



New approach to the retrieval of AOD and its uncertainty from MISR observations over dark water

Marcin L. Witek¹, Michael J. Garay¹, David J. Diner¹, Michael A. Bull¹, and Felix C. Seidel¹

¹Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

Correspondence to: Marcin L. Witek (marcin.l.witek@jpl.nasa.gov)

Abstract

A new method for retrieving aerosol optical depth (AOD) and its uncertainty from Multi-angle Imaging SpectroRadiometer (MISR) observations over dark water is outlined. MISR's aerosol retrieval algorithm calculates cost functions between observed and pre-simulated radiances for a range of AODs (from 0.0 to 3.0) and a prescribed set of aerosol mixtures. The previous Version 22 (V22) operational algorithm considered only the AOD that minimized the cost function for each aerosol mixture, then used a combination of these values to compute the final, "best estimate" AOD and associated uncertainty. The new approach considers the entire range of cost functions associated with each aerosol mixture. The uncertainty of the reported AOD depends on a combination of a) the absolute values of the cost functions for each aerosol mixture, b) the widths of the cost function distributions as a function of AOD, and c) the spread of the cost function distributions among the ensemble of mixtures. A key benefit of the new approach is that, unlike the V22 algorithm, it does not rely on arbitrary thresholds imposed on the cost function to determine the success or failure of a particular mixture. Furthermore, a new Aerosol Retrieval Confidence Index (ARCI) is established that can be used to screen high-AOD retrieval blunders caused by cloud contamination or other factors. Requiring $ARCI \geq 0.15$ as a condition for retrieval success is supported through statistical analysis and outperforms the thresholds used in the V22 algorithm. The described changes to the MISR dark water algorithm will become operational in the new MISR aerosol product (V23), planned for release in 2017.

The author's copyright for this publication is transferred to California Institute of Technology

1. Introduction

Uncertainty estimation in satellite remote sensing is a highly challenging endeavor that requires diligent assessment of many potential sources of error. Some of these errors are random and well understood and generally follow a traditional error propagation pathway. Other errors—which in underconstrained retrievals of geophysical quantities can be the dominant sources of uncertainty—are not readily assessed as they stem from various approximations and assumptions that lead to non-linear systematic responses. Additionally, there are sources of error that cannot be quantified at all that are attributable to the resolution of an instrument and variability at scales smaller than those observed.



Recently, Povey and Grainger (2015) provided a detailed overview and classification of possible sources of error in satellite retrievals. They clearly distinguish between pixel-level uncertainties associated with individual retrievals and bulk uncertainty metrics originating from validation studies. The two terms, however, are often not properly differentiated in the literature.

- 5 This is in part due to a lack of awareness about the issue as most studies focus on external validation of satellite data products, which often do not contain information about the uncertainties associated with individual retrievals. This status quo started changing due to growing pressure from the data assimilation community, which requires information on pixel-level uncertainties in order to make proper use of satellite-derived geophysical quantities within
- 10 the framework of a numerical model.

- An example of a remotely sensed geophysical parameter that is being increasingly used in data assimilation applications is the aerosol optical depth (AOD) (Lynch et al., 2016; Shi et al., 2011, 2013; Zhang et al., 2008; Zhang and Reid, 2006, 2009, 2010). AOD represents the total extinction by aerosol particles in the atmospheric column from the surface to the top of the
- 15 atmosphere. Model data assimilation generally requires a value of AOD and an uncertainty associated with this value in order to determine whether or by how much the model should be adjusted to agree with the retrieval. AODs are readily available in many satellite data products, such as those from the Moderate Resolution Imaging Spectroradiometer (MODIS), the Multi-angle Imaging SpectroRadiometer (MISR), and the Visible Infrared Imaging Radiometer Suite (VIIRS),
- 20 whereas individual retrieval uncertainties, in many cases, are not. Pixel-level information often consists of retrieval quality assurance flags, which categorize the “usability” of a retrieval based on a qualitative judgment of the algorithm performance. For example, MODIS Collection 6 uses four coarse labels to communicate retrieval quality, ranging from 0 to 3, with 0 being the least trusted and 3 indicating the highest quality (Hubanks, 2015; Levy et al., 2013; Remer et al., 2005).
- 25 Other metrics conveying some useful information about AOD retrieval uncertainty are the proximity of clouds and cloud coverage in the retrieval region (Shi et al., 2011, 2013, 2014; Witek et al., 2013; Zhang and Reid, 2006). While such metrics are very valuable, they comprise only crude proxies for pixel-level uncertainties and, therefore, have limited quantitative utility.

- The most frequently used metric that quantitatively characterizes the quality of a
- 30 particular AOD dataset as a whole is the error envelope (EE) (Bréon et al., 2011; Garay et al., 2016; Kahn et al., 2010; Levy et al., 2010, 2013; Limbacher and Kahn, 2014; Omar et al., 2013; Remer et al., 2008; Sayer et al., 2012, 2013; Witek et al., 2013). EE results from a validation study where a particular satellite AOD dataset is compared against another AOD dataset, typically ground-based information from the Aerosol Robotic Network (AERONET) (Dubovik et al., 2000; Holben et al.,
- 35 1998), which is considered to represent the “truth”. Taking the general form $\max[\pm a, \pm(b \times \text{AOD})]$, where a and b are empirically determined constants, EE represents the range required so that about 68% (one standard deviation) of the matched data agree, providing an overall characterization of the entire dataset. The EE is often regarded as an expected retrieval error, or confidence envelope. External validation is fundamental to assessing a dataset’s overall quality
- 40 and serves as a useful guide for identifying the presence of systematic errors in the dataset (e.g., Kahn et al., 2010). The EE, which describes the performance of the dataset as a whole relative to a



reference dataset, however, does not represent and should not be confused with the uncertainty characteristics of an individual retrieval.

More appropriate attempts to address the true AOD retrieval uncertainty involved sensitivity tests of the retrieval algorithm with respect to varying external factors such as lower boundary conditions or aerosol microphysical properties (Kahn et al., 2001; Kalashnikova and Kahn, 2006; Sayer et al., 2016). These studies were limited in scope as they were unable to address all possible sources of error. Povey and Grainger (2015) suggested the use of ensemble techniques as a comprehensive means of quantifying uncertainties in satellite remote sensing of the environment. Each member of the ensemble adds a random perturbation to the measurements, ancillary parameters, and underlying retrieval assumptions in order to comprehensively map the probability distribution of the retrieved quantity. Even though such an approach would face significant conceptualization and computational challenges, especially in operational data processing, there are already examples of ensembles being used in aerosol remote sensing. One example is MISR (Diner et al., 1998), which employs many different particle microphysical mixtures as part of its operational aerosol retrieval process (Kahn et al., 2001; Martonchik et al., 1998). The resulting spread of AOD solutions for this ensemble of aerosol mixtures offers quantitative insight into the uncertainty of the individual retrieval. Such an approach, if extended to all possible sources of error and all physically plausible realizations of parameter space, has the potential of providing a robust and comprehensive measure of retrieval uncertainty in the manner suggested by Povey and Grainger (2015).

This study describes a new approach to determining AODs and AOD uncertainties in MISR retrievals. MISR is an instrument aboard the National Aeronautics and Space Administrations (NASA) Earth Observing System (EOS) Terra satellite that performs radiometric observations of the Earth using nine pushbroom cameras pointing at nine different angles (Diner et al., 1998, 2005). All nine cameras observe the same area on Earth within seven minutes, at four different wavelengths. The multi-angle viewing capability allows MISR to sample portions of the scattering phase function simultaneously, providing information that helps distinguish between different aerosol types (Kahn et al., 2001; Kalashnikova and Kahn, 2006). The retrieval process is based on the minimization of a cost function evaluated between the instrument observations and pre-calculated radiometric look-up tables (LUTs) that are in turn based upon a prescribed set of aerosol mixtures. Version 22 (V22) of the algorithm, which has been in operational production since December 2007, considers 74 aerosol mixtures with different microphysical properties that are intended to represent typical atmospheric conditions found on Earth (Kahn et al., 2009, 2010). MISR has two aerosol processing pathways, one for dark water (oceans, seas, deep lakes) and the other for the land surface (Martonchik et al., 2002). The modifications described in this study currently apply only to the dark water algorithm (Kalashnikova et al., 2013; Witek et al., 2013).

The paper is organized as follows. Section 2 describes the V22 operational approach to determining AODs and their uncertainties in MISR aerosol retrievals over dark water. In section 3 critical modifications and changes, which form the basis of a new methodology, are presented. Section 4 introduces an important new metric that is employed to assess the quality of retrievals.



This criterion is used for distinguishing between “good” and “poor” retrievals. Section 5 offers statistical analysis of the new AOD retrieval uncertainty and comparisons against the V22 approach. Finally, a short summary of the study follows in section 6.

5

2. MISR V22 dark water algorithm

A detailed description of the MISR retrieval strategy is described in the Level 2 Aerosol Retrieval Algorithm Theoretical Basis document (Diner et al., 2008). The MISR aerosol retrieval algorithm follows two separate lines of processing depending on the surface type: dark water and land. The two retrieval types are independent of each other and largely rely on different physical and empirical assumptions. Only the dark water algorithm (Kalashnikova et al., 2013) is considered in this study. Here some key elements of the algorithm relevant to the new approach are reviewed.

The problem of retrieving aerosol properties over large water bodies, such as oceans, seas, or deep lakes, is greatly simplified by the reflectance uniformity associated with such surfaces and the fact that deep water bodies are essentially black at red and near-infrared (NIR) wavelengths. One-dimensional radiative transfer (RT) theory is sufficient for determining the relationship between top-of-atmosphere (TOA) radiances and AOD. However, this calculation assumes an aerosol model that specifies the particle size distribution, shape, complex refractive index, and vertical distribution. Additional assumptions are made about the gaseous concentration in the atmosphere (ozone absorption, Rayleigh scattering) and surface whitecap fraction (i.e., the area of the surface covered by white foam from breaking waves). MISR’s ability to observe multi-angle radiances, which are in large part governed by the shape of the aerosol scattering phase functions, provides a wealth of information with which to refine aerosol retrievals over dark water (Kalashnikova and Kahn, 2006).

The MISR aerosol retrieval algorithm relies on a LUT generated for a predefined set of mixtures with known optical properties. The V22 operational algorithm defines 74 aerosol mixtures, each of which consists of up to three unique particle types having prescribed optical and microphysical properties (Kahn et al., 2010; Kahn and Gaitley, 2015). The 74 mixtures consist of combinations of eight primary particle types. The MISR particle types and mixtures are designed to represent several compositional categories typically found in the atmosphere, such as sea spray, sulfate/nitrate, mineral dust, carbonaceous, and urban soot aerosols. Recently, Kahn and Gaitley (2015) provided a thorough verification of MISR’s aerosol type retrieval capabilities and Lee et al. (2016) demonstrated that the MISR aerosol particle climatologies regionally showed good agreement with the results of chemical transport models. The current 74 mixture set, however, is not complete and has some documented deficiencies (Kahn et al., 2010; Kalashnikova et al., 2013; Limbacher and Kahn, 2014). A more comprehensive set of aerosol types and mixtures will be considered in a future release of the MISR’s aerosol algorithm.

For each of the 74 mixtures, forward RT calculations are performed to provide top-of-atmosphere (TOA) radiances for the 36 MISR channels (9 angles and 4 wavelengths). These



radiance is stored in the form of “equivalent reflectances” in the Simulated MISR Ancillary Radiative Transfer (SMART) Dataset, where equivalent reflectance is defined as

$$\rho = \frac{\pi L}{E_0}, \quad (1)$$

where L is the radiance and E_0 is the exo-atmospheric solar irradiance determined for each MISR spectral band. The RT calculations of TOA radiances are carried out for discrete values of mixture optical depth, from 0 to 3, referenced to MISR’s 558 nm (green) band for all plausible combinations of view and solar geometries. The simulations incorporate a modified linear mixing theory for mixtures containing more than one aerosol type, a wind-speed-driven glitter and whitecap model, ozone correction, and Rayleigh scattering (Abdou et al., 1997). The modeled TOA radiances are then directly compared with the MISR observations. The criterion used to find the best-fitting optical depth for a particular mixture is minimization of the reduced χ^2_{abs} parameter, calculated as a function of green-band AOD (τ)

$$\chi^2_{abs}(\tau) = \frac{\sum_{l=1}^4 w_l \cdot \left[\sum_{j=1}^9 v(l,j) \cdot \frac{[\rho_{MISR}(l,j) - \rho_m(l,j)]^2}{\sigma_{abs}^2(l,j)} \right]}{\sum_{l=1}^4 w_l \cdot \left[\sum_{j=1}^9 v(l,j) \right]}. \quad (2)$$

In Eq. 2, ρ_{MISR} are MISR equivalent reflectances, ρ_m are modeled TOA equivalent reflectances for a given aerosol mixture, and σ_{abs} are the absolute radiometric uncertainties in ρ_{MISR} calculated as $\sigma_{abs}(l,j) = 0.05 \max(\rho_{MISR}(l,j), 0.04)$. The summation index l is over the 4 MISR wavelengths and j is over the 9 MISR cameras. The parameter $v(l,j) = 1$ if a valid value of $\rho_{MISR}(l,j)$ exists, and is set to 0 otherwise. The weights w_l are always equal 1 for the red and NIR bands, and are ≤ 1 for the blue and green bands, depending on τ . These weights allow individual bands to contribute varying amounts to the χ^2_{abs} calculation as a function of optical depth. For example, at low τ (<0.5) the MISR 446 nm (blue) and 558 nm (green) bands are not used in χ^2_{abs} calculations since at these wavelengths and τ values the water leaving radiance could be a significant contributor to the TOA signals.

For each mixture, the best fitting value of τ_{mix} is taken to be the value that minimizes χ^2_{abs} using a parabolic fitting approach (Diner et al., 2008). Additional parameters are used to determine the goodness of fit of the particular aerosol mixture to the MISR data. Those are χ^2_{geom} , χ^2_{spec} , and $\chi^2_{max dev}$ which are calculated at the previously obtained value of τ_{mix} . Definitions of these parameters can be found in Diner et al. (2008) and Kalashnikova et al. (2013). An aerosol mixture is considered “successful” if all four metrics χ^2_{abs} , χ^2_{geom} , χ^2_{spec} , and $\chi^2_{max dev}$ do not exceed certain empirically established thresholds. In V22, these thresholds are set to 2, 3, 3, and 5, respectively.



As a result of this procedure, each retrieval region is assigned a set of mixtures and associated AODs that pass all the threshold criteria. In the special case when none of the 74 mixtures pass the threshold tests, the retrieval is considered unsuccessful and no AOD is reported. In most cases, however, multiple mixtures are deemed successful. They typically have somewhat different τ_{mix} values corresponding to their minimum χ_{abs}^2 . The arithmetic mean of all AODs from the passing mixtures is reported as the “best estimate” AOD in the V22 Level 2 aerosol product with the field name “RegBestEstimateSpectralOptDepth”. The retrieval uncertainty, with the field name “RegBestEstimateSpectralOptDepthUnc”, is determined from the standard deviation of the AODs from the passing mixtures. In the case where only a single mixture is successful, the uncertainty is determined directly from the parabolic fit for that mixture (Diner et al., 2008).

A critical aspect of the retrieval process and its outcome is its dependence on a number of empirically determined thresholds. The specific numerical values for the χ_{abs}^2 , χ_{geom}^2 , χ_{spec}^2 , and χ_{maxdev}^2 thresholds were chosen based on pre-production tests aimed at eliminating obvious blunders while maintaining adequate spatial coverage. However, the reported AOD and its uncertainty are directly linked to these thresholds, which entails a certain degree of subjectivity. For example, a mixture and AOD combination resulting in a χ_{abs}^2 value of 1.99 would be considered successful, while a mixture/AOD combination with a χ_{abs}^2 value of 2.01 would not be. Alternatively, mixtures with very different AODs, for example a non-absorbing and an absorbing mixture, might both be considered successful and be included in the uncertainty calculation, but have dramatically different χ_{abs}^2 values, something which is not taken into account when determining the uncertainty. Mitigating such issues was an important driver for developing a new approach to AOD determination and its uncertainty for MISR dark water retrievals in a more objective manner.

3. New approach to AOD retrieval and its uncertainty

The presence of empirical thresholds in goodness-of-fit parameters in the V22 MISR dark water aerosol retrieval algorithm affects the frequency of retrieval success and the resulting AODs, AOD uncertainties, and aerosol properties through the selection of successful mixtures. A more desirable approach would minimize the reliance on arbitrary thresholds. In this section, a new method is described that meets this objective and simplifies the retrieval of the “best estimate” AOD and its associated uncertainty. Furthermore, it results in a single parameter that enables screening of retrieval blunders and AOD outliers and which turns out to be superior to the results derived using the original V22 thresholds. This algorithm revision has been implemented in the software used to generate the next version of the MISR operational aerosol product, V23, scheduled for public release in 2017.

The new method relies solely on the χ_{abs}^2 values calculated using Eq. 2. The other goodness-of-fit metrics are retained solely for diagnostic purposes. Extensive testing showed that AOD



selection in V22 was governed primarily by χ_{abs}^2 , with the other parameters typically having little effect due to the magnitude of their thresholds, except in a limited set of cases. The key elements of the new method are visualized in Figure 1. First, the values of χ_{abs}^2 for each mixture are calculated as continuous functions of τ . The result is then inverted, yielding the distribution of $1/\chi_{abs}^2$. Taking the reciprocal has two primary benefits. First, it gives a smaller weight to retrievals with large values of χ_{abs}^2 that represent poor agreement between the model and the MISR observations. Second, the distribution of $1/\chi_{abs}^2$ tends to look Gaussian, with a peak at τ_{mix} . In the next step, these functions are averaged over all $N=74$ mixtures, leading to:

$$f(\tau) = \frac{1}{N} \sum_{m=1}^N \frac{1}{\chi_{abs,m}^2(\tau)}. \quad (3)$$

- 10 The position of the peak of the average distribution f is the new retrieved AOD
 $AOD = \tau(\max(f(\tau)))$. (4)

The function f , which in most cases closely resembles a Gaussian (normal) distribution, has a peak that is narrow or wide depending on how closely the individual τ_{mix} from the 74 mixtures are clustered. Furthermore, because the absolute values of $1/\chi_{abs,m}^2$ from each mixture contribute to the overall shape of the distribution, aerosol models fitting the observations well (low χ_{abs}^2) dominate the shape of f and the position of its peak, whereas mixtures with poor fits (high χ_{abs}^2) contribute less. The retrieval uncertainty (σ) is determined from the full width at half maximum (FWHM) of f , and scaled to a standard deviation under the assumption that the functional form of f in the vicinity of its peak can be approximated by a Gaussian distribution:

$$20 \quad \sigma = \frac{FWHM(f)}{2\sqrt{2 \ln 2}} \approx \frac{FWHM(f)}{2.3548}. \quad (5)$$

- Equations 4 and 5 form the backbone of the new approach for determining AODs and their uncertainties in MISR retrievals over dark water in the V23 algorithm. One important benefit of the method is that it does not rely on empirically determined thresholds. In all cases, all 74 mixtures contribute to the retrieved AOD, but the amount they contribute depends on how well they agree with the MISR observations. The retrieval uncertainty is then related to the degree to which the AODs associated with the entire ensemble of aerosol mixtures cluster around a specific AOD. If all mixtures are consistent with the same AOD and are highly sensitive to its specific value, the peak in f will be narrow and the uncertainty low. If mixtures disagree as to a single value of AOD, or the χ_{abs}^2 parameter is relatively insensitive to the AOD, the distribution will be broad and the reported uncertainty will be larger.

While the width of the average distribution $f(\tau)$ contains information about the retrieval uncertainty, the peak of the distribution, $\max(f(\tau))$, has an additional important benefit: it can be utilized as a retrieval screening parameter. $\max(f(\tau))$ represents the overall agreement of the



TOA equivalent reflectances from the aerosol mixtures in the LUT with the MISR observations, which can be considered a measure of the confidence in the retrieval. $\max(f(\tau))$ is designated the “Aerosol Retrieval Confidence Index”, or ARCI. Low ARCI implies that generally high χ_{abs}^2 were obtained, indicating that the aerosol models fit the MISR observations poorly. Large ARCI, on the other hand, means that sufficiently many models had low enough χ_{abs}^2 , signifying good agreement with the observations. In a sense, a threshold on ARCI is similar to a threshold on χ_{abs}^2 , except that the former incorporates all aerosol mixtures simultaneously while the latter is applied mixture by mixture. Furthermore, as will be shown in section 4, ARCI is more effective than χ_{abs}^2 in filtering out retrieval blunders and other obvious outliers.

Figure 1 visualizes the important steps of the method using actual MISR data. In this example, the new retrieved AOD is 0.182, whereas the V22 method gave a value of 0.174. The new retrieval uncertainty is 0.049, which is more realistic than the 0.003 uncertainty reported by the V22 algorithm. The very small uncertainty in V22 is due to the fact that only two mixtures were considered successful by passing the V22 thresholds. This example highlights a deficiency in the V22 assessment of retrieval uncertainty as the uncertainty is highly dependent on the number of passing mixtures as well as the value of the four separate thresholds used to determine which mixtures are considered successful. The new procedure eliminates the need for thresholds in determining AOD and its uncertainty, and the only threshold involved is applied to the single ARCI parameter, which is used as a retrieval quality indicator.

4. Retrieval quality assessment

In the MISR V22 retrieval algorithm several thresholds were set to filter out mixtures that do not provide a good match to the instrument observations. The threshold that provides the most strict screening in V22 is $\chi_{abs}^2 \leq 2$, which is applied individually to each aerosol mixture. Because the thresholds provide an additional line of defense against clouds that were not screened by other procedures in the aerosol retrieval process, elimination of these thresholds can result in a large number of high-AOD retrievals in areas that are notorious for frequent cloud cover, but have climatologically very low AODs.

This situation is illustrated in Figure 2, which shows the average AOD for the combination of January and July of 2007 obtained with the new retrieval methodology. Vast areas of the high-latitude oceans are speckled with unrealistically high AODs, clearly indicating an issue with cloud contamination. In V22 the χ_{abs}^2 and other thresholds are able to limit the occurrence of such blunders. In the new algorithm, which performs aerosol retrievals on a 4.4 km grid, in contrast to the coarser 17.6 m grid used in V22, the problem of cloud contamination is further amplified due to closer proximity to cloud edges. Applying the same thresholds as in V22 does not fully mitigate the issue: substantially more 4.4 km retrievals remain cloud contaminated than in V22 (results not shown). Fortunately, the ARCI metric introduced in the previous section proves to be extremely effective at filtering out potentially cloud-contaminated AOD retrievals.



Figure 3a shows average AOD from 4.4 km retrievals as a function of the minimum value of χ_{abs}^2 . In total about 49 million retrievals were analyzed here. After a rapid initial drop due to a similar rapid increase in sampling (Fig. 3b), the average AOD increases gradually with increasing $\min(\chi_{abs}^2)$, while the sampling continues decreasing. The AOD increase could be due to a

5 combination of the increasing number, magnitude, and relative occurrence of cloud-contaminated, high-AOD retrievals with increasing $\min(\chi_{abs}^2)$. Based on the V22 χ_{abs}^2 threshold approach, all retrievals with $\min(\chi_{abs}^2) \leq 2.0$ would have been considered successful. However, no clear justification for a threshold, either at 2.0, or any other value, is evident in the average AOD data. Choosing a value for the threshold that minimizes the average AOD would screen clouds but

10 also potentially screen optically thick aerosol plumes, such as the heavy dust that is prevalent off the west coast of Africa. The picture looks different, however, when one considers the average AOD as a function of ARCI (Fig. 3c). After excluding the initial fluctuation for extremely small ARCI caused by poor sampling, two distinct regimes in the trend of average AOD can be noticed. In the first regime, the average AOD is highly sensitive to the specific value of ARCI, characterized by a

15 sharp decrease in AOD with increasing ARCI between about 0.03 and 0.13. This suggests that a decreasing number of cloud-contaminated, high-AOD retrievals are included in the average as the ARCI is increased. In the second regime, there are relatively small changes in the average AOD as ARCI increases above 0.15. The low AOD gradient in the second regime suggests a low prevalence of cloud contaminated or erroneous AODs. The retrieval count decreases slowly with increasing

20 ARCI (Fig. 3d), indicating that the observed trends in the average AOD are statistically robust. Conveniently, the number of screened retrievals with $\text{ARCI} \geq 0.15$ is similar to the number of retrievals that do not pass the $\chi_{abs}^2 \leq 2.0$ threshold. Out of about 49 million retrievals, 35.9% are below the ARCI threshold (not passing), and 37.1% are above the $\chi_{abs}^2 \leq 2.0$ threshold. We set

25 $\text{ARCI} \geq 0.15$ as the value to be used as a threshold for screening retrieval blunders due to potential cloud contamination or other factors.

Another way to look at the difference between the two screening approaches is presented in Fig. 4a, which shows the two-dimensional distribution of average AOD as a function of $\min(\chi_{abs}^2)$ and ARCI. Figure 4b shows the respective retrieval count. The previous χ_{abs}^2 threshold limit at 2.0 is marked with the black vertical line. All retrievals to the left of this line would have been

30 considered successful in the V22 algorithm. This includes a group of high-AOD (likely cloud-contaminated) retrievals with $\min(\chi_{abs}^2)$ close to 0.1 and ARCI about 0.1. Another suspicious group of retrievals with high average AODs that would have passed the previous threshold is close to $\min(\chi_{abs}^2) = 2.0$. The new ARCI threshold limit, marked with the gray horizontal line, eliminates most of the suspiciously high-AOD regions. All retrievals above the gray horizontal line

35 are considered to be of sufficiently good quality. Of course, more complicated relationships could be investigated, but the use of ARCI as a single screening parameter proves to be highly efficient and furthermore has the advantage of simplicity. The V23 MISR aerosol product will provide the



values of both ARCI and χ_{abs}^2 for use in exploring custom-made cloud screenings and for other purposes.

Figure 5 presents the average AOD distribution obtained using the combination of the January and July 2007 data with retrieval screening based on the ARCI metric ($ARCI \geq 0.15$). This result is directly comparable to Fig. 2, which uses the unscreened data. The benefit of ARCI screening is readily apparent. AODs in large swaths of remote oceans are now represented by smaller and more realistic values (Witek et al., 2013). At the same time, climatologically large AODs off the coasts of Africa and South and East Asia are retained, indicating that the new screening method does not unintentionally remove all high AODs that are likely valid. The global average AOD is reduced from 0.295 for the unscreened data to 0.141 with ARCI screening. However, speckles of high AOD values are still present in many remote and cloudy regions. The majority of these retrievals are visibly cloud contaminated. This demonstrates that the ARCI screening is not ideal as some erroneous AODs pass the threshold. Increasing the threshold reduces the appearance of blunders, but also decreases the number of valid low- and moderate-AOD retrievals, reducing the overall coverage. Because the choice of setting the ARCI threshold limit at 0.15 is well supported statistically (see Figs. 3c, 4a), the remaining cloud-contaminated AOD retrievals should be addressed using another screening method. A possible approach is to employ the clear flag fraction metric discussed in Witek et al. (2013). The application of this approach to removing the remaining cloud-contaminated retrievals in the MISR V23 aerosol product will be discussed in a separate paper.

5. Statistical assessment of AOD retrieval uncertainty.

The AOD retrieval uncertainty described by Eq. 4 is a measure of the sensitivity of the algorithm to the assumed aerosol microphysical properties. This is an important factor affecting retrieval uncertainty (Li et al., 2009; Povey and Grainger, 2015), but, as mentioned in the introduction, there are many other sources of error not accounted for in this approach. Hence, the AOD uncertainty obtained from the algorithm should not be interpreted as a measure of how far the retrieved AOD deviates from the “truth”. This is an important distinction that needs to be properly understood. The calculated uncertainty is purely algorithmic and depends on the initial choice of aerosol mixtures that go into the MISR SMART LUT. Any changes in the prescribed mixtures would lead to different uncertainty estimates. Furthermore, if Eq. 2 is modified such that a different goodness-of-fit metric is used, a different uncertainty result would be expected. For these reasons, interpretation of the established uncertainty does not extend beyond the algorithm’s performance. It does, however, help establish confidence intervals on the retrievals when comparing one pixel to another.

In Figure 6, AOD uncertainty is plotted as a function of AOD using the combined data from January and July of 2007. Only ARCI-screened retrievals are considered. Reported uncertainties are generally much smaller than their associated AODs. Only for very low AOD values do the uncertainties exceed the retrieved AODs. The linear fit to the data indicates that the uncertainty is about 12% of the AOD and has an offset of 0.012. This offset is much smaller than what has been



traditionally discussed in the literature, specifically the 0.03 or 0.05 in the smaller and larger EE, respectively (e.g. Kahn et al., 2010). At low AODs (<0.03) the reported uncertainty is almost always below these offset levels. At higher AODs, uncertainties are often smaller than 20% or even 10% of the retrieved AODs. Note that there is always a substantial spread of AOD

5 uncertainties at any given AOD level, often over an order of magnitude. This shows that, at least from a statistical perspective, the algorithm is capable of representing variability in retrieval confidence. For example, a retrieved AOD of 0.1 can have an uncertainty of 0.05 or of 0.005 depending on circumstances. Assigning physical meaning to a particular uncertainty value as a departure from the true value, as stated earlier, is a task that needs to be addressed separately.

10 A comparison of the new AOD uncertainties against their V22 predecessors reveals many similarities between the two, as evidenced in Figure 7. Recall that the V22 algorithm calculated uncertainties based on the AOD that minimized χ^2_{abs} for each mixture, while the V23 algorithm evaluates the full range of χ^2_{abs} as a function of AOD, so this agreement is not accidental. On average the V23 uncertainty is larger than that reported in V22. There does not seem to be a
15 lower limit on uncertainties in V22, often exhibiting values below 10^{-3} whereas in V23 they mostly stay above 2×10^{-3} . The small values reported in V22 may be due to situations where only a single mixture was considered successful. Furthermore, there is discernable quantization, or clustering, of uncertainties in V22, visible as vertical striping in Fig. 7. This quantization is clearly eliminated in V23. Overall, the new AOD uncertainties appear to have a more reasonable
20 statistical behavior compared to the uncertainties obtained in V22.

6. Conclusions and summary.

Ensemble techniques have been widely used in weather forecasting applications and climate
25 research. They are indispensable in characterizing uncertainties and errors of highly non-linear systems, where standard error propagation techniques cannot be applied. These techniques are also useful tools for quantifying uncertainties in satellite remote retrievals of geophysical quantities (Povey and Grainger, 2015). MISR's aerosol retrieval strategy is a good example of the application of ensemble techniques to retrieval uncertainty assessment in operational data
30 processing.

MISR's aerosol retrieval algorithm uses minimization of a cost function between observations and pre-calculated signals as a function of AOD. The spread of the cost function around a particular AOD value is one indication of the uncertainty of the retrieved solution. Additionally, an ensemble of cost functions for different aerosol mixtures samples sensitivity of
35 the retrieval process to the assumed aerosol optical and microphysical properties. This is one of the major sources of uncertainty in passive remote sensing of AOD. By including an ensemble of aerosol types in the retrieval approach, an algorithmic measure of AOD retrieval uncertainty that includes the impacts of measurement uncertainties, model errors, and aerosol type variability can be effectively derived using MISR data.



This study presents a new approach to determining AODs and AOD uncertainties in MISR retrievals. The new method will become operational for dark water aerosol processing in the upcoming release of V23 of the MISR aerosol product (scheduled for 2017). Unlike the V22 algorithm, the new approach eliminates several arbitrary thresholds. Instead, the AOD and AOD uncertainty determination relies solely on the χ_{abs}^2 metric defined by Eq. 1. All considered mixtures contribute to the final result with a varying influence depending on the shape and magnitude of the associated cost functions. This approach allows for a consistent calculation of AOD and AOD uncertainty without the need for screening acceptable mixture solutions based on a complex interplay of multiple, and somewhat arbitrary, thresholds.

An unintended side effect of the new retrieval approach is an increased abundance of (mostly) cloud-contaminated, high-AOD retrievals in oceanic areas where very low aerosol concentrations are expected. Those blunders—remnants of imperfect cloud screening—were also present in V22, but many were rejected through the use of thresholds on different cost functions. They are more apparent in the V23 results due to the increase in the spatial resolution of the product from 17.6 km in V22 to 4.4 km in V23. Fortunately, an effective screening criterion has been established that filters out most cloud-contaminated retrievals. An analysis of the ARCI metric strongly suggests a specific threshold value, below which the retrievals become increasingly contaminated by clouds. Although this screening method does not eliminate all AOD outliers, it has a strong statistical foundation. Additional cloud screening making use of the clear fraction flag with retrieval regions is built into the V23 algorithm, and will be described separately. Comparison of the new V23 algorithm results (AODs and AOD uncertainties, in particular) to other products, specifically AERONET sunphotometer measurements, will be addressed in future publications.

Acknowledgement.

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Support from the MISR Project is acknowledged. We thank Ralph Kahn for providing helpful comments on the manuscript.



References:

- Abdou, W. A., Martonchik, J. V., Kahn, R. A., West, R. A. and Diner, D. J.: A modified linear-mixing method for calculating atmospheric path radiances of aerosol mixtures, *J. Geophys. Res. Atmos.*, 102(D14), 16883–16888, doi:10.1029/96JD03434, 1997.
- 5 Bréon, F. M., Vermeulen, A. and Desclotres, J.: An evaluation of satellite aerosol products against sunphotometer measurements, *Remote Sens. Environ.*, 115(12), 3102–3111, doi:10.1016/j.rse.2011.06.017, 2011.
- Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V., Ackerman, T. P., Davies, R., Gerstl, S. A. W., Gordon, H. R., Muller, J. P., Myneni, R. B., Sellers, P. J.,
10 Pinty, B. and Verstraete, M. M.: Multiangle Image Spectroradiometer (MISR) instrument description and experiment overview, *IEEE Trans. Geosci. Remote Sens.*, 36(4), 1072–1087, 1998.
- Diner, D. J., Martonchik, J. V., Kahn, R. A., Pinty, B., Gobron, N., Nelson, D. L. and Holben, B. N.: Using angular and spectral shape similarity constraints to improve MISR aerosol and surface retrievals over land, *Remote Sens. Environ.*, 94(2), 155–171, doi:10.1016/j.rse.2004.09.009, 2005.
- 15 Diner, D. J., Abdou, W. A., Ackerman, T. P., Crean, K., Gordon, H. R., Kahn, R. A., Martonchik, J. V., McMuldroch, S., Paradise, S. R. and Pinty, B.: Level 2 aerosol retrieval algorithm theoretical basis, *Jet Propuls. Lab. Calif. Inst. Technol.*, 2008.
- Dubovik, O., Smirnov, a., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F. and Slutsker, I.: Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network
20 (AERONET) Sun and sky radiance measurements, *J. Geophys. Res.*, 105(D8), 9791, doi:10.1029/2000JD900040, 2000.
- Garay, M. J., Kalashnikova, O. V. and Bull, M. A.: Development and Assessment of a High Spatial Resolution (4.4 km) MISR Aerosol Product Using AERONET-DRAGON Data, *Atmos. Chem. Phys. Discuss.*, 1–28, doi:10.5194/acp-2016-569, 2016.
- 25 Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenue, F., Jankowiak, I. and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, *Remote Sens. Environ.*, 66(1), 1–16, doi:10.1016/S0034-4257(98)00031-5, 1998.
- Hubanks, P. A.: MODIS Atmosphere QA Plan for Collection 006, Greenbelt, MD 20071, USA.
- 30 [online] Available from: https://modis-atmos.gsfc.nasa.gov/reference_atbd.html, 2015.
- Kahn, R., Banerjee, P. and McDonald, D.: Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean, *J. Geophys. Res.*, 106(D16), 18219, doi:10.1029/2000JD900497, 2001.
- Kahn, R. A. and Gaitley, B. J.: An analysis of global aerosol type as retrieved by MISR, *J. Geophys. Res. Atmos.*, 120(9), 4248–4281, doi:10.1002/2015JD023322, 2015.
- 35 Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V., Paradise, S. R., Hansen, E. G. and Remer, L. A.: MISR aerosol product attributes and statistical comparisons with MODIS, *IEEE Trans. Geosci. Remote Sens.*, 47(12), 4095–4114, doi:10.1109/TGRS.2009.2023115, 2009.
- Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A. and Holben, B. N.:
40 Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network, *J. Geophys. Res. Atmos.*, 115(23), doi:10.1029/2010JD014601,



2010.
Kalashnikova, O. V and Kahn, R.: Ability of multiangle remote sensing observations to identify and distinguish mineral dust types: 2. Sensitivity over dark water, *J. Geophys. Res. Atmos.*, 111(D11), doi:10.1029/2005JD006756, 2006.
- 5 Kalashnikova, O. V., Garay, M. J., Martonchik, J. V. and Diner, D. J.: MISR Dark Water aerosol retrievals: Operational algorithm sensitivity to particle non-sphericity, *Atmos. Meas. Tech.*, 6(8), 2131–2154, doi:10.5194/amt-6-2131-2013, 2013.
Lee, H., Kalashnikova, O. V., Suzuki, K., Braverman, A., Garay, M. J. and Kahn, R. A.: Climatology of the aerosol optical depth by components from the Multi-angle Imaging SpectroRadiometer
10 (MISR) and chemistry transport models, *Atmos. Chem. Phys.*, 16(10), 6627–6640, doi:10.5194/acp-16-6627-2016, 2016.
Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R. and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos. Chem. Phys.*, 10(21), 10399–10420, doi:10.5194/acp-10-10399-2010, 2010.
- 15 Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F. and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, *Atmos. Meas. Tech.*, 6(11), 2989–3034, doi:10.5194/amt-6-2989-2013, 2013.
Li, Z., Zhao, X., Kahn, R., Mishchenko, M., Remer, L., Lee, K. H., Wang, M., Laszlo, I., Nakajima, T. and Maring, H.: Uncertainties in satellite remote sensing of aerosols and impact on monitoring its
20 long-term trend: A review and perspective, *Ann. Geophys.*, 27(7), 2755–2770, doi:10.5194/angeo-27-2755-2009, 2009.
Limbacher, J. A. and Kahn, R. A.: MISR research-aerosol-algorithm refinements for dark water retrievals, *Atmos. Meas. Tech.*, 7(11), 3989–4007, doi:10.5194/amt-7-3989-2014, 2014.
Lynch, P., Reid, J. S., Westphal, D. L., Zhang, J., Hogan, T. F., Hyer, E. J., Curtis, C. A., Hegg, D. A., Shi, Y., Campbell, J. R., Rubin, J. I., Sessions, W. R., Turk, F. J. and Walker, A. L.: An 11-year global
25 gridded aerosol optical thickness reanalysis (v1.0) for atmospheric and climate sciences, *Geosci. Model Dev.*, 9(4), 1489–1522, doi:10.5194/gmd-9-1489-2016, 2016.
Martonchik, J. V., Diner, D. J., Kahn, R. A., Ackerman, T. P., Verstraete, M. M., Pinty, B. and Gordon, H. R.: Techniques for the Retrieval of aerosol properties over land and ocean using multiangle
30 data, *IEEE Trans Geosci Remote Sens*, 36(4), 1212–1227, 1998.
Martonchik, J. V., Diner, D. J., Crean, K. A. and Bull, M. A.: Regional aerosol retrieval results from MISR, *IEEE Trans. Geosci. Remote Sens.*, 40(7), 1520–1531, doi:10.1109/TGRS.2002.801142, 2002.
Omar, A. H., Winker, D. M., Tackett, J. L., Giles, D. M., Kar, J., Liu, Z., Vaughan, M. A., Powell, K. A. and
35 Trepte, C. R.: CALIOP and AERONET aerosol optical depth comparisons: One size fits none, *J. Geophys. Res. Atmos.*, 118(10), 4748–4766, doi:10.1002/jgrd.50330, 2013.
Povey, A. C. and Grainger, R. G.: Known and unknown unknowns: Uncertainty estimation in satellite remote sensing, *Atmos. Meas. Tech.*, 8(11), 4699–4718, doi:10.5194/amt-8-4699-2015, 2015.
- 40 Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote, E. and Holben, B. N.: The MODIS Aerosol Algorithm,



- Products, and Validation, *J. Atmos. Sci.*, 62(4), 947–973, doi:10.1175/JAS3385.1, 2005.
- Remer, L. A., Kleidman, R. G., Levy, R. C., Kaufman, Y. J., Tanré, D., Mattoo, S., Martins, J. V., Ichoku, C., Koren, I., Yu, H. and Holben, B. N.: Global aerosol climatology from the MODIS satellite sensors, *J. Geophys. Res. Atmos.*, 113(14), 1–18, doi:10.1029/2007JD009661, 2008.
- 5 Sayer, A. M., Hsu, N. C., Bettenhausen, C., Ahmad, Z., Holben, B. N., Smirnov, A., Thomas, G. E. and Zhang, J.: SeaWiFS Ocean Aerosol Retrieval (SOAR): Algorithm, validation, and comparison with other data sets, *J. Geophys. Res. Atmos.*, 117(3), 1–17, doi:10.1029/2011JD016599, 2012.
- Sayer, A. M., Hsu, N. C., Bettenhausen, C. and Jeong, M. J.: Validation and uncertainty estimates for MODIS Collection 6 “deep Blue” aerosol data, *J. Geophys. Res. Atmos.*, 118(14), 7864–7872, doi:10.1002/jgrd.50600, 2013.
- 10 Sayer, A. M., Hsu, N. C., Bettenhausen, C., Lee, J., Redemann, J., Schmid, B. and Shinozuka, Y.: Extending “deep Blue” aerosol retrieval coverage to cases of absorbing aerosols above clouds: Sensitivity analysis and first case studies, *J. Geophys. Res. Atmos.*, 121(9), 4830–4854, doi:10.1002/2015JD024729, 2016.
- 15 Shi, Y., Zhang, J., Reid, J. S., Holben, B., Hyer, E. J. and Curtis, C.: An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation, *Atmos. Chem. Phys.*, 11(2), 557–565, doi:10.5194/acp-11-557-2011, 2011.
- Shi, Y., Zhang, J., Reid, J. S., Hyer, E. J. and Hsu, N. C.: Critical evaluation of the MODIS Deep Blue aerosol optical depth product for data assimilation over North Africa, *Atmos. Meas. Tech.*, 6(4), 949–969, doi:10.5194/amt-6-949-2013, 2013.
- 20 Shi, Y., Zhang, J., Reid, J. S., Liu, B. and Hyer, E. J.: Critical evaluation of cloud contamination in the MISR aerosol products using MODIS cloud mask products, *Atmos. Meas. Tech.*, 7(6), 1791–1801, doi:10.5194/amt-7-1791-2014, 2014.
- Witek, M. L., Garay, M. J., Diner, D. J. and Smirnov, A.: Aerosol optical depths over oceans: A view from MISR retrievals and collocated MAN and AERONET in situ observations, *J. Geophys. Res. Atmos.*, 118(22), 12620–12633, doi:10.1002/2013JD020393, 2013.
- 25 Zhang, J. and Reid, J. S.: MODIS aerosol product analysis for data assimilation: Assessment of over-ocean level 2 aerosol optical thickness retrievals, *J. Geophys. Res. Atmos.*, 111(22), 1–17, doi:10.1029/2005JD006898, 2006.
- 30 Zhang, J. and Reid, J. S.: An analysis of clear sky and contextual biases using an operational over ocean MODIS aerosol product, *Geophys. Res. Lett.*, 36(15), 1–6, doi:10.1029/2009GL038723, 2009.
- Zhang, J. and Reid, J. S.: A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products, *Atmos. Chem. Phys.*, 10(22), 10949–10963, doi:10.5194/acp-10-10949-2010, 2010.
- 35 Zhang, J., Reid, J. S., Westphal, D. L., Baker, N. L. and Hyer, E. J.: A system for operational aerosol optical depth data assimilation over global oceans, *J. Geophys. Res. Atmos.*, 113(10), 1–13, doi:10.1029/2007JD009065, 2008.

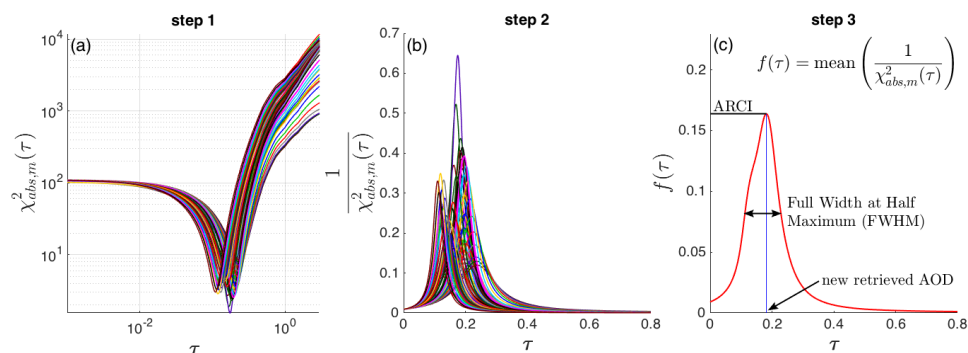


Figure 1 Example of calculation steps performed in the new methodology for determining AOD and its uncertainty. (a) χ_{abs}^2 values for 74 MISR mixtures as a function of AOD (τ) (Eq. 1); (b) inverse (reciprocal) values for the 74 mixtures; and (c) inverse residuals averaged over all mixtures (Eq. 2), with the new retrieved AOD, ARCI, and FWHM indicated on the distribution.

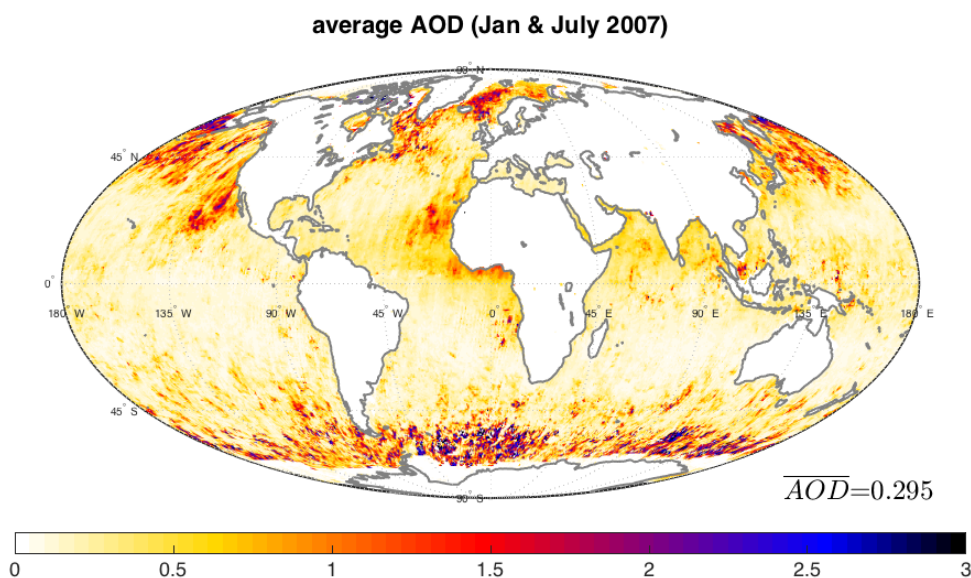


Figure 2 Average AOD obtained using unscreened data from January and July of 2007. The data are mapped to a $0.5^\circ \times 0.5^\circ$ grid. High AOD values over remote oceans indicate issues with cloud contamination in the retrieval process.

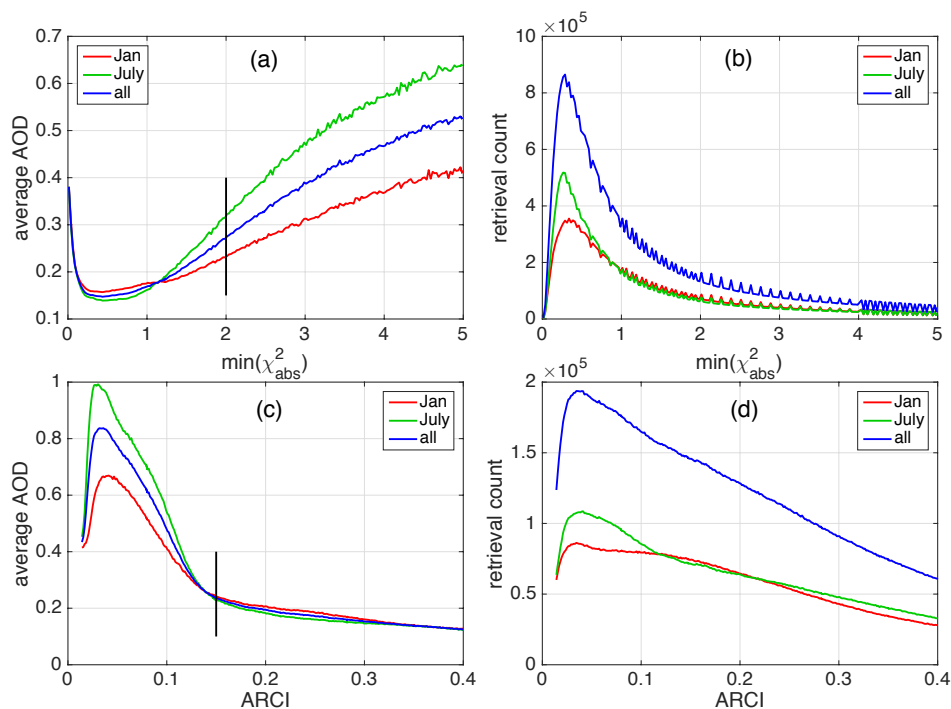


Figure 3 Histograms of average AOD as a function of (a) $\min(\chi^2_{abs})$, and (c) ARCI for January 2007, July 2007, and the two

5 months combined. Panels (b) and (d) are histograms of retrieval counts corresponding to $\min(\chi^2_{abs})$ and ARCI values, respectively.

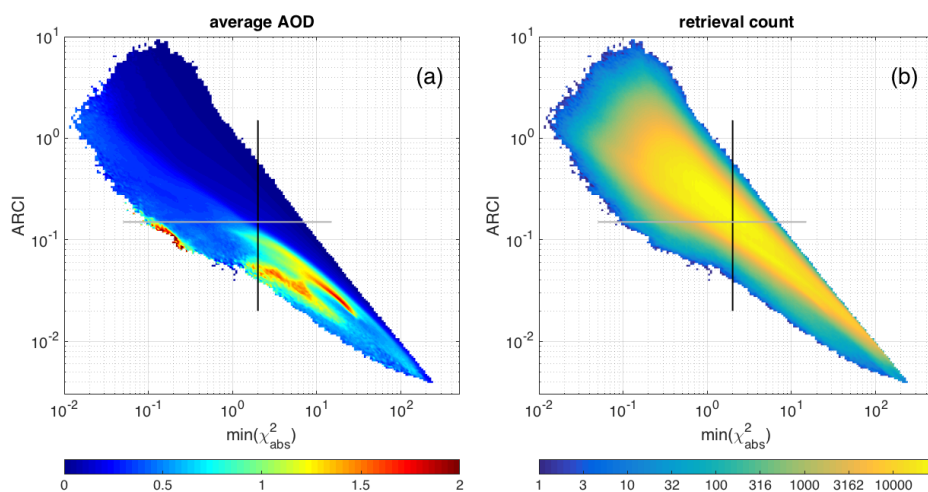


Figure 4 (a) average AOD as a function of ARCI and $\min(\chi^2_{abs})$, (b) retrieval count for the data plotted in panel (a).



average AOD with ARCI screening (Jan & July 2007)

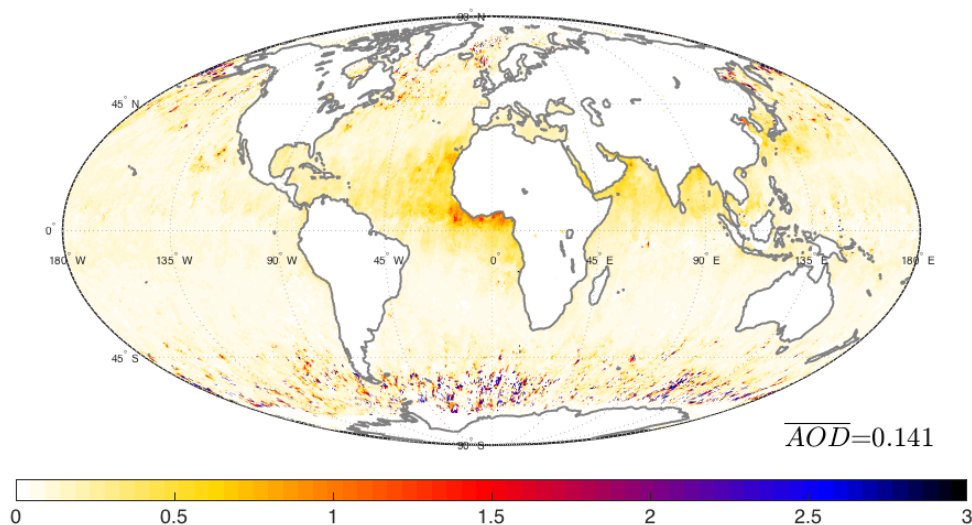


Figure 5 Average AOD distribution with $ARCI \geq 0.15$ screening for the combination of January and July 2007. The data are mapped to a $0.5^\circ \times 0.5^\circ$ grid. There is substantial improvement in the global distribution of mean AODs when compared to the unscreened data in Fig. 2. However, some residual high AOD values remain over remote oceans. They can be further screened using other approaches.

5

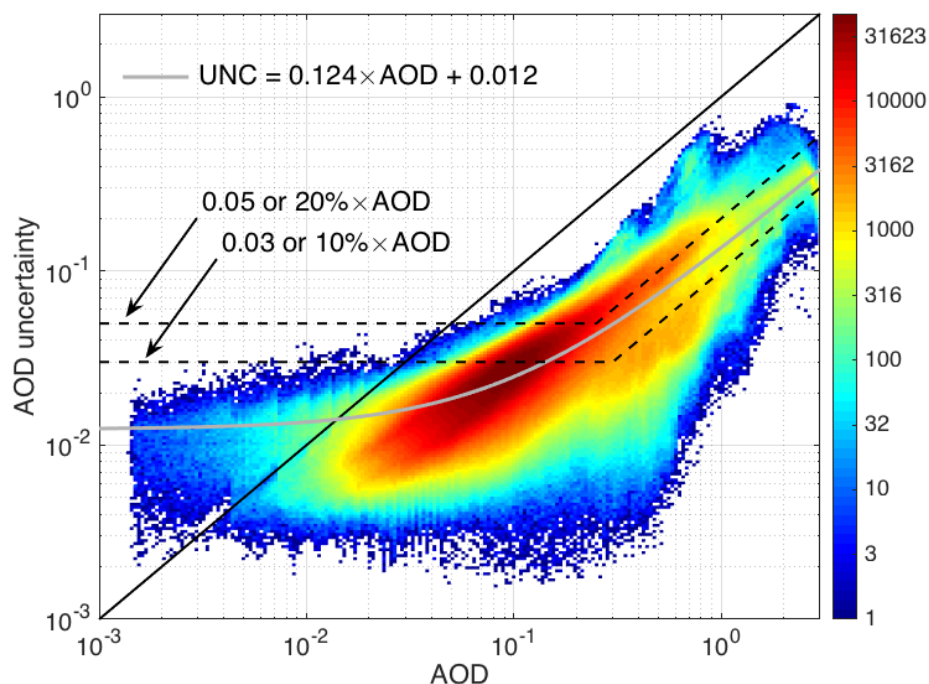


Figure 6 Density plot of AOD uncertainty in log-log space as a function of AOD for the combined January and July 2007 data with ARCI screening. The black line is the 1-to-1 line, included as a visual guide to illustrate that, over most of the AOD range, the uncertainties are smaller than the AOD values themselves. The gray line is a linear fit to the data. Two dashed lines represent two arbitrary uncertainty envelopes.

5

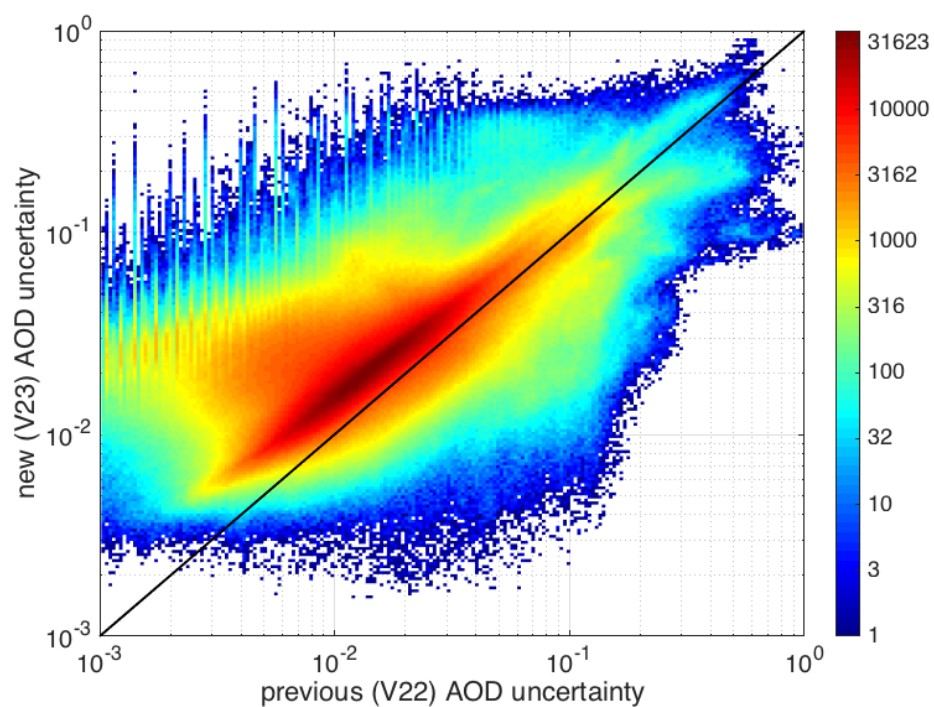


Figure 7 Density plot showing comparison between the previous (V22) AOD uncertainty and the new (V23) AOD uncertainty. The black line is the 1-to-1 line.