Note that even in such regions, only a small fraction of retrievals is responsible for the large discrepancies. Again taking January and July 2006 as examples, about 1% and 0.6%, respectively, of mid-visible AOD for over-ocean MISR-MODIS coincidences have differences that exceed 0.125; over land, 10% and 6% of coincident-AOD differences exceed 0.2 for these two sample months [29]. Identification of the locations where outliers concentrate, and analysis of underlying causes, which range from unscreened bright coastal water or land surface, to missing mixtures of smoke and dust optical analogs in the algorithm climatologies, are given in detail elsewhere [11, 12, 29].

3. *AOD Confidence Envelopes.* Both the MISR and MODIS teams define confidence envelopes to assess statistical agreement between satellite and near-coincident AERONET sun photometer AOD results [12, 20]. For example, the MISR team typically reports the fraction of cases falling within the larger of (0.05 or 0.2 x AOD), and the larger of (0.03 or 0.1 x AOD), representing broader and more restrictive envelopes, respectively [10]. One attribute of the confidence envelope approach is that it *weights all events equally*, so a small fraction of outliers does not exert a disproportionately large influence on the comparison statistic. Another aspect of the envelopes used is that they apply a *combination of absolute and relative criteria*. This is important when using a single metric to assess agreement over the full range of AOD values.

In different plots, *Mishchenko et al.* [28] apply relative criteria only, or absolute criteria only, to the entire high-and-low-AOD MISR and MODIS time series. These choices contribute to the discrepancy patterns appearing in the plots.

3.1. *The need for an Absolute Criterion.* An *absolute criterion* (e.g., 0.05 or 0.03) is required at very low AOD, to account for measurement error that does not disappear as AOD approaches zero. All real observations have uncertainty associated with the limits of measurement sensitivity. So applying only a relative condition (such as RSTD) when making low-AOD comparisons can produce arbitrarily large *percentage* differences, even when the absolute differences between the satellite-retrieved and ground-based validation values are very small. For example, if AOD retrieval sensitivity is 0.03, for a case where the actual AOD is 0.01, the calculated *relative* error can exceed 200%, and yet the reported AOD can be well within reasonable measurement accuracy.

The effect of applying a relative-only criterion is illustrated by comparing the Envelope and Relative-only Percent-agreement metrics in Table 1. Over ocean, for example, where AOD frequently approaches the limit of retrieval sensitivity, the Relative-only percent agreement statistics are factors of two or three poorer. (Note that the specific confidence envelopes given in Table 1 are used for comparing MISR AOD directly with AERONET ground-truth. These envelope metrics produce low values when applied to MISR-MODIS AOD comparisons over land without reference ground-truth, due to systematic differences in algorithm assumptions [12, Section IIIC]. Quantitative assessment of the MISR and MODIS product uncertainties must rely on direct comparisons with ground-truth, which produce much higher envelope metric values [e.g., 4; 10].)

3.2. *The need for a Relative Criterion*. A *relative criterion* (e.g., 20% or 10% of AOD) is needed because typically, AOD variability in natural settings, as well as some components of retrieval uncertainty, scale as the AOD itself. For example, spatial sampling differences occur even for nearly collocated satellite data, and especially when comparisons are made with surface-based point measurements [9-11]. These sampling differences, convolved with AOD variability on spatial scales up to about ten kilometers, can produce retrieved AOD differences about as large as the AOD itself, even when no algorithm error is involved [e.g., 11, Section 3.2]. Such variability is also demonstrated by the observation that, away from sources and aerosol plumes, improved agreement is obtained when regional-average rather than single, minimally distant retrievals are compared with point data [10; 28, Figures 2a and 2b; 29]. A relative criterion is an appropriate way to perform AOD comparisons in these situations, provided the AOD is significantly above the limit of measurement sensitivity.

Figures 2a, 2b, and 4 of *Mishchenko et al.* [28] apply only absolute criteria to the full MISR and MODIS time series. They point out that the seasonal variability in fractional agreement correlates well with the MODIS fraction over ocean [28, Figure 2b], with peaks in northern winter, and is anti-correlated with the monthly AOD time series [28, Figure 2c]. (Note that for relative metrics, shown in [28] Figures 1c, 5b, and 6, higher values indicate poorer agreement, whereas for absolute metrics, plotted in Figures 2a, 2b, and 4, higher values indicate better agreement.) By applying only an absolute criterion, the variations in [28, Figures 2a and 2b] are skewed by discrepancies at the high-AOD end of the range of values, and the observed correlations indicate poorest agreement systematically when AOD is highest. But other factors, such as sampling, affect the coincident data time-series as well.

4. *Seasonal AOD Discrepancy Patterns, and Apparent RSTD Stability*. The seasonal behavior of the time-series plots in [28] is heavily influenced by the spatial distribution of MISR and MODIS retrieval coincidences. The coincident sampling pattern is mediated by the geographic location of continents, convolved with the seasonally varying distributions of solar illumination, cloud, sun glint, snow extent, and ice cover [12, Figure 4]. Seasonally varying coverage of specific regions where MISR-MODIS AOD discrepancies tend to occur [11, 12, 29] also contributes to the observed patterns. As such, the aggregated AOD discrepancy patterns shown in [28] cannot be interpreted simply in terms of satellite retrieval error.

For example, in the joint MISR-MODIS retrieval data set, observations of the northern mid-latitude ocean dominate global ocean coverage in January, those of the extra-tropical

oceans dominate coverage in March, and sun glint affecting MODIS precludes coincident retrievals over nearly all the northern mid-to-low-latitude ocean in July [12, Figure 4e]. Due to the peculiarities of this complex coincidence sampling pattern, "global" mean over-ocean AOD in the coincident data set is systematically higher in January than in July (Table 1; [28, Figure 1b]), despite typically higher global average over-ocean AOD in July due to contributions from northern hemisphere dust and smoke.

Over land, the patterns are yet more complex. Taking 2006 as an example, the number of MISR-MODIS coincidences is nearly the same in January and March, but doubles for July; AOD is nearly the same in January and July, but peaks in March (Table 1). Some aspects of MISR-MODIS coincidence coverage patterns are described elsewhere [e.g., 12], but it is beyond the scope of the current response to assess their impact on specific, globally aggregated metrics in detail.

In [28, Figures 1c and 6c], an envelope of seasonally repeating RSTD values appears, despite large fluctuations in monthly mean AOD values and variations in the number of AERONET sites involved. This is interpreted as a demonstration that the RSTD provides a robust metric for "100%" of AODs. However, the seasonal pattern of MISR-MODIS coincident retrievals, convolved with the seasonal patterns of high- and low-AOD values and locations where MISR-MODIS AOD outliers occur most often, create a regular pattern of RSTD values as well. It remains to be shown whether the repeatability of the RSTD has any deeper statistical significance.

5. *MISR Outliers and Data Stratification*. In [28, Figure 6c and Table 1], spikes in the MISR-AERONET AOD deviation time series, as compared with that for MODIS are highlighted, and it is argued that they are caused by MISR's inability to handle cases of large AOD over land, due to blurring of surface contrast features that the MISR algorithm uses to separate atmospheric from surface signals. MISR AOD underestimation over land in high-AOD situations has been documented in previous publications [9, 10, 12], and representative cases were analyzed in detail using the MISR Research aerosol retrieval algorithm [7]. To isolate places and times when algorithm assumptions are violated, we *stratified the satellite data* according to different environmental conditions that can affect the retrieval [10, 14, 29]. This is a key step toward identifying underlying causes that goes beyond looking only at global statistical aggregates. Based on this work, the MISR AOD underestimation tends to occur when mid-visible AOD exceeds about 0.4, and actually, several factors contribute. The MISR team is exploring options for improving the Standard algorithm in light of the factors producing the high-AOD underestimation [29].

Regarding [28, Figure 6c], only a relative criterion is applied to the full range of AOD, with multiple comparisons between MISR, MODIS, and AERONET. But unlike [28, Figure 1c], land and ocean are combined, and near-coincidence with AERONET is required for most of the Figure 6c plots, so the data set is apparently smaller, and the seasonal pattern is much noisier. It should also be mentioned that compared to MISR, there are between three- and over-five times more samples in the MODIS-AERONET data set, due to the much wider swath of the MODIS instrument [28, Figure 7]. This preferentially diminishes the statistical contributions of outliers to the MODIS time series relative to MISR, disproportionately increasing the relative magnitude of the MISR spikes.

6. ***Deriving Trends from AOD Time Series.*** The green linear trend line shown in [28, Figure 6a] was derived from MISR-MODIS AOD absolute differences constrained to have coincident AERONET AOD retrievals. It is presented as a statistically significant indication that AERONET cloud screening improves MISR-MODIS agreement, because the number of AERONET sites available for the study increased with time over most of the study period [28, Figure 3]. Having additional cloud screening from AERONET should help reduce MISR-MODIS AOD differences, but one would expect smaller deviations later in the time series on statistical grounds alone, due in part to the larger number of counts. If the number of MISR-MODIS coincidences associated with the green curve scales roughly with the number of AERONET sites (the numbers of MISR-AERONET and MODIS-AERONET pairs scale this way according to [28, Figure 7]), there would be many more points between July 2005 and July 2007 than early in the time series.

Instrument radiometric calibration can also contribute to apparent trends, though this factor has been studied with a range of methods [33-36], and the uncertainties are too small to account for the ~0.05 AOD trend shown in [28, Figure 6]. For example, [14] reports a decrease of ~0.02 in MODIS-Terra AOD average absolute error over land relative to AERONET between 2000 and 2007, which might be explained by a known <2% drift in the MODIS blue-band calibration. No such trend is seen for MODIS AOD over ocean, which does not use blue-band radiances. More importantly, these plots show aggregated land and ocean data, and there are many more counts over ocean (e.g., Table 1).

The varying number of counts has the additional effect of making earlier times more sensitive to large outliers. The three largest peaks in the MODIS-MISR time series with AERONET actually occur in April 2001 and in June and August 2002, near the beginning of the series, whereas the smallest values appear at the very end of the curve, between September 2008 and February 2009. If the trend line represents a linear fit with equal weight to the entire series, the high early peaks and low values at the end would exert leverage on the resulting slope. Small AERONET AOD values at the end [28, Figure 2c] are associated with a precipitous decrease in available AERONET stations after April 2008 [28, Figure 3], which, according to the text, reduced the sampling to a small number of what are apparently low-AOD sites. No explanation for the high early peaks is given in [28]; however, these could come, for example, from preferential sampling of the variability in high-AOD events at the AERONET sites used. Lacking further information, limited sampling could allow outliers or unique events to dominate the beginning of the time series, whereas greater sampling would damp outliers later in the time series and a small number of low-AOD cases dominate the very end, combining to produce the apparent trend shown by the green curve.

7. ***Statistical Characterization of Measurement Uncertainty.*** In the discussion following [28, Figures 6 and 7], it is asserted that the MISR and MODIS accuracy statements published in [4, 10], giving the percent of retrieval results falling within a confidence envelope when compared to AERONET validation data "... is both misleading... and inconvenient. Indeed, it cannot be applied to an arbitrary MODIS or MISR pixel since one does not know whether this pixel is one of the 67% 'good' pixels or one of the 33% 'bad' pixels." As summarized in the Introduction to this response, validation of the retrieval results has been done both statistically and through detailed, quantitative analysis of individual, representative cases. The two approaches are complementary; statistical analyses have the advantage of identifying broad trends and patterns, but in themselves,

provide neither explanations nor details about specific cases. We have used, and continue to use, the *combination* of approaches to assess data quality, identify issues, and upgrade the retrieval algorithms, to the extent possible.

8. ***Application of MISR and MODIS Aerosol Products.*** Effective use of the MISR and MODIS aerosol products depends upon (1) choosing problems appropriate to the available measurements, and (2) properly using the data, in light of the known strengths and limitations. Many successful applications of these data to global- and regional-scale questions are already presented in the literature. They range from assessing zonal mean or global aerosol short-wave forcing [37-42] and regional long-wave forcing [43], to improving aerosol forecasting through data assimilation [44, 45], monitoring dust and pollution plume evolution [46, 47] and air quality [48, 49], mapping aerosol air mass type evolution [50], and validating aerosol transport model AOD simulations [51, 52]. In each case, ways of exploiting the strengths of the MISR and MODIS data have been found, and in many cases, independent validation was performed specific to the application. Often, these data sets are combined with each other, with other data, and with model simulations, as required to deal with the coverage and content limitations documented in the data product quality statements and validation papers cited above.

This integrated approach is especially needed for long-term trend analysis, which is extremely sensitive to variability in observing conditions, as well as aerosol amount and type variations combined with sampling biases inherent in the global data sets (Section 6 above; see [53] for a demonstration of how this can be done). Similarly, given sampling biases and limited particle microphysical property information in the satellite data, major advances in assessing aerosol direct radiative forcing globally will likely require the application of data from multiple sources and models [12, 54]. We continue to work toward developing constructive approaches to these cutting-edge questions, taking advantage of the complementary nature of MISR, MODIS, and other satellite and suborbital data.

References

[1] Kaufman, Y. D. Tanre, L.A. Remer, E.F. Vermote, A. Chu, and B.N. Holben, 1997. Operations remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. *J. Geophys. Res. 102*, 17051-17067.

[2] 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, doi:10.1029/2006JD007811.

[3] Martonchik, J.V., R.A. Kahn, and D.J. Diner, 2009. Retrieval of Aerosol Properties over Land Using MISR Observations. In: Kokhanovsky, A., ed., Satellite Aerosol Remote Sensing Over Land. Springer, Berlin.

[4] Remer, L. A., Y.J. Kaufman, D. Tanre, S. Mattoo, D.A. Chu, J.V. Martins, R.-R. Li, C. Ichoku, R.C. Levy, R.G. Kleidman, T.F. Eck, E. Vermote, and B.N. Holben, 2005. The MODIS aerosol algorithm, products, and validation, *J. Atmos. Sci., 62*, 947–973.

[5] Tanre, D., Y.J. Kaufman, M. Herman, and S. Mattoo, 1997. Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances. *J. Geophys. Res. 102*, 16971-16988.

[6] Abdou, W.A., D.J. Diner, J.V. Martonchik, C.J. Bruegge, R.A. Kahn, B.J. Gaitley, K.A. Crean, L.A. Remer, and B. Holben, 2005, Comparison of coincident MISR and MODIS aerosol optical depths over land and ocean scenes containing AERONET sites, *J. Geophys. Res.*, doi:10.1029/2004JD004693.

[7] Chen, W-T, R. Kahn, D. Nelson, K. Yau, and J. Seinfeld, 2008. Sensitivity of multi-angle imaging to optical and microphysical properties of biomass burning aerosols, *J. Geophys. Res. 113*, D10203, doi:10.1029/2007JD009414.

[8] Di Girolamo, L., T.C. Bond. D. Bramer. D.J. Diner, F. Fettinger, R.A. Kahn, J.V. Martonchik, M.V. Ramana, V. Ramanathan, and P.J. Rasch, 2004, Analysis of Multi-angle Imaging SpectroRadiometer (MISR) aerosol optical depths over greater India during winter 2001-2004, *Geophys. Res. Let.* 31, L23115, doi:10.1029/2004GL021273.

[9] Jiang, X., Y. Liu, B. Yu and M. Jiang, 2007. Comparison of MISR aerosol optical thickness with AERONET measurements in Beijing metropolitan area, *Rem. Sens. Environ.*, 107, 45-53.

[10] Kahn, R., B. Gaitley, J. Martonchik, D. Diner, K. Crean, and B. Holben, 2005. MISR global aerosol optical depth validation based on two years of coincident AERONET observations. *J. Geophys. Res. 110, doi:10:1029/2004JD004706.*

[11] Kahn, R. A., M. J. Garay, D. L. Nelson, K. K. Yau, M. A. Bull, B. J. Gaitley, J. V. Martonchik, and R. C. Levy, 2007. 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.

[12] Kahn, R.A., D.L. Nelson, M. Garay, R.C. Levy, M.A. Bull, D.J. Diner, J.V. Martonchik, S.R. Paradise, and E.G. Hansen, and L.A. Remer, 2009. MISR aerosol product attributes, and statistical comparisons with MODIS. *IEEE Trans. Geosci. Remt. Sens.*, 4095-4114.

[13] Kalashnikova O. V., and R. Kahn 2006. Ability of multiangle remote sensing observations to identify and distinguish mineral dust types: Part 2. Sensitivity over dark water, *J. Geophys. Res. 111*, D11207, doi:10.1029/2005JD006756.

[14] 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. Atm. Chem. Phys. Discus., *submitted.*

[15] Liu, Y., J.A. Sarnat, B.A. Coull, P. Koutrakis, and D.J. Jacob, 2004, Validation of MISR aerosol optical thickness measurements using Aerosol Robotic Network (AERONET) observations over the continuous United States, *J. Geophys. Res.*109, doi:10.1029/2003JD003981.

[16] Martonchik, J.V., D.J. Diner, R.A. Kahn, B.J. Gaitley, and B.N. Holben, 2004, Comparison of MISR and AERONET aerosol optical depths over desert sites, *Geophys. Res. Let. 31,* doi:10.1029/2004GL019807.

[17] Pierce, J. R., R. A. Kahn, M. R. Davis, and J. M. Comstock (2010), Detecting thin cirrus in Multiangle Imaging Spectroradiometer aerosol retrievals, J. Geophys. Res., 115, D08201, doi:10.1029/2009JD013019.

[18] Redemann, J., B. Schmid, J.A. Eilers, R. Kahn, R.C. Levy, P.B. Russell, J.M. Livingston, P.V. Hobbs, W.L. Smith, Jr., and B.N. Holben, 2005, Suborbital measurements of spectral aerosol optical depth and its variability at sub-satellite-grid scales in support of CLAMS, 2001, *J. Atmosph. Sci.*, 62(4):993-1007, doi: 10.1175/JAS3387.1.

[19] Reidmiller, D.R., P.V. Hobbs, and R. Kahn, 2006, Aerosol optical properties and particle size distributions on the east coast of the United States, derived from airborne in situ and remote sensing measurements, *J. Atmosph., Sci. 63*, 785–814.

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

[21] Schmid B., J. Redemann, P. B. Russell, P. V. Hobbs, D. L. Hlavka, M. J. McGill, B. N. Holben, E. J. Welton, J. R. Campbell, O. Torres, R. A. Kahn, D. J. Diner, M. C. Helmlinger, D. A. Chu, C. Robles Gonzalez, and G. de Leeuw, 2003. Coordinated airborne, spaceborne, and ground-based measurements of massive, thick aerosol layers during the dry season in southern Africa, *J. Geophys. Res*, 108(D13), 8496, doi:10.1029/2002JD002297.

[22] Xia, X.G., Wang, P.C., Wang, Y.S., et al., 2008. Aerosol optical depth over the Tibetan Plateau and its relation to aerosols over the Taklimakan Desert. *Geophysical Research Letters*, **35** (16): L16804

[23] Xiao, N., T. Shi, C.A. Calder, D.K. Munroe, C. Berrett, S. Wolfinbarger, and D. Li, 2009. Spatial characteristics of the difference between MISR and MODIS aerosol optical depth retrievals over mainland Southeast Asia, *Remt. Sens. Envoron. 113*, 1-9.

[24] Kaufman, Y. J., L. A. Remer, D. Tanre, R. R. Li, R. Kleidman, S. Mattoo, R. Levy, T. Eck, B. N. Holben, C. Ichoku, J. Martins, and I. Koren, 2005. A critical examination of the residual cloud contamination and diurnal sampling effects on MODIS estimates of aerosol over ocean. *IEEE Trans. Geosci. Remote Sens.*, 43 (12), 2886-2897.

[25] Li, R.-R., L. Remer, Y. J. Kaufman, S. Mattoo, B.-C. Gao, and E. Vermote, 2005. Snow and Ice Mask for the MODIS Aerosol Products. *IEEE Geo. and Rem. Sens. Lett.* 2, No. 3, 306-310.

[26] Redemann, J., Q. Zhang, B. Schmid, P. B. Russell, J. M. Livingston, H. Jonsson, and L. A. Remer, 2006. Assessment of MODIS-derived visible and near-IR aerosol optical properties and their spatial variability in the presence of mineral dust. *Geophys. Res. Lett.* 33, L18814, doi:10.1029/2006GL026626.

[27] Russell, P. B., J. M. Livingston, J. Redemann, B. Schmid, S. A. Ramirez, J. Eilers, R. Khan, D. A. Chu, L. Remer, P. K. Quinn, M. J. Rood, and W. Wang, 2007. Multi-grid-cell validation of satellite aerosol property retrievals in INTEX/ITCT/ICARTT 2004. *J. Geophys. Res.* 112, D12S09, doi: 10.1029/2006JD007606.

[28] Mishchenko, M.I., L. Liu, I.V. Geogdzhayev, L.D. Travis, B. Cairns, and A.A. Lacis, 2010. Toward unified satellite climatology of aerosol properties. 3. MODIS versus MISR versus AERONET. *J. Quant. Spectro. Rad. Transf. 111*, 540-552, doi:10.1016/j.jqsrt.2009.11.0003.

[29] Kahn, R.A., B.J. Gaitley, M. Garay, D.J. Diner, J.V. Martonchik, T. Eck, A. Smirnov, and B.N. Holben, 2010. MISR global aerosol product assessment by comparison with AERONET. *J. Geophys. Res.*, submitted.

[30] Hoaglin, D.C., F. Mosteller, and J.W. Tukey, 1983. *Understanding Robust and Exploratory Data Analysis.* John Wiley & Sons.

[31] Willmott, C.J., and K. Matsuura, 2005. Advantages of the meas absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research 30*, 79-82.

[32] Levy, R., G.G. Leptoukh, R. Kahn, V. Zubko, A. Gopalan, and L.Remer, 2009. A critical look at deriving monthly aerosol optical depth from satellite data. *IEEE Trans. Geosci. Remt. Sens.*, 47, 2942-2956.

[33] Bruegge, C.J., W.A. Abdou, D.J. Diner, B.J. Gaitley, M.C. Helmlinger, R.A. Kahn, and J.V. Martonchik, 2004. Validating the MISR radiometric scale for the ocean aerosol science communities. In: Post-launch calibration of satellite sensors, Stanley A. Morain and Amelia M. Budge, editors. A.A. Balkema Publishers, Leiden, Netherlands, pp.103-115.

[34] Kahn, R., W-H. Li, J. Martonchik, C. Bruegge, D. Diner, B. Gaitley, W.Abdou, O. Dubovik, B. Holben, S. Smirnov, Z. Jin, and D. Clark, 2005. MISR low-light-level calibration, and implications for aerosol retrieval over dark water, *J. Atmosph, Sci.* 62, 1032-1062.

[35] Lallart, P., R. Kahn, and D. Tanré, 2008. POLDER2/ADEOSII, MISR, and MODIS/Terra reflectance comparisons, *J. Geophys. Res. 113*, D14S02, doi:10.1029/2007JD009656.

[36] Lyapustin, A., Y. Wang, R. Kahn, J. Xiong, A. Ignatov, R. Wolfe, A. Wu, B. Holben, C. Bruegge, 2007. Analysis of MODIS-MISR calibration differences using surface albedo around AERONET sites and cloud reflectance, <u>Remt. Sensing Enviorn. 107</u>, 12-21, doi:10.1016/j.rse.2006.09.028.

 [37] Kim, D., and V. Ramanathan, 2008. Solar radiation budget and radiative forcing due to aerosol and clouds. *J. Geophys. Res. 113*, D02203, doi:10.1029/2007JD008434.

[38] Kishcha, P., B. Starobinets, O. Kalashnikova, C. N. Long, and P. Alpert, 2009. Variations of meridional aerosol distribution and solar dimming. *J. Geophys. Res.,* 114, doi:10.1029/2008JD010975.

[39] Bellouin, N., O. Boucher, J. Haywood, and M. S. Reddy, 2005. Global estimate of aerosol direct radiative forcing from satellite measurements, *Nature 438*, 1138– 1141.

[40] Christopher, S. A., J. L. Zhang, Y. J. Kaufman, and L. A. Remer, 2006. Satellite-based assessment of top of atmosphere anthropogenic aerosol radiative forcing over cloud-free oceans. *Geophys. Res. Lett.,* **33** (15), L15816.

[41] Loeb, N.G., and Manalo-Smith, N., 2005. Top-of-Atmosphere direct radiative effect of aerosols over global oceans from merged CERES and MODIS observations, *J. Climate 18*, 3506–3526.

[42] Myhre, G., T. F. Berglen, M. Johnsurd, C. Holye, T. K. Berntsen, S. A. Christopher, I. S. A. Isaksen, T. A. Jones, R. A. Kahn, N. Loeb, P. Quinn, L. Remer, and K. E. Yttri, 2008. Radiative forcing of the direct aerosol effect using a multi-observation approach. *Atmos. Chem. Phys. 8*, 12823-12886.

 [43] Zhang J., and S. Christopher, 2003. Longwave radiative forcing of Saharan dust aerosols estimated from MODIS, MISR, and CERES observations on Terra, *Geophys. Res. Lett.* 30:2188, doi:10.1029/2003GL018479.

 [44] 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.

[45] Zhang, J. J.S. Reid, D.L. Westphal, N.L. Baker, and E.J. Hyer, 2008. A system for operational aerosol optical depth data assimilation over global oceans. *J. Geophys. Res. 113*, doi:10.1029/2007JD009065.

[46] Kalashnikova, O.V., and R.A. Kahn, 2008. Mineral dust plume evolution over the Atlantic from combined MISR/MODIS aerosol retrievals. *J. Geophys. Res. 113*, D24204, doi:10.1029/2008JD010083.

[47] Yu, H., L. A. Remer, M. Chin, H. Bian, R. G. Kleidman, and T. Diehl, 2008. A satellite-based assessment of transpacific transport of pollution aerosol. *J. Geophys. Res. 113*, D14S12, doi:10.1029/2007JD009349.

[48] Al-Saadi, J., J. Szykman, R.B. Pierce, C. Kittaka, D. Neil, D.A. Chu, L. Remer, L. Gumley, E. Prins, L. Weinstock, C. McDonald, R. Wayland, F. Dimmick, and J. Fishman, 2005. Improving national air quality forecasts with satellite aerosol observations. *Bul. Am. Meteorol. Soc.* 86(9), 1249.

[49] Liu, Y., P. Koutrakis, R. Kahn, S. Turquety, and R.M. Yantosca, 2007. Estimating PM2.5 component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 2 - A case study, *J. Air & Waste Management Assoc.* 57, 1360-1369.

[50] Dey, S., and L. Di Girolamo, 2010. A climatology of aerosol optical and microphysical properties over the Indian subcontinent from nine years (2000-2008) of Multi-angle Imaging SpectroRadiometer (MISR) data. *J. Geophys. Res.*, in press.

[51] Yu, H., et al., 2006. A Review of measurement-based assessment of aerosol direct radiative effect and forcing. *Atmosph. Chem. and Phys. 6*, 613-666.

[52] Kinne S., M. Schulz, C. Textor, et al., 2006. An AeroCom initial assessment -- optical properties in aerosol component modules of global models. *Atmos. Chem. Phys.* **6**:1815-1834.

[53] Zhang, J. and J.S. Reid, 2010. A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over water MODIS aerosol product. *Atmosph. Chem. Phys.*, submitted.

[54] CCSP (US Climate Change Program), 2009. Synthesis and Assessment Product (SAP) 2.3 – Atmospheric aerosols: Properties and Climate Impacts, pp. 116.

FIGURE CAPTION

Figure 1. January, March, and July 2006 global RSTD values for MISR-MODIS AOD discrepancies, over ocean and land separately. The absolute deviations were placed in rank order for each population, and RSTDs were calculated for progressively larger fractions of the total. The data shown are from MISR Version 22 and MODIS Collection 5. Only the values at the extreme right of each plot are considered in [28].

Table 1. Example values of MISR-MODIS discrepancy and agreement metrics[†]

	Jan 06 – Ocean	Mar 06 – Ocean	Jul 06 - Ocean	Jan 06 - Land	Mar 06 - Land	Jul 06 - Land	Jan 06 – Land+Oce.	Mar 06 – Land+Oce	Jul 06 – Land+Oce.
Counts	9,450,478	9,638,995	10,806,131	2,647,938	2,637,051	5,293,638	12,098,416	12,276,046	16,099,769
MISR <AOD>	0.140	0.148	0.108	0.162	0.203	0.147	0.145	0.160	0.120
MODIS <AOD>	0.138	0.149	0.094	0.120	0.198	0.125	0.134	0.160	0.104
Measures of Discrepancy (%) [Higher values indicate poorer agreement]									
RSTD	32	37	37	85	90	110	50	61	81
RSTD from [24]	31[§]	44[§]	36[§]	87[§]	92[§]	110[§]	69[#]	70[#]	77[#]
RSTD (0.05 δ)*	30	29	38	27	22	29	29	27	34
AAD	21	22	23	62	54	61	31	31	38
Measures of Agreement (%) [Higher values indicate better agreement]									
Envelope 0.05^20%	85	83	88	32	34	34	74	72	71
Envelope 0.03^10%	64	59	67	18	19	19	54	50	51
Abs. <= 0.05	84	81	88	30	30	32	72 (62[¶])	70 (47[¶])	70 (40[¶])
Abs. <= 0.03	63	58	67	18	17	18	53 (44[¶])	50 (35[¶])	51 (28[¶])
Rel. <=20%	54	50	40	15	19	15	46	43	32
Rel. <=10%	30	27	22	8	10	7	25	23	19

[†]RSTD, the Relative Standard Deviation, is a measure of discrepancy, defined by Equation 1. The AAD is defined by Equation 2. The measures of agreement report the percent of MISR mid-visible AOD values falling within the envelope of (1) 0.05 or 20% MODIS AOD, (2) 0.03 or 10% MODIS AOD, (3) within the Absolute-only AOD criterion of 0.05 or (4) 0.03, and (5) within the Relative-only criterion of 20% MODIS AOD or (6) 10% MODIS AOD, respectively. MISR Version 22 and MODIS Collection 5 data are used throughout. All statistics are evaluated for total monthly MISR-MODIS coincident AOD data sets over land, ocean, or land + ocean.
[§]Read from *Mishchenko et al.*, 2010, Figure 1c.
[#]Read from *Mishchenko et al.*, 2010, Figure 6c.

¶Read from *Mishchenko et al.*, 2010, Figure 4a or 4b.

*RSTD calculated by comparing MISR perturbed by $0.05 \times R + 0.05$ (see text)