Accuracy Assessment of Aqua-MODIS Aerosol Optical Depth over Coastal Regions: Importance of Quality Flag and Sea Surface Wind Speed

Jacob Anderson
University of Nebraska-Lincoln, jacoba@huskers.unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/geoscidiss

Part of the Atmospheric Sciences Commons, Environmental Monitoring Commons, and the Other Earth Sciences Commons

http://digitalcommons.unl.edu/geoscidiss/29

This Article is brought to you for free and open access by the Earth and Atmospheric Sciences, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Dissertations & Theses in Earth and Atmospheric Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Accuracy Assessment of Aqua-MODIS Aerosol Optical Depth over Coastal Regions: Importance of Quality Flag and Sea Surface Wind Speed

By

Jacob C. Anderson

A THESIS

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirement
For the Degree of Master of Science

Major: Earth and Atmospheric Sciences

Under the Supervision of Professor Jun Wang

Lincoln, Nebraska

August, 2012
ACCURACY ASSESSMENT OF AQUA-MODIS AEROSOL OPTICAL DEPTH
OVER COASTAL REGIONS: IMPORTANCE OF QUALITY FLAG AND SEA SURFACE WIND SPEED

Jacob C. Anderson, M.S.
University of Nebraska, 2012

Advisor: Jun Wang

Using data collected from 62 coastal stations worldwide from the Aerosol Robotic Network (AERONET) from 2002-2011, accuracy assessments are made for coastal aerosol optical depth (AOD) retrieved from MODIS aboard the Aqua satellite. It is found that coastal AODs (at 550 nm) characterized respectively by the MODIS Dark Land (Land) surface algorithm, the Open Ocean (Ocean) algorithm, and AERONET all exhibit a log-normal distribution. After filtering by quality flags, the coastal MODIS AODs retrieved from the Land and Ocean algorithms are highly correlated with AERONET (with R^2~0.8), but only the Land algorithm AODs fall within the expected error envelope greater than 66% of the time. Furthermore, the MODIS AODs from the Land algorithm, Ocean algorithm, and combined Land_And_Ocean product show statistically significant discrepancies from their AERONET counterparts in terms of mean, probability density function, and cumulative density function, which suggest a need for future improvement. Without filtering with quality flag, the MODIS Land_And_Ocean AOD dataset can be degraded by 30-50% in terms of mean bias. Overall, the MODIS Ocean algorithm overestimates (underestimates) the coastal AOD by 0.021 (0.029) for AOD < 0.25 (>
0.25), which is shown to be related to the ocean surface wind speed and cloud contamination. The Modern Era Retrospective-Analysis for Research and Applications (MERRA) reveals that wind speeds over the global coastal region (with a mean and median value of 2.94 m s\(^{-1}\) and 2.66 m s\(^{-1}\) respectively) are often slower than the constant 6 m s\(^{-1}\) assumed in the MODIS Ocean algorithm. As a result of high correlation (R\(^2\)>0.98) between the bias in binned MODIS AOD and the corresponding binned wind speed over the coastal sea surface, an empirical scheme for correcting the bias of AOD retrieved from the MODIS Ocean algorithm is formulated and is shown to be effective over the majority of the coastal AERONET stations, and hence can be used in future analysis of AOD trend and MODIS AOD data assimilation.
# TABLE OF CONTENTS

1. Introduction .............................................................................................................. 1

2. Data .......................................................................................................................... 7
   2.1 MODIS and AERONET AOD products ................................................................. 7
   2.2 Sea Surface Wind Speed Data .............................................................................. 9
   2.3 MODIS-AERONET Collocation and Coastal Site Classification ....................... 10

3. Methodology ............................................................................................................. 12
   3.1 Ordinary Statistics .............................................................................................. 12
   3.2 Log-Normal Distributions .................................................................................. 13
   3.3 Robust Statistical Techniques ............................................................................ 14
      3.3.1 T-Test .......................................................................................................... 14
      3.3.2 Likelihood Ratio Test .................................................................................. 15
      3.3.3 Kolmogorov-Smirnov Test ......................................................................... 15

4. Results and Discussion ............................................................................................ 17
   4.1 Coastal vs. Non-Coastal MODIS AOD Evaluation ............................................... 17
   4.2 Land_And_Ocean Quality Control ..................................................................... 22
   4.3 Wind and Cloud Impact on the MODIS Ocean Algorithm .................................. 24

5. Summary and Conclusions ....................................................................................... 31

Acknowledgements ...................................................................................................... 33

References .................................................................................................................... 34

Tables .......................................................................................................................... 39
Figures ................................................................................................................................. 43
Appendix A ................................................................................................................................ 52
LIST OF TABLES AND FIGURES

TABLES

Table 1. MODIS AOD mean bias ................................................................. 39
Table 2. MODIS AOD bias after cloud filtering ........................................ 40
Table 3. Regression statistics of for the MODIS AOD products .................. 41
Table 4. Bias result from empirical correction .......................................... 42

FIGURES

Figure 1. Frequency of coastal AOD .......................................................... 43
Figure 2. Probability density functions of AOD ......................................... 44
Figure 3. Regressions of MODIS AOD and AERONET AOD .................... 45
Figure 4. Station plots of MODIS AOD bias ............................................. 46
Figure 5. Taylor diagram ........................................................................ 47
Figure 6. Cumulative density functions of AOD ........................................ 48
Figure 7. Station plots of wind speed and MODIS AOD regressions ........... 49
Figure 8. Scatter plots of empirical corrections ......................................... 50
Figure 9. Improvements from the empirical corrections ........................... 51
1. Introduction

Aerosols play an important role in the Earth’s energy balance and hydrological cycle (Charlson et al., 1992) through scattering and absorbing radiation, as well as acting as cloud condensation nuclei in the atmosphere. These airborne particles also reduce visibility and affect human health (Samet et al., 2000). The Intergovernmental Panel on Climate Change (IPCC), in their fourth assessment, reports that the direct aerosol radiative forcing is best estimated as -0.5 W m\(^{-2}\) (IPCC, 2007), rendering a cooling powerful enough to offset the warming from CO\(_2\) by almost one-third (Myhre, 2012). However, the uncertainties associated with this best estimate are close to 80%, i.e., the range of aerosol radiative forcing is from -0.1 to -0.9 Wm\(^{-2}\) (IPCC, 2007). Further reduction of such large uncertainties, especially through observation-based characterization of aerosol properties on a global scale, is needed for improved prediction of climate change (IPCC, 2007). Since the IPCC report in 2007, various studies have evaluated the trends of aerosol optical depth (AOD), the first-order indication of columnar aerosol mass and aerosol forcing, using the data retrieved from different satellite sensors, including the moderate resolution imaging spectroradiometer (MODIS), advanced very high resolution radiometer (AVHRR), and multi-angle imaging spectroradiometer (MISR). However, these past evaluations only studied the trend of aerosols over the open ocean and show inconsistent results (Zhang and Reid, 2006; Mishchenko and Geogdzhayev, 2007; Remer et al., 2008; Zhao et al., 2011), with both
positive and negative trends estimated in global averages reported by Mishchenko et al. (2007) and Zhang and Reid (2010), respectively.

In the context of trend analysis of satellite-based AOD, this study focuses on the evaluation and refinement of the accuracy of AOD retrieved from MODIS over the coastal regions. Since its inception, the MODIS level 2 aerosol product, due to its daily and nearly global coverage, has been used widely in research for studying aerosol processes ranging from source, transport, and deposition to impacts on air quality and climate (Remer et al., 2005; Kahn et al., 2005, 2007, 2011; Levy et al., 2007a, 2010; Mi et al., 2007; Chatterjee et al., 2010; Bhaskaran et al., 2011; Bréon et al., 2011; Smirnov et al., 2011). World-wide comparisons of MODIS AOD with those measured from AERONET show that the MODIS AOD product overall has an accuracy of $\pm(0.05 + 0.15\cdot AOD_{\text{actual}})$ over the land and $\pm(0.03 + 0.05\cdot AOD_{\text{actual}})$ over the ocean (Remer et al., 2005; Kahn et al., 2005, 2007, 2011; Levy et al., 2007a, 2010). Such uncertainty brackets the value of the AOD trend, at 550 nm, of -0.03 from 1991 to 2005 reported by Mishchenko et al. (2007) and of 0.003 per decade found by Zhang and Reid (2010).

Furthermore, the MODIS AOD uncertainty at regional scales on either land or ocean can be much greater than on a global scale (Levy et al., 2005) because the accuracy of aerosol retrievals is subject to the change of boundary conditions and surface types (Levy et al., 2007b). In this regard, the MODIS AOD products over the coastal regions deserve special attention because they are a simple union of the retrievals from algorithms that are designed for either over land only or over open ocean only and (as discussed below)
neither algorithm has a dedicated scheme to characterize the surface reflectances at the coast that are often influenced by a sand-water mixture and water-leaving radiance contributed by the underlying sea shore and suspended matter in shallow ocean water. Furthermore, over half the world’s population resides in the coastal region (Tibbetts, 2002), which makes assessment of AOD over the coastal region critical for understanding the global trend of AOD, especially anthropogenic AOD, and may have important implications for future air quality studies.

Both MODIS aerosol retrieval algorithms over dark surfaces, i.e., the Open Ocean (hereafter Ocean) algorithm and the Dark Land (hereafter Land) algorithm, use the cloud-free top of the atmosphere (TOA) reflectances that are measured at resolutions ranging from 250 m in the shortwave visible wavelengths to 500 m in the near-infrared and are then aggregated to boxes of 20 by 20 (500 m resolution) pixels or equivalent to 10 by 10 km resolution at nadir for aerosol retrieval (Remer et al., 2005). The 10 km resolution AOD products are analyzed in this study. The Ocean algorithm is used for retrieval if all pixels within the 20x20-pixel box are water; otherwise, the Land algorithms are used. Determining if any pixel within the 20x20-pixel box is over land is based upon the MOD35 1km data that contains information about surface type (Remer et al., 2005). To date, a simple union of the AODs retrieved from the Land and Ocean algorithms make up the MODIS, level 2, “Land_And_Ocean” AOD product that is popularly used in the research community.
However, such a simple union of separate retrievals from the Land and Ocean algorithms renders several difficulties for the evaluation of MODIS AOD retrievals over the coastal regions. First, the algorithm used for AOD retrieval in the same location can vary with the time or with the MODIS ground track. This can be understood because the pixel resolution at the ground is a function of satellite viewing zenith angle (Figure A1). With a repeat cycle of 16 days, a box of MODIS 20x20-pixels can be exactly equal to 10 by 10 km when viewed by MODIS at nadir, but can also be equivalent to 20 by 48 km area when viewed by MODIS at the high viewing zenith angle. In the former condition, the AOD for 20x20-pixels over a coastal ocean region might be retrieved by the Ocean algorithm. In the latter case, the 20x20-pixels on which the MODIS algorithm operates may possibly contain one or more land pixel(s) and, therefore, the Land algorithm would be applied over a larger area containing the same coastal region. Secondly, the quality flags that are stored with the Land and Ocean retrievals are not passed through to the Land_And_Ocean data set, rendering a loss of the quality information associated with each AOD retrieval.

Besides the aforementioned two issues associated with the simple data merge of MODIS AOD over the coast, this study will also look into the assumptions made by the MODIS Look Up Tables (LUTs) corresponding to the ocean surface characteristics. Inherent in the MODIS LUT are assumptions about the ocean surface (Kleidman et al., 2012). Three ocean surface properties are particularly important in the Ocean algorithm: water-leaving radiance, rough ocean surface causing changes in sun glint patterns, and
white caps (Kleidman et al., 2012). The spectral water-leaving radiances are determined by suspended material in the surface layer, and can be influenced by shallow ocean floor reflectance. Therefore, these values can vary significantly from open ocean to coastal ocean. However, such variation is not considered in the current MODIS aerosol algorithm that assumes 0.0 water leaving radiances for all but the 550 nm channel where a value of 0.005 is assumed (Remer et al., 2005).

The Ocean algorithm uses a Cox and Munk (1954) rough ocean surface model to provide the sun glint pattern. The algorithm masks all geometry within 40 degrees of specular reflection. Both the glint patterns and the white caps are determined by surface wind speed. The MODIS aerosol algorithm assumes a constant 6 m s\(^{-1}\) wind speed for all retrievals. A Koepke (1984) model is used to estimate the white cap reflectance contribution. Thus, the assumed wind speed of 6 m s\(^{-1}\) in the algorithm needs to be evaluated. Kleidman et al. (2012) demonstrate the dependence of the MODIS aerosol accuracy over the open ocean on surface wind speed, and show that significant error can arise when wind speeds are faster or slower than 6 m s\(^{-1}\). This can be understood because the ocean surface has greater reflectance with higher wind speeds and lower reflectance at slower wind speeds. The change in the ocean surface reflectance due to wind speed is not accounted for in the Ocean algorithm leading to an overestimation of AOD for wind speeds greater than 6 m s\(^{-1}\) and an underestimation for wind speeds less than 6 m s\(^{-1}\). Previous evaluations of the Ocean algorithm have studied the effect of wind speed (Zhang and Reid, 2006; Shi et al., 2011; Kleidman et al., 2012) and cloud contamination
(Zhang and Reid, 2006 and 2010) on retrievals of aerosols, however, these are restricted to the open ocean. Hence, the evaluation of the effect of cloud contamination and assumed sea surface wind speed on aerosol retrievals over coastal oceans needs further assessment.

This paper is designed to address three issues, the overall accuracy of MODIS AOD over the coast, the impact of MODIS quality flags, and the sea surface wind speed impact on Ocean retrievals. To avoid the issues related to Terra-MODIS calibration (Remer et al., 2005), we here only evaluate the accuracy of Aqua-MODIS AOD.
2. Data

2.1 MODIS and AERONET AOD products

MODIS AOD is reported at 7 wavelengths (470, 550, 660, 870, 1200, 1600, and 2100 nm) for the Ocean algorithm and three wavelengths (470, 550, and 660 nm) for the Land algorithm. The 550 nm wavelength is used for comparison with AERONET because it is consistent with the primary wavelength used by many climate and chemistry transport models (Kinne et al., 2003) as well as previous MODIS validation studies (Remer et al., 2005; Levy et al., 2007a, 2010). Note that vegetated surfaces are not “dark” in the 550 nm wavelength, and therefore, AOD cannot be directly retrieved over land. The AOD at 550 nm over land is derived from the retrieved AODs at multiple MODIS channels, usually from the 470 nm and 660 nm wavelengths when those data are available (Levy et al., 2010).

The MODIS aerosol algorithms operate through the use of look-up tables (LUTs) in order to estimate AOD (Levy et al., 2010). The LUTs store the top-of-atmosphere (TOA) spectral reflectance from pre-computed radiative transfer equations with the assumptions that only a small set of aerosol types, loadings, and geometries can represent the global range of aerosol conditions (Kaufman et al., 1997; Levy et al., 2010). The LUTs represent spectrally consistent atmospheric properties computed for different aerosol types that are classified according to AERONET characterization and vary with geographical location and season (Levy et al., 2010). The spectral surface reflectance contribution is constrained by the 2100/670 nm reflectance ratio, and further description
of the parameterization is described by Levy et al. (2007b). Thus, with the surface contribution “known”, the MODIS observed TOA reflectance at the 470 nm and 660 nm wavelengths can be compared to the LUT to determine the best-fit AOD. The best fit represents the solution that provides the smallest fitting error when matching the LUT spectral reflectances to MODIS observations (Levy et al., 2010).

MODIS collection 5.1, quality assured, level 2 data from 4th of July, 2002 through 10th of January, 2011 are used (http://giovanni.gsfc.nasa.gov/mapss/). MODIS uses quality flags to represent the accuracy of AOD retrievals. The quality flags range from 3 (high confidence) to 0 (low or no confidence) (Levy et al., 2010). The quality flags are assigned to each MODIS AOD retrieval based on the number and quality of pixels used in the AOD algorithms (Remer et al., 2005; Levy et al., 2010). It has been shown in previous research that the quality flags associated with MODIS retrievals play a significant role in MODIS AOD error approximation (Remer et al., 2005; Levy et al., 2010). The Expected Error (EE) envelope for the MODIS aerosol algorithms are represented by $EE = \pm 0.05 \pm 0.15 \cdot \text{AOD}_{\text{actual}}$ for the Land algorithm, and $EE = \pm 0.03 \pm 0.05 \cdot \text{AOD}_{\text{actual}}$ for the Ocean algorithm (Remer et al., 2005). The retrievals that fall within the EE bound (greater than 66% of the time), on a global average, are represented by the quality flag 3 for the Land algorithm (MODIS collection 5), and the flags 1, 2, and 3 for the Ocean algorithm (MODIS collection 4) (Remer et al., 2005; Levy et al., 2010). Therefore, all MODIS data in this study is filtered by quality flag.
AERONET derives AOD from direct sun photometer measurements in some or all of the following seven different spectral bands centered at 340, 380, 440, 500, 670, 940, and 1020 nm (Holben et al., 1998). AERONET measures the extinction of direct beam solar radiation, and applies the Beer-Lambert-Bouguer law to determine AOD (Holben et al., 1998) with uncertainties on the order of 0.01-0.02 (Eck et al., 1999). Only quality assured, cloud screened, AERONET Level 2 data are used in this study to evaluate the MODIS aerosol product (http://giovanni.gsfc.nasa.gov/mapss/). To facilitate the comparison with MODIS, AERONET AOD measurements are interpolated to the 550 nm wavelength from multiple AERONET wavelengths using a quadratic fit on a log-log scale (Eck et al., 1999).

2.2 Sea Surface Wind Speed Data

The Modern Era Retrospective-analysis for Research and Applications (MERRA) meteorological database (tavg1_2d_flx_Nx; http://disc.sci.gsfc.nasa.gov/mdisc/) is used in the evaluation of the AOD retrieval uncertainty due to the surface wind speed assumption in the MODIS Ocean algorithm. MERRA has 0.5 degree latitude by 0.66 degree longitude resolution and provides an extensive source of meteorological variables (Rienecker et al., 2011). It uses the Goddard Earth Observing System-5 Data Assimilation System (GEOS-5 DAS) and a new set of physics packages for the atmospheric general circulation model (AGCM). The wind-related inputs into the MERRA system include wind speed data from radiosondes, pilot balloon (PIBAL) measured winds, MODIS, Geostationary Operational Environmental Satellites (GOES),
Special Sensor Microwave/Imager (SSM/I), Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), NASA’s Quick Scatterometer (QuickSCAT) and others. More information on the MERRA inputs can be found in the MERRA file specification document (http://gmao.gsfc.nasa.gov/merra/file_specifications.php).

MERRA has been evaluated for accuracy and found to be one of the “best performing” reanalysis products for ocean surface turbulent flux and wind stress parameters (Brunke et al., 2011). Kennedy et al. (2011) show the MERRA near-surface wind speeds have biases within 0.5 m s\(^{-1}\) when compared to the Atmospheric Radiation Measurement Program’s (ARM) Cloud Modeling Best Estimate (CMBE) soundings during 1999-2001.

### 2.3 MODIS-AERONET Collocation and Coastal Site Classification

The spatially and temporally collocated MODIS and AERONET data pairs are acquired through the Multi-Sensor Aerosol Product Sampling System (MAPSS, http://giovanni.gsfc.nasa.gov/mapss/) (Ichoku et al., 2002; Petrenko et al., 2012). The database is a result of collocating the MODIS and AERONET measurements over the full record of MODIS (both Aqua and Terra) spanning the years 2002-2011 for this study for the Aqua satellite. Two methods are used for collocating the MODIS and AERONET AOD data. Firstly, AERONET measurements within +/- 30 minutes of the MODIS overpass time are averaged and compared against MODIS AOD retrievals averaged within a 55 km diameter centered over the AERONET sites (Figure A2, mean method, Ichoku et al., 2002). Using the mean method MAPSS also saves the mode of the quality flags from each pixel within the averaging region (55 km) to represent the quality flag for
the collocated MODIS retrieval (Petrenko et al., 2012). The quality flag mode is then used to filter the AOD data for only high quality retrievals. Secondly, the MODIS AOD retrieval closest to the AERONET site is paired with the AERONET measurement that is closest to the MODIS overpass time (central method). Petrenko et al. (2012), find that there is little difference between the central and mean methods, therefore, to be consistent with previous research and maximize data volume the mean method is used for of this research.

After acquiring the collocated MODIS retrievals at all AERONET sites a quality filter is run. AERONET sites must have at least 15 high quality (flag 3 for Land, or flags 1, 2, or 3 for Ocean) MODIS retrieval pairs to be used in this study. To be considered a coastal site, AOD measurements from AERONET must be paired with both the MODIS Ocean algorithm at least once and the MODIS Land algorithm at least once over the full time period of this study (2002-2011). For the roughly 9 year record of Aqua-MODIS and AERONET AOD pairs from MAPSS, the result from the mean collocation method, that is consistent with Ichoku et al. (2002), shows that 62, or roughly 26%, of the AERONET stations have MODIS retrievals from both the Land and Ocean algorithms, and consequently those sites are designated as coastal. All other AERONET sites are designated as non-coastal, and are either paired only with the Land algorithm or the Ocean algorithm.
3. Methodology

The MODIS-AERONET pairs are examined after they are split into three categories. The first includes all AERONET sites (global), the second consists of only coastal AERONET sites (coastal), and the third is made up of only non-coastal sites (non-coastal). Each MODIS algorithm is separately evaluated within each category. The land-only and ocean-only retrievals are grouped into the non-coastal category because the focus here is on the difference between coastal and non-coastal MODIS AOD performance. Evaluations of the MODIS AOD performance over the ocean and over land have already been attempted (Remer et al., 2005; Levy et al., 2010). We utilize multiple metrics to statistically evaluate the MODIS AOD accuracy with respect to AOD measured by AERONET.

3.1 Ordinary Statistics

The first type of metric is a combination of parameters that are commonly used to describe the relationship between two variables including: bias, mean, standard deviation, correlation, statistical significance, and best-fit (ordinary-least-square) regressions. MODIS AOD bias is calculated by subtracting AERONET AOD from the paired MODIS AOD, and then averaging the differences for each AERONET site over the full time period (2002-2011) to obtain the mean bias for each site. In addition, the biases are calculated for the Land, Ocean, and Land_And_Ocean MODIS products respectively. Station plots are then created to reveal the MODIS bias along the coast. Furthermore, the correlation, variance and root mean square difference (RMSD) between MODIS AOD
and AERONET AOD are combined to generate the well known Taylor Diagram to aid the visualization of the differences found in the comparison. Designed by Taylor (2001), the Taylor Diagram uses a 2D polar plot to demonstrate three pieces of information that are interconnected, in which radius represents normalized standard deviations, cosine of the angle represents correlation, and the radius of the circles centered on point “REF” (e.g., radius of 1) along the x-axis indicates normalized RMSD. As will be shown in Section 4, the Taylor Diagram is particularly useful for visualizing the error characteristics of each of the MODIS aerosol algorithms over varying surface types.

### 3.2 Log-Normal Distributions

While the first type of metric is useful, it may not be sufficient to fully describe the goodness of fit between two data sets, especially when the population in the datasets are not normally distributed (Wilks, 2011). In other words, statistically significant correlation and/or small bias does not necessarily warrant that the fit between the two datasets is statistically significant. As shown in Figure 1, the AOD frequencies over coastal sites (and non-coastal sites) are not normally distributed; indeed, they are log-normal (Figure 2), which is consistent with O’Neill et al. (2000). Note that the MODIS Land algorithm allows negative values when retrieving AOD, even though negative AOD values are not physically possible. From 2002-2011 approximately 400 retrievals from MODIS when paired with AERONET over the coastal regions resulted in negative AODs and those retrievals are not included in the frequency/PDF/CDF analyses. Two parameters, $\mu$ and $\sigma$, represent the location parameter and scale parameter respectively,
and can be identified to fully describe a log-normal probability density function (PDF):

\[ f(\tau; \mu, \sigma) = \frac{1}{\tau \sqrt{2\pi \sigma^2}} e^{-\frac{(\ln(\tau) - \mu)^2}{2\sigma^2}}. \]

Where \( \mu \) is the mean of the logarithm of AODs, \( \sigma \) is the standard deviation of the logarithm of AODs, and \( \tau \) is AOD. The actual frequency for AOD values between \( \tau \) and \( \tau + \Delta \tau \) can be obtained by integrating the PDF over the range \( \tau \) to \( \tau + \Delta \tau \), and then multiplying the integral by the total number of sample data points.

### 3.3 Robust Statistical Techniques

To evaluate if the (log-normal) PDFs of MODIS AOD data fit with that of the AERONET measurements at a statistically significant level, a second type of statistic metric is used that consists of a t-test for difference of mean for paired data, a likelihood ratio test, and a Kolmogorov-Smirnov (K-S) test. Briefly, after the difference (e.g. bias) in each data pair \( \Delta_i \) is computed, the mean difference (or mean bias) \( \bar{\Delta} \) can be estimated and then compared with the difference between the means for each variable (e.g. MODIS AOD or AERONET AOD), \( \mu_\Delta \).

#### 3.3.1 T-Test

The t-test for statistical significance is then applied to

\[ z = \frac{\bar{\Delta} - \mu_\Delta}{s_\Delta / \sqrt{n}}, \]

where \( s_\Delta^2 \) is the sample variance of the \( \Delta_i \) differences for a total of \( n \) pairs (Wilks, 2011). The p-value indicates at which statistically significant level the null hypothesis is not true. In this case the p-value can be used to evaluate the significance of the difference between MODIS and AERONET.
3.3.2 Likelihood Ratio Test

A likelihood ratio test is a parametric test to determine the likelihood that the MODIS-derived AODs could have been drawn from the same log-normal distribution as the AERONET AODs. To perform this test it is necessary to fit log-normal distributions separately to each MODIS algorithm and AERONET, and compare these two distributions with the single log-normal distribution fit using both sets of data (Wilks, 2011). The general form of the likelihood test statistic is \( \phi^* = 2 \cdot \ln \left( \frac{\varphi(H_A)}{\varphi(H_0)} \right) = 2 \cdot [L(H_A) - L(H_0)] \), where \( \varphi(H_A) \) and \( \varphi(H_0) \) are the likelihood functions and \( L \) is the log-likelihood. For our case the test statistic is equal to \( \phi^* = 2 \cdot \left( \sum_{i=0}^{T} PDF_{MODIS} \right) + \left( \sum_{i=0}^{T} PDF_{AERONET} \right) \) - \left( \sum_{i=0}^{T} PDF_{MODIS and AERONET} \right) \) where the PDFs are a function of \( \mu \), \( \sigma \), and \( \tau \). The test statistic has two degrees of freedom since there are four parameters used to estimate the individual AERONET and MODIS distributions and two for the null hypothesis that MODIS and AERONET AOD data are from the same PDF \( (PDF_{MODIS} and AERONET) \), thus, \( \phi^* \) is evaluated with the \( \chi^2 \) table with two degrees of freedom.

3.3.3 Kolmogorov-Smirnov Test

The K-S test compares the cumulative density functions (CDFs, running integral of PDFs) of each of the MODIS algorithms to that of AERONET. The test statistic is represented by the maximum difference between the MODIS and AERONET CDFs, \( D = \max |CDF_{MODIS} - CDF_{AERONET}| \). When \( D \) is greater than the critical value \( 1.36/\sqrt{n} \),
the null hypothesis that the two CDFs show a good fit is rejected at the 99% confidence level. By analyzing the degree of fit between the MODIS and AERONET PDFs and CDFs, our evaluation goes beyond the bias and correlation tests that have been used commonly in the past to evaluate MODIS AOD accuracy and, hence, provides a more robust statistical technique that is needed to move toward a more complete description of the uncertainty in MODIS AOD retrievals.
4. Results and Discussion

4.1 Coastal vs. Non-Coastal MODIS AOD Evaluation

Although previous MODIS analysis, over a global average, was valuable for understanding MODIS error characteristics (Kahn et al., 2005, 2007, 2011; Remer et al., 2005; Mi et al., 2007; Levy et al., 2007a, 2010), an examination of coastal regions shows a reduction in MODIS accuracy. It is visible that the MODIS and AERONET AOD frequencies are not Gaussian (Figure 1). Indeed, a log-normal distribution is found for MODIS and AERONET AODs over the coastal (Figure 1), non-coastal (Figure A3), and global stations (Figure A4), and the corresponding parameters (μ, σ) for coastal retrievals are shown in Figure 1, with the PDFs shown in Figure 2. Using a χ² test we find that the log-normal PDFs fit each distribution at a statistically significant level (Figure 2).

After removing retrievals with quality flags below 3 for Land and below 1 for Ocean, MODIS AODs are highly correlated with the paired AOD from AERONET with R² of 0.8, regardless if AODs are retrieved over coastal or non-coastal region (Figure 3, Table 3), and/or from Ocean algorithm, Land algorithm, and Land_And_Ocean combined (respectively shown in three columns in Figure 3). Compared with AERONET AOD, MODIS AOD retrievals from the Ocean algorithm have a correlation coefficient (R²) of 0.809 and a regression equation of \( \tau_M = 0.913 \cdot \tau_A + 0.028 \) (where subscripts A and M represent AERONET and MODIS respectively) at global scale (Figure A5, Table 3), R² of 0.804 and \( \tau_M = 0.863 \cdot \tau_A + 0.034 \) for all the coastal sites (Figure 3f, Table 3), and a larger R² of 0.854 and \( \tau_M = 1.115 \cdot \tau_A - 0.001 \) for all non-coastal (e.g., open ocean) sites.
(Figure 3c, Table 3). In comparison, MODIS AODs from the Land algorithm have an $R^2$ of 0.793 and linear regression $\tau_M = 0.979 \cdot \tau_A + 0.008$ for all AERONET stations (Figure A5, Table 3), $R^2$ of 0.795 and $\tau_M = 1.027 \cdot \tau_A + 0.016$ for coastal stations (Figure 3e, Table 3), and $R^2$ of 0.795 and $\tau_M = 0.971 \cdot \tau_A + 0.004$ for non-coastal (inland) sites (Figure 3b, Table 3). In contrast to the Ocean algorithm AOD, there seems to be little change in correlation between the coastal, non-coastal and global evaluations of the Land algorithm AOD. However, the Ocean AOD correlation is consistently greater than the Land AOD correlation in all respective categories.

The AODs over coastal and non-coastal regions retrieved from the Land algorithm both fall within the expected error (EE) envelope greater than 66% of the time (Figure 3b and 3e), but the counterparts from the Ocean algorithm (Figure 3c and 3f) only fall within the EE envelope ~58% of time, which is lower than 66% that is revealed from the past studies of MODIS collection 4 that did not separate the AERONET-MODIS AOD comparisons into coastal and non-coastal regions (Remer et al., 2005). Nevertheless, while the EE for the Ocean algorithm is smaller than that for the Land algorithm, the bias is found to be less (Table 1).

A small bias (often < 0.03) of AOD overall is consistent with past research (Levy et al., 2010; Remer et al., 2005), however, for the same type of product (e.g., from Ocean algorithm, Land algorithm, and combined Land_And_Ocean), a larger bias of AOD is apparent over the coastal regions than over non-coastal regions (Figure 3). It is noted that for AOD from the Ocean algorithm, the overall bias (0.012) along the coast is larger than
the counterparts (0.006) over the open ocean (Figure 3f vs. 3c). Although, this is a bit misleading because of two counteracting effects over the coast where AOD larger (smaller) than 0.25 are underestimated (overestimated) by 0.029 (0.021) (Table 1). The AOD value of 0.25 was chosen as the cutoff because Levy et al. (2010) suggest that regardless of quality flag, AOD less than ~0.2 may represent an aerosol signal that is too low to retrieve meaningful aerosol size information from MODIS, and the Ocean regressions are found to cross the 1:1 line around an AOD of 0.25. Using a t-test for difference we find that regardless of the MODIS product (i.e., Ocean, Land, Land_And_Ocean), the AOD biases over coastal regions are statistically significant beyond the 99% confidence level with a p-value much less than 0.01.

The Land algorithm has a significantly larger bias than the Ocean algorithm for most coastal AERONET sites (Figure 4). This is expected because of the inherent difficulties in characterizing land surfaces in general. The average MODIS AOD bias for the Land algorithm over coastal sites is 0.026 and shows little dependence on AOD amount (Table 1); again using a t-test for difference we find p-values much less than 0.01 revealing that the bias in AOD retrieval form Land algorithm is statistically significant. However, the bias results show large variations amongst different coastal AERONET sites (Figure 4).

The Taylor Diagram (Figure 5) visualizes the overall performance of the MODIS aerosol algorithms in a single figure. The MODIS correlation coefficient with AERONET decreases for coastal retrievals compared to non-coastal retrievals (Figure 5).
Furthermore, the normalized standard deviations of MODIS AOD increase from roughly 0.8 for non-coastal retrievals to 1.3 for coastal retrievals (Figure 5), indicating that MODIS AOD is less capable of capturing the temporal variation of AERONET AOD over the coastal sites. By the same token, the Ocean algorithm over the open ocean (non-coastal) captures the variation in AOD better than the other algorithms (Figure 5), because its resultant representation in the Taylor diagram is closest to the point “REF” indicating the best performance with respect to AERONET. It is shown that all of the MODIS AOD retrievals over the coast, regardless of algorithm, cluster farthest away from the “REF” point, indicating a need for refinement of the MODIS product over coastal regions (Figure 5).

To further evaluate if MODIS aerosol products represent the statistics observed from AERONET, the PDFs from the MODIS products are used to compare against the PDF from AERONET to determine the fit of the AOD frequencies. The t-test for difference of mean, described in Section 3.3.1, reveals p-values of much less than 0.01 for the difference of each of the MODIS AOD products to AERONET, demonstrating that at a statistically significant level the mean of MODIS AOD products are different from the mean of paired AERONET AOD over the coast. Going beyond the t-test for difference of mean we apply the likelihood ratio test and the K-S test for goodness-of-fit.

The likelihood test returns a test statistic $\phi$ as described in Section 3.3.2. The test statistic is compared to a critical value to determine the likelihood that the MODIS AOD PDF fits the PDF from AERONET AOD. The critical value for the $\chi^2$ statistics with two
degrees of freedom at the 99% confidence level is 9.210, where anything greater than this value results in rejecting the null hypothesis that the PDFs may come from the same distribution. We find that the test statistics are 23.03, 29.77, and 22.98 for the quality filtered MODIS Land, Ocean, and Land_And_Ocean products, respectively. Therefore, we conclude that the PDFs from the MODIS algorithms statistically differ from the PDFs from AERONET over coastal regions. The likelihood test is useful to compare PDFs. However, to more fully describe the fit between MODIS and AERONET data, our analysis is extended to CDFs as well.

The results of the K-S test and maximum difference for the CDFs from each quality filtered MODIS algorithms to the CDF from AERONET (Figure 6). The critical values (described in Section 3.3.3) needed to reject the null hypothesis that the MODIS Land, Ocean, and Land_And_Ocean AOD CDFs fit the counterpart of the AERONET AOD, at a 99% confidence level, are respectively 0.013, 0.009, and 0.008 (Figure 6). If the maximum difference is greater than the respective critical values then the null hypothesis must be rejected and the CDFs are significantly different. It is clear in Figure 6 that the maximum departures of the CDFs from each of the MODIS AOD products differs from the counterpart of the AERONET AOD CDF, indeed all are greater than the corresponding critical values (Figure 6). Therefore, it is concluded that the CDFs from each of the MODIS algorithms do not match the AERONET CDF at a statistically significant level; in fact, they differ beyond the 99% confidence level. This difference
suggests that MODIS does not accurately estimate AOD over the coastal regions and leaves room for improvement in the operation aerosol algorithms.

4.2 Land_And_Ocean Quality Control

The Land_And_Ocean product reported by the MODIS science team is created by merging (a simple union) retrievals from the MODIS Land algorithm and the MODIS Ocean algorithm to provide a more complete spatial coverage of AOD over the globe (Levy et al., 2010). However, the quality flag that is reported with each retrieval is not carried through to the Land_And_Ocean product in the combination process. Using the mean method to collocate the Land_And_Ocean product with AERONET renders difficulties in quality control because regions within the 55 km diameter around the AERONET site may contain retrievals from both the Land algorithm and the Ocean algorithm (Petrenko et al., 2012) (Figure A6). Thus, a more robust method is required to filter the MODIS Land_And_Ocean data using only “high quality” retrievals based upon the original quality flags. In other words, AODs reported in the Land_And_Ocean data set originating from the Land algorithm must have a quality flag of 3 in order to be considered a “high quality” retrieval for the evaluation, while those originating from the Ocean algorithm must have a flag of greater than or equal to 1 (Figure A6). The meaning of each quality flag category is explained above in Section 2. Note that when using the mean collocation method, AOD reported by the Land_And_Ocean product may not be equal to the AOD reported by either the Land or Ocean algorithms because the average of the Land_And_Ocean product within the 55 km region around AERONET may include
MODIS boxes originating from the Ocean and Land algorithms. It follows that the quality flags associated with the Land_And_Ocean product may originate from the Land and/or Ocean retrievals as well. Thus, the filtered Land_And_Ocean collocation with AERONET uses only those “high quality” retrievals from either or both the Land and Ocean algorithms within the 55 km diameter (described here in Section 2 and also in Petrenko et al., 2012).

The QA filtering as described above improves the global MODIS Land_And_Ocean correlation ($R^2$) with AERONET from .737 to .804 (Figure A5, Table 3), and reduces the AOD bias by 34% for coastal regions from 0.029 to 0.019 (Table 1). Focusing on the high AOD events (AOD > 0.25) over the coast the reduction in bias is even greater at 62% from 0.026 to 0.010 (Table 1). Bias values were evaluated using a t-test for difference and are found to have p-values much less than 0.01 and, therefore, are statistically significant. However, the number of MODIS-AERONET pairs is reduced from 113,152 to 71,303 globally (or by 37%) after applying the quality flag filter. For the purposes of this research, the increase in accuracy outweighs the reduction in data. The Land_And_Ocean, quality assured, data set has a regression equation of $\tau_M = 0.964 \cdot \tau_A + 0.014$ on a global scale over the full record of MODIS (Figure A5, Table 3), and over coastal regions has an $R^2$ of 0.818, and a regression of $\tau_M = 0.933 \cdot \tau_A + 0.028$ (Figure 3, Table 3). Furthermore, the t-test for difference reveals that the closest or statistically best fit to the AERONET data over the coast is observed from the quality filtered Land_And_Ocean product, with p-values changing from $\sim 1 \times 10^{-5}$ to $\sim 1 \times 10^{-6}$.
after the filter. The reduction in bias from the quality filter can be observed in Figure 4d vs. 4e and an increase in correlation is found on a global scale. However, even after the quality flag filter, the coastal regions still show poorer MODIS performance compared to the non-coastal retrievals (Figure 5). The result suggests that a dedicated algorithm for coastal retrievals may be needed in lieu of the current Land and Ocean algorithms used for MODIS aerosol retrievals.

4.3 Wind and Cloud Impact on the MODIS Ocean Algorithm

Different sources of error arise in the MODIS Ocean retrievals because of surface characteristic assumptions made to create the LUTs used by the algorithm, and the validity of a pixel for the AOD retrieval depends on the result of cloud-mask algorithm designed specifically for the MODIS Ocean product. We examine the potential sources of error separately to more accurately describe the MODIS performance over the coastal regions. We focus on the possible cloud contamination described by Zhang and Reid (2006) and the surface wind speed assumption of 6 m s\(^{-1}\) in the MODIS Ocean algorithm.

It is known that cloud contamination can affect MODIS AOD accuracy over the ocean (Zhang and Reid, 2006) for the collection 4 product. We investigate the possible cloud contamination for the collection 5 MODIS product for coastal waters by using the MODIS reported cloud fraction for over-ocean retrievals. Multiple thresholds for cloud fraction are tested and Table 2 contains results from the 80% threshold, 70% threshold, and the standard quality flag filtered MODIS product. The threshold analysis reveals that the 70% threshold can greatly reduce bias while maintaining a sufficient number of
retrievals, with the number of retrievals reduced by only 16% globally and 14% over coastal regions (Table 2). For the cloud fraction threshold of 70% (80%) the reduction of bias for coastal sites is 100% (67%) and for non-coastal sites is 58% (33%) (Table 2). At first glance it may seem the filter removes all bias from coastal retrievals, however, this is not the case. MODIS bias for AOD events greater than 0.25 actually increases by 21% for both the 80% and 70% cloud fraction filters.

While Table 2 shows results consistent with those of Zhang and Reid (2006) who showed that the removal of MODIS over-ocean AODs where pixels correspond to a cloud fraction larger than a threshold of 80% can significantly reduce the errors in AOD estimates, a more detailed examination also shows that the cloud fraction filter leads to an even more negative bias for AOD events over 0.25 but reduces the positive bias for AOD events less than 0.25 (Table 2). Zhang and Reid (2006, 2010) demonstrate that the cloud contamination causes MODIS overestimation due to the high reflectivity of clouds in the visible spectrum and, therefore, filtering AOD retrievals by cloud fraction would lead to an overall decrease in MODIS AOD. The same physical interpretation is true for MODIS collection 5; however, the negative bias persistent for AOD over 0.25 requires another explanation.

In addition to cloud contamination, error may be introduced into the MODIS Ocean algorithm by inherent assumptions in calculating the LUTs, specifically, the assumption of a constant surface wind speed of 6 m s\(^{-1}\) (Zhang and Reid, 2010; Shi et al., 2011; Kleidman et al., 2012). Kleidman et al. (2012) show a systematic increase of
MODIS error as a function of wind speed for retrievals over the open ocean. This dependence is most apparent when wind speed deviates from the 6 m s\(^{-1}\) speed assumed for the rough ocean surface and white cap parameterizations within the MODIS AOD retrievals (Kleidman et al., 2012). To our knowledge, the impact of varying wind speeds on the MODIS retrieval accuracy over coastal waters has not been studied, although previous work on wind climatologies suggests that surface wind speeds over coastal regions are frequently slower than 6 m s\(^{-1}\) (Lavagnini et al., 2005; Martin et al., 1999; Maryland Department of Natural Resources, 2011). To quantify the impact of the surface wind speed on coastal aerosol retrievals, we stratify the analysis of MODIS-AERONET biases (before and after cloud-contamination filtering) as a function of ocean surface wind speed. At every coastal AERONET site, each MODIS AOD bias is paired spatially and temporally with the corresponding 2 meter wind speed from the MERRA re-analysis.

We first conduct the correlation analysis for MERRA 2 meter wind speed and the bias in MODIS AOD without filtering by cloud fraction (but with QA flag assured). A positive correlation between wind speed and MODIS AOD bias for all coastal AERONET sites is shown in Figure 8, with a linear best fit of \(\tau_{\text{bias}} = 0.010 \cdot v - 0.020\) before cloud filtering, where \(\tau_{\text{bias}}\) is MODIS bias and \(v\) is wind speed. The positive correlation between bias and wind speed can be understood from the following two factors: (1) wind speeds over coastal regions are frequently (94% of the time) less than 6 m s\(^{-1}\) at the time of MODIS retrievals (Figure 8b), and (2) slower wind speeds lead to more negative MODIS bias while faster wind speeds lead to positive bias (Figure 8).
While factor (1) explains, in part, the negative bias for the AOD (greater than 0.25) retrieved from the Ocean algorithm, factor (2) can be used to interpret the overestimation in MODIS AOD for AOD less than 0.25 over the coast. High AOD near the coast presumably occurs during high wind conditions that can generate more sea salt particles or can be associated with frontal passage moving aerosols from land to ocean; in either case, such high winds, if they are larger than 6 m s\(^{-1}\), can lead to overestimation in MODIS AOD retrievals.

The overall MODIS AOD bias is reduced from 0.006 with the quality flag filter alone to 0.004 (33%) after adding the wind speed correction (Table 4). Furthermore the low AOD (less than 0.25) bias is reduced by 33% from 0.021 to 0.014, and the high AOD (greater than or equal to 0.25) bias is reduced from -0.029 to -0.020 a 31% reduction (Table 4). However, with the known impact of cloud contamination, we conduct a similar analysis after filtering out the retrievals with cloud fractions greater than 70%.

The linear regression equation between wind speed and MODIS bias, after accounting for cloud contamination (Figure 8c), is then used as an empirical correction for the MODIS Ocean algorithm retrievals over the coast. The linear regression found for the MODIS bias dependence on wind speed is \( \tau_{\text{bias}} = 0.010 \cdot v – 0.024 \), after applying the 70% cloud fraction (Figure 8c). By including the wind speed at the time of each MODIS AOD retrieval, we can estimate the MODIS bias by using the \( \tau_{\text{bias}} \) equation and subtract it from the reported AOD to improve the AOD accuracy. The empirical correction yields a reduction in overall MODIS AOD bias for the Ocean algorithm over the coast from
+0.006 for the standard quality flag filtered MODIS product to -0.0005 (Table 4) for the cloud and wind corrected AOD. Furthermore, for AOD events less than 0.25 the bias is reduced from +0.021 to +0.0098 (Table 4), and for AOD events greater than 0.25 the bias is reduced from -0.029 to -0.027 (Table 4). In addition to reducing bias, the number of retrievals that fall within the EE is increased over the coast from 58% to 63%. Although this is still below expected 66%, it is a marked improvement. The success of the correction scheme suggests that bias introduced into AOD retrievals by wind speed assumptions can be reduced by using real or modeled wind speed data. Along with the reduced bias, the correlation and best fit regression show a better fit to the AERONET AOD data for the corrected results when compared to the original MODIS Ocean product (Figure 5). Figure 5 is particularly useful for observing the benefits from the correction scheme because it shows the shift (represented by the black arrow) from the QA flag assured Ocean product to the corrected Ocean product. From Figure 5 it is immediately clear that the correction improves the MODIS AOD correlation with AERONET and reduces the variance in observation, indicating that the temporal variation is better captured by the corrected product. Furthermore, after both cloud and wind correction the MODIS frequency shows a better fit to the AERONET distribution than the standard MODIS Ocean product (Figure 9). Although the corrected MODIS AOD CDF does not pass the K-S test with a maximum difference of 0.024 and a critical value of 0.011 at the 99% confidence level, it does show an improvement with a reduction in maximum
difference from the AERONET CDF over the standard MODIS product and the cloud fraction filter alone (Figure 9).

For high AOD events (greater than 0.25) the MODIS bias is reduced by applying the empirical wind speed correction. However, because the cloud fraction filter results in a more negative trend to the already negative MODIS bias, the wind speed correction is less visible. A possible cause of the more negative bias after cloud filtering is that cloud contamination has a greater affect, proportionally, for low AOD retrievals than for higher AODs. The increased reflection from cloud contamination has a proportionally reduced impact on total reflection (Kleidman et al., 2012). Thus, the cloud contamination filter removes some of the high AOD events that are minimally impacted by high cloud fractions, and may skew the results to a more negative bias. This impact will be evaluated in future studies.

An assessment of the covariance between cloud fraction and wind speed is provided in Figure 8d. The scatter plot of wind speed and cloud fraction for coastal retrievals clearly shows that the wind speed and cloud fraction are not correlated (Figure 8d) and, therefore, a correction scheme that accounts for each independently, such as above, is an appropriate method.

To further evaluate the effectiveness of our empirical method to correct the AOD bias due to assumed wind speed in the MODIS algorithm, we show the correlation between the MERRA 2 meter wind speed and MODIS AOD bias, as well as the slope and intercept from the best linear fit between the two at each coastal AERONET site.
Most of the coastal AERONET sites (46 out of the possible 62) show a statistically significant correlation between MODIS bias and wind speed (Figure 7b). Furthermore, from those statistically significant sites, 40 are found to have a negative MODIS bias as the wind speed approaches zero (Figure 7c), and 45 are found to have a regression with a positive slope that indicates a systematic positive bias in MODIS AOD as wind speeds increase (Figure 7d). Due to the high proportion of sites showing a statistically significant dependence on wind speed, a generalization is extended to all coastal sites.
5. Summary and Conclusions

Nearly 9.5 years of Aqua-MODIS aerosol retrievals are evaluated using AERONET aerosol measurements for validation. Specific focus in the analysis is given to the coastal regions of the world due to their complex surface characteristics and their dominant contribution to the loading of anthropogenic aerosols in the atmosphere. Over the coast the MODIS aerosol algorithms show increased uncertainty with respect to non-coastal regions. It is shown that the quality flags reported by the MODIS Land and Ocean algorithms should be applied to the MODIS Land_And_Ocean aerosol product to reduce the statistically significant difference and improve the PDF match between MODIS AOD and AERONET aerosol observations. After filtering by quality flags, the MODIS AODs respectively retrieved from the Land and Ocean algorithms are highly correlated with AERONET (with $R^2 \approx 0.8$), but only the Land algorithm AODs fall within the expected error envelope greater than 66% of the time. Furthermore, MODIS AODs after quality flag filtering, regardless from Land algorithm, Ocean algorithm, or combined Land_And_Ocean product, show statistically significant discrepancies with the respective counterparts from AERONET in terms of mean, probability density function, and cumulative density function, which suggests the need for improvement in coastal retrieval algorithms.

The MODIS over-ocean algorithm is found to have two main sources of error over coastal regions. The first is cloud contamination that leads to an overestimation of AOD, and this result is in agreement with Zhang and Reid (2006, 2010) in their global
analysis without separation of assessment of AOD accuracy between coastal and non-coastal sites. The second source of error is dependent on near-surface wind speed due to the assumptions made in the MODIS AOD retrievals over ocean. Based upon MERRA data, we found that wind speeds over the coastal ocean are frequently slower than the 6 m s\(^{-1}\) assumed by the MODIS Ocean algorithm. This high bias in wind speed pre-described in the MODIS Ocean algorithm often leads to an overestimation of the surface reflectance contribution to the radiance measured by the satellite at the top of atmosphere and so an underestimation of the aerosol reflectance contribution, leading to an underestimation of AOD. After applying the equation \(\tau_{\text{bias}} = 0.010 \cdot v - 0.024\), where \(\tau_{\text{bias}}\) is estimated MODIS bias and \(v\) is surface wind speed, to the MODIS AOD retrievals from the Ocean algorithm, it is found that the corresponding AOD bias is reduced from +0.006 to -0.0005, which should be useful for future analysis of AOD trend and MODIS AOD data assimilation.

The results of this research can also be used to improve the input into global climate models, providing a more accurate assessment of the effects aerosols have on climate change. Previous AOD trend analyses have found AOD trends to be on the order of 0.01 decade\(^{-1}\) or less (Mishchenko et al., 2007; Zhang and Reid, 2010). Understanding these trends may have important implications in understanding both aerosol emission and also climate forcing from aerosols. With such small trends in AOD, the reduced bias gained from our empirical correction may become increasingly important to the evaluation of regional AOD trends.
Acknowledgements

Funding for this research was provided by NASA GSFC, NASA Nebraska Space Grant Consortium, and the Department of Earth and Atmospheric Sciences at the University of Nebraska-Lincoln. I would like to thank Gregory Leptoukh, who tragically passed away during this research. Without his vision and foresight this project would have never been possible. His efforts in data quality and data fusion not only made this work a success but provide a framework for numerous research opportunities in the field of remote sensing. I would also like to thank the faculty, staff, and fellow students in the Earth and Atmospheric Science department at the University of Nebraska-Lincoln for their support in graduate school and with this project. Finally, I would like to thank my sister and parents for their support in all facets of my life.
References


IPCC: Climate Change 2007: The physical science basis. Contribution of working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate


Maryland Department of Natural Resources: Maryland Offshore Wind Climatology with Application to Wind Power Generation January 21, 2011, State Report.


Shi, Y., Zhang, J., Reid, R., et al.: An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation,


Zhang, J., and Reid, J.: 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, Atmospheric Chemistry and Physics, 10, 18879-18917, doi:10.5194/acp-10-10949-2010, 2010.

Tables

Table 1. Aqua-MODIS AOD mean bias over the full data record (2002-2010) for all AERONET coastal stations. 62 coastal AERONET sites were identified and the results are an average of all the sites. Each of the MODIS aerosol algorithms are shown with the recommended quality control except for the Land_And_Ocean product which is shown without any quality control (default MODIS product) and the results of our quality control technique. Bias results are separated into Low AOD and High AOD events as classified by AERONET measurements with the cutoff at 0.25.

<table>
<thead>
<tr>
<th>All Coastal Sites</th>
<th>Land Algorithm</th>
<th>Ocean Algorithm</th>
<th>Land_And_Ocean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QA Filtered</td>
<td>QA Filtered</td>
<td>No Filter</td>
</tr>
<tr>
<td>Total Bias</td>
<td>0.026</td>
<td>0.006</td>
<td>0.029</td>
</tr>
<tr>
<td>Low AOD Bias</td>
<td>0.024</td>
<td>0.021</td>
<td>0.033</td>
</tr>
<tr>
<td>High AOD Bias</td>
<td>0.026</td>
<td>-0.029</td>
<td>0.026</td>
</tr>
</tbody>
</table>
Table 2. Aqua-MODIS Ocean algorithm AOD bias with respect to AERONET AOD for both coastal and open ocean sites. The bias is listed for three categories on how MODIS AOD is used in the evaluation. The first is filtering of data with quality control flag; the second builds upon the first but also removes MODIS AOD data with cloud fraction larger than 80%; the third is the same as second except the threshold for cloud fraction is now decreased to 70%. The number of AOD retrievals used in the different analyses (last row in Table 2) is also shown to display the reduction in data size associated with each category. In each category, bias is further analyzed in terms of low AOD conditions (AOD < 0.25) and high AOD conditions. In addition, the relative percent change of bias due to the filtering of data with cloud fraction is shown in in parentheses, negative percentages indicate an increase in bias. See text for further details.

<table>
<thead>
<tr>
<th>Aqua-MODIS Cloud Contamination</th>
<th>Normal QA</th>
<th>80% Threshold</th>
<th>70% Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coastal</td>
<td>Open Ocean</td>
<td>Coastal</td>
</tr>
<tr>
<td>Total Bias</td>
<td>0.006</td>
<td>0.012</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(67%)</td>
</tr>
<tr>
<td>Low AOD Bias</td>
<td>0.021</td>
<td>0.018</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(14%)</td>
</tr>
<tr>
<td>High AOD Bias</td>
<td>-0.029</td>
<td>-0.022</td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-21%)</td>
</tr>
<tr>
<td>Number of Retrievals</td>
<td>18,001</td>
<td>4,190</td>
<td>17,104</td>
</tr>
</tbody>
</table>
Table 3. Regression statistics of for the MODIS AOD products with respect to AERONET. Data span 2002-2011.

<table>
<thead>
<tr>
<th>Regression Statistics</th>
<th>Land Coastal</th>
<th>Land Non-Coastal</th>
<th>Land Global</th>
<th>Ocean Coastal</th>
<th>Ocean Non-Coastal</th>
<th>Ocean Global</th>
<th>Land_And_Ocean No QA Coastal</th>
<th>Land_And_Ocean No QA Non-Coastal</th>
<th>Land_And_Ocean No QA Global</th>
<th>Land_And_Ocean With QA Coastal</th>
<th>Land_And_Ocean With QA Non-Coastal</th>
<th>Land_And_Ocean With QA Global</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.795</td>
<td>0.795</td>
<td>0.793</td>
<td>0.804</td>
<td>0.854</td>
<td>0.809</td>
<td>0.753</td>
<td>0.73</td>
<td>0.737</td>
<td>0.818</td>
<td>0.801</td>
<td>0.804</td>
</tr>
<tr>
<td>R²</td>
<td>0.863</td>
<td>1.115</td>
<td>0.948</td>
<td>0.968</td>
<td>0.962</td>
<td>0.933</td>
<td>0.992</td>
<td>0.964</td>
<td>0.964</td>
<td>0.964</td>
<td>0.964</td>
<td>0.964</td>
</tr>
<tr>
<td>Slope</td>
<td>1.027</td>
<td>0.971</td>
<td>0.979</td>
<td>0.863</td>
<td>1.115</td>
<td>0.913</td>
<td>0.948</td>
<td>0.968</td>
<td>0.962</td>
<td>0.933</td>
<td>0.982</td>
<td>0.964</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.016</td>
<td>0.004</td>
<td>0.008</td>
<td>0.034</td>
<td>-0.001</td>
<td>0.028</td>
<td>0.037</td>
<td>0.026</td>
<td>0.03</td>
<td>0.029</td>
<td>0.003</td>
<td>0.014</td>
</tr>
</tbody>
</table>
Table 4. Bias result of corrections to the MODIS Ocean algorithm. Normal QA is quality flag filter only, 70% Cloud Filter is quality flag plus 70% cloud fraction filter, Wind Correction is the quality flag filter plus wind speed correction, Cloud and Wind Correction is quality flag filter plus the 70% cloud fraction filter plus the empirical wind speed correction. Values in parenthesis are percent reduction in bias from the Normal QA baseline.

<table>
<thead>
<tr>
<th>Aqua-MODIS Correction</th>
<th>Normal QA</th>
<th>70% Cloud Filter</th>
<th>Wind Correction</th>
<th>Cloud and Wind Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Bias</td>
<td>0.006</td>
<td>0.000 (100%)</td>
<td>0.004 (33%)</td>
<td>-0.0005 (92%)</td>
</tr>
<tr>
<td>Low AOD Bias</td>
<td>0.021</td>
<td>0.016 (24%)</td>
<td>0.014 (33%)</td>
<td>0.0098 (53%)</td>
</tr>
<tr>
<td>High AOD Bias</td>
<td>-0.029</td>
<td>-0.035 (-21%)</td>
<td>-0.020 (31%)</td>
<td>-0.027 (7%)</td>
</tr>
<tr>
<td>Number of Retrievals</td>
<td>18,001</td>
<td>15,768</td>
<td>18,001</td>
<td>15,768</td>
</tr>
</tbody>
</table>
**Figures**

![Figure 1](image)

**Figure 1.** Frequency of coastal AOD (left vertical axis) and relative frequency of AOD (right vertical axis) at AERONET sites over the ~9 year period from 2002-2011. Plots are derived from AOD data at 62 coastal AERONET sites and collocated Aqua-MODIS retrievals over these sites. $\mu$ is the log-normal location parameter and $\sigma$ is the log-normal scale parameter, mean is the average AOD over the whole time period. Quality assured and quality flag filtered frequencies of AOD from (A) AERONET, (B) Land_And_Ocean product, (C) AERONET AODs paired with only the MODIS Land algorithm, (D) Land algorithm, (E) AERONET AODs paired with only the MODIS Ocean algorithm, and (F) Ocean algorithm.
Figure 2. Probability density functions (right vertical axis) and relative frequency (left vertical axis) of the coastal AODs from (A) AERONET, (B) MODIS Land_And_Ocean, (C) MODIS Ocean algorithm, and (D) MODIS Land algorithm. All MODIS algorithms are quality flag filtered and the data span from 2002-2011. The p-values indicate statistical significance of fitness between frequency distributions derived from PDFs (e.g., product of total number of data points, PDFs, and AOD bin interval) and actual frequency distribution (e.g., the bars in red). (A) Has only those AOD observations when there was a corresponding MODIS AOD retrieval.
Figure 3. Scatter plot of AERONET AOD (x-axis) and the quality flag filtered Aqua-MODIS AOD (y-axis) from 2002-2011. In (A), (B), and (C), AODs in y-axis are respectively derived from MODIS Land_And_Ocean, Land, and Ocean products over the non-coastal AERONET stations. (D) (E) and (F) are respectively the same as (A), (B), and (C) but over the coastal AERONET stations. In each scatter plot, also shown is the correlation coefficient ($R^2$), mean bias, the number of MODIS-AERONET collocated data points (N), and the best-fit linear regression equation (solid black line), the 1:1 line (dashed black line), and the bounds (red dashed line) for MODIS AOD variation after consideration of MODIS AOD expected error (EE). See details in Section 2.1.
Figure 4. (A) Map of the location of all coastal AERONET sites; (B)-(E) show the maps of MODIS AOD bias (with respect to AERONET AOD) at each coastal AERONET site respectively for: (B) MODIS Land AOD product filtered with quality flag, (C) MODIS Ocean AOD product filtered with quality flag; (D) MODIS Land_And_Ocean AOD product without any quality filtering; (E) MODIS Land_And_Ocean aerosol product after using the method described in the Section 4 for quality filtering. Bias calculations are based on ~9 years (2002-2011) of collocated Aqua-MODIS and AERONET AOD data. Blue indicates MODIS underestimation of AOD (e.g., negative bias) and red is overestimation (positive bias). The area of the circles is proportional to bias magnitude.
Figure 5. Taylor diagram comparing 2002-2011 quality flag filtered MODIS AOD retrievals and AERONET AOD observations. The MODIS Ocean, Land, Land_And_Ocean, and empirically corrected Ocean products are represented by dark blue, red, green, and light blue respectively. Coastal MODIS AOD retrievals are listed with a 1 and Non-Coastal AODs are shown with a 2. The arrow represents the effect of the empirical correction on the MODIS Ocean product.
Figure 6. Cumulative density functions (CDF) of AOD derived from AERONET (black), and corresponding paired MODIS AODs respective derived from MODIS Land (red), Ocean (blue), and Land_And_Ocean (green) products after filtering with quality flag. These CDFs are based upon the log-normal distributions with parameters shown in Figure 2. Maximum differences ($\Delta_{\text{max}}$) between the AERONET CDF and Aqua-MODIS CDFs are shown by two dashed lines in their respective colors. Data are from Aqua-MODIS from 2002-2011 over coastal regions. Critical values described in Section 3.3.3.
**Figure 7.** (A) Station locations for each coastal AERONET site, (B) the correlation between quality flag filtered Aqua-MODIS Ocean algorithm AOD bias and sea-surface wind speed, (C) the y-intercept from the linear regression of bias and wind speed, (D) the slope from the linear regression of bias and wind speed. Blue colors represent statistically significant values in (B) and negative intercepts and slopes for (C) and (D), respectively. Red represents statistically insignificant values in (B) and positive intercepts and slopes for (C) and (D), respectively. Magnitude scales are provided in each panel for clarity. (C) and (D) show only sites with p-value less than or equal to 0.05 (46 out of the possible 62 sites). Data are from 2002-2011.
Figure 8. (A) Scatter plot of quality flag filtered MODIS Ocean AOD bias against the 2 m wind speed from MERRA, (B) the frequency of coastal wind speeds, (C) same as (A) but for the bias of MODIS AOD after 70% cloud fraction filter, (D) scatter plot of the wind speed and cloud fraction pairs for each retrieval. The analysis is for all coastal sites (62 AERONET sites) and for the years 2002-2011. R is the Pearson linear correlation coefficient, N is the number of retrievals and Y is the regression equation. In (B), the right vertical axis is the cumulative density function for the coastal wind speeds (represented by the black curve). In (A) and (C) red is MODIS bias binned to 1 m s⁻¹ intervals along with regression and correlation corresponding to those bins, and the blue dotted line is a reference to 0 MODIS bias.
Figure 9. Frequency distribution of (A) AERONET AOD over coastal regions that have an Aqua-MODIS Ocean algorithm collocated retrieval, (B) AOD from Aqua-MODIS Ocean algorithm after cloud fraction filtering (70%) and wind speed bias correction, (C) AOD from Aqua-MODIS Ocean algorithm after cloud fraction and quality flag filtering only. (D) Cumulative density functions (CDF) derived from the frequency distributions respectively in (A)-(C), along with their respective maximum difference ($\Delta_{\text{max}}$) from the AERONET CDF. Quality flag filtering is applied for all algorithms and the Aqua-MODIS data span 2002-2011. The critical values and K-S test from panel (D) are described in Section 3.3.3 and Section 5 in the text.
Appendix A

Figure A1. MODIS spatial resolution as a function of viewing zenith angle. The region bounded by the dark “bow-tie” is 10 km along the flight direction at the center, and is 2330 km across (along the scan direction).
Figure A2. Schematic of the mean collocation method from MAPSS (http://giovanni.gsfc.nasa.gov/mapss/).
Figure A3. Frequency of non-coastal AOD (left vertical axis) and relative frequency of AOD (right vertical axis) at AERONET sites over the ~9 year period from 2002-2011. μ is the log-normal location parameter and σ is the log-normal scale parameter, mean is the average AOD over the whole time period. Quality assured and quality flag filtered frequencies of AOD from (A) AERONET, (B) Land_And_Ocean product, (C) Land algorithm, (D) Ocean algorithm.
Figure A4. Frequency of global AOD (left vertical axis) and relative frequency of AOD (right vertical axis) at AERONET sites over the ~9 year period from 2002-2011. $\mu$ is the log-normal location parameter and $\sigma$ is the log-normal scale parameter, mean is the average AOD over the whole time period. Quality assured and quality flag filtered frequencies of AOD from (A) AERONET, (B) Land_And_Ocean product, (C) Land algorithm, (D) Ocean algorithm.
Figure A5. Scatter plots for all AERONET stations where AERONET AOD is the x-axis and the quality flag filtered Aqua-MODIS AOD is the y-axis. Data span 2002-2011. In (A), (B) AODs are derived respectively from the Land algorithm and the Ocean algorithm. In (C) and (D) AODs are derived from the Land_And_Ocean product where (C) is not quality flag filtered and (D) is after quality flag filtering. Shown in each panel are the correlation coefficient ($R^2$), mean bias, number of MODIS-AERONET collocated data points (N), and the best-fit linear regression equation (solid black line), the 1:1 line (dashed blue line), and the expected error (EE) bounds (red dashed line) for MODIS AOD Land and Ocean algorithms in (A) and (B) respectively.
Figure A6. An example of the quality filter for the Land_And_Ocean MODIS product over a coastal site with both Ocean and Land algorithm AOD retrievals. Numbers represent quality flags associated with Land and Ocean algorithm retrievals. Red numbers indicate all retrievals greater than or equal to the mode of the quality flag separately for each retrieval algorithm (Land and Ocean). (A) is prior to filter and (B) is after filter. An average of all AODs represented by red quality flags are used to represent the Land_And_Ocean AOD for a scene.