

A spatio-temporal approach for global validation and analysis of MODIS aerosol products

Charles Ichoku,¹ D. Allen Chu,¹ Shana Mattoo,¹ Yoram J. Kaufman,²
Lorraine A. Remer,² Didier Tanré,³ Ilya Slutsker,¹ and Brent N. Holben⁴

Received 20 March 2001; revised 13 December 2001; accepted 27 December 2001; published 29 June 2002.

[1] With the launch of the MODIS sensor on the Terra spacecraft, new data sets of the global distribution and properties of aerosol are being retrieved, and need to be validated and analyzed. A system has been put in place to generate spatial statistics (mean, standard deviation, direction and rate of spatial variation, and spatial correlation coefficient) of the MODIS aerosol parameters over more than 100 validation sites spread around the globe. Corresponding statistics are also computed from temporal subsets of AERONET-derived aerosol data. The means and standard deviations of identical parameters from MODIS and AERONET are compared. Although, their means compare favorably, their standard deviations reveal some influence of surface effects on the MODIS aerosol retrievals over land, especially at low aerosol loading. The direction and rate of spatial variation from MODIS are used to study the spatial distribution of aerosols at various locations either individually or comparatively. This paper introduces the methodology for generating and analyzing the data sets used by the two MODIS aerosol validation papers in this issue. *INDEX*

TERMS: 0305 Atmospheric Composition and Structure: Aerosols and particles (0345, 4801); 1610 Global Change: Atmosphere (0315, 0325); 1640 Global Change: Remote sensing; 0394 Atmospheric Composition and Structure: Instruments and techniques

1. Introduction

[2] Aerosol physical parameters are among the numerous products operationally retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Earth Observing System's (EOS) Terra satellite launched on 18 December 1999. The MODIS aerosol algorithm routinely retrieves aerosol optical thickness (AOT or $\tau_{a\lambda}$) at 0.47 and 0.66 μm wavelengths (and interpolates at 0.55 μm) over vegetated land surfaces, and at seven spectral bands (0.47, 0.55, 0.66, 0.87, 1.2, 1.6, and 2.1 μm) over oceans [Kaufman *et al.*, 1997; Tanré *et al.*, 1997]. Other essential parameters retrieved include Angstrom exponent over land and ocean, and effective radius and ratio of small to large size modes over ocean [Chu *et al.*, 2002; Remer *et al.*, 2002].

[3] MODIS aerosol products are validated with ground-based sunphotometer data, particularly those of the Aerosol Robotic Network (AERONET), comprising automatic tracking Sun photometers/sky radiometers located at over 100 sites around the world [Holben *et al.*, 1998]. AERONET sunphotometers derive

AOT at 0.34, 0.38, 0.44, 0.50, 0.67, 0.87, and 1.02 μm wavelengths from direct solar radiation measurements. These data (<http://aeronet.gsfc.nasa.gov/>) undergo cloud screening (Level 1.5) and quality assurance (Level 2.0) [Smirnov *et al.*, 2000], although for near real-time data availability, we use the Level 1.5 data for our validation. AERONET data are very widely used for various aerosol-related studies including satellite retrieval validation [e.g. Zhao *et al.*, 2002].

[4] The objective of this paper is to present the data structure developed for comprehensive and rapid global validation of MODIS aerosol products at near real time. A general assessment of the spatial characteristics of the products is also presented. Specific validations of the MODIS aerosol inversion schemes for retrievals over land and ocean are discussed in separate papers in this issue [Chu *et al.*, 2002; Remer *et al.*, 2002], which also show validation results for different geographical areas where different aerosol types (urban, industrial, marine, and smoke) are dominant.

2. Validation Database Design Concept

[5] Whereas MODIS achieves an almost complete global coverage once or twice a day, sunphotometer retrievals (example, AERONET) occur several times a day, but only over individual instrumented locations. The MODIS Level 2 aerosol products are raster data sets with a pixel size of 10×10 Km. AERONET AOT data are acquired at 15-minute intervals on the average. It would be incongruous to compare single MODIS pixel values directly to AERONET point measurements for many reasons. First, the parameter value in an image pixel represents a spatial average over the pixel surface (an area of 10×10 Km for MODIS aerosol), and cannot be justifiably equated with a point value measured with a sunphotometer. Secondly, even if the pixel is small enough to represent a point, it is extremely unlikely that it would represent the same conditions as a sunphotometer point data, since their observation axes are different and the atmosphere is constantly in motion. Again, clouds may obscure a MODIS pixel directly over a sunphotometer site, but may not affect nearby pixels. In the same way, times of sunphotometer measurements seldom coincide exactly with MODIS overpass times, the closest pair being often larger than 5 min apart. With cloud contamination and consequent data filtering, the shortest time difference can stretch even longer. Therefore, to achieve a meaningful and balanced validation, we compare spatial statistics from MODIS with corresponding temporal statistics from sunphotometers. The justification is that, since airmasses are constantly in motion, an airmass captured by MODIS across a certain horizontal span over a sunphotometer site, will be sampled by the sunphotometer during a certain time period.

[6] Corresponding MODIS and sunphotometer (currently only AERONET) data subsets are collocated in time and space, and extracted. Spatial and temporal statistics are computed and stored for use in validation and other studies. An automated system known as MAPSS (MODIS Aerosol and associated Parameters Spatio-temporal Statistics) has been developed for this purpose,

¹Science Systems and Applications Inc., NASA/GSFC, Greenbelt, MD, USA.

²Laboratory for Atmospheres, NASA/GSFC, Greenbelt, MD, USA.

³Laboratoire d'Optique Atmosphérique, Centre National de la Recherche Scientifique et Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, France.

⁴Laboratory for Terrestrial Physics, NASA/GSFC, Greenbelt, MD USA.

and the data can be accessed at <http://modis-atmos.gsfc.nasa.gov/mapss.html>.

3. MODIS Data Subsetting and Statistics

[7] The basis for the procedure is to identify within each MODIS aerosol image each pixel falling over a validation point by its longitude and latitude. Then an $N \times N$ Km subset centered on that pixel is extracted and its mean (μ_s) and standard deviation (σ_s) are computed. In addition, a linear multiple regression plane is fitted to each subset data and its greatest slope (θ_s) and azimuth of this slope (α_s) as well as the multiple correlation coefficient (R_s) are computed. The subscript ‘s’ designates spatial distribution. Standard procedures have been used for linear multiple regression plane fitting as well as for computing μ_s , σ_s , and R_s [e.g. Fox, 1997]. As regards θ_s and α_s , they are derived from the equation of the fitted plane, which can be expressed as:

$$ax + by + cz = k \quad (1)$$

where, for each pixel of the MODIS data subset being processed, x and y represent longitude and latitude values; z is the aerosol parameter value (e.g. AOT); a , b , and c are equation coefficients derived from the regression fit; and k is a constant. A vector perpendicular to the plane is referred to as the normal vector. The cosines of the angles the normal vector makes with the x , y , and z axes (termed the x , y , and z direction cosines) can be used to define the slope and slope azimuth of the plane as follows:

$$\begin{aligned} \theta_s &= \cos^{-1}(d_z) \\ \alpha_s &= \cos^{-1}\left(d_y / \sqrt{(d_x^2 + d_y^2)}\right) \end{aligned} \quad (2)$$

where, $d_x = a/r$, $d_y = b/r$, and $d_z = c/r$ are the x , y , z direction cosines derived from the equation of the plane coefficients, and $r = \sqrt{a^2 + b^2 + c^2}$.

[8] In this work, the computed θ_s and α_s are expressed in degrees. However, for slope, θ_s , it should be noted that the parameter represented by the z component of coordinate system are not in the same units as the x and y axes. It is necessary to remember the respective units when evaluating the physical gradient of the parameter variation. Thus, for instance, since a slope of 30° arbitrarily represents a vertical change of 0.577 units per unit horizontal distance, in a longitude-latitude coordinate system, a 30° slope in AOT represents a change of 0.577 per degree of geographic coordinates.

[9] All the computed quantities (μ_s , σ_s , θ_s , α_s , and R_s) are useful indicators of the local spatial characteristics of the aerosol parameter being analyzed, at a given location and time. Thus, μ_s represents the average value of the parameter, σ_s expresses its local spatial variability, θ_s would indicate the degree of the spatial trend if any, α_s gives the direction of the trend, and R_s shows how smooth the trend is.

[10] To evaluate the effect of window size on parameter statistics, four window sizes (30×30 , 50×50 , 70×70 , and 90×90 Km) were tested. Larger window sizes could introduce undesirable errors due to topographic or aerosol type heterogeneity. Figure 1 shows a random sample of MODIS AOT at $0.55 \mu\text{m}$ for April 19, 2000 over different sites (land and ocean). The bar heights represent the means, while the spikes atop them represent the standard deviations. We found that the window-size dependence is generally small and has no specific trend, at least for the window-size range tested. However, we use 50×50 Km window-size for calculating all validation spatial statistics, because for the 10-Km aerosol products, 30×30 Km corresponds to only 3×3 pixels, which is a small statistical sample, insufficient for plane

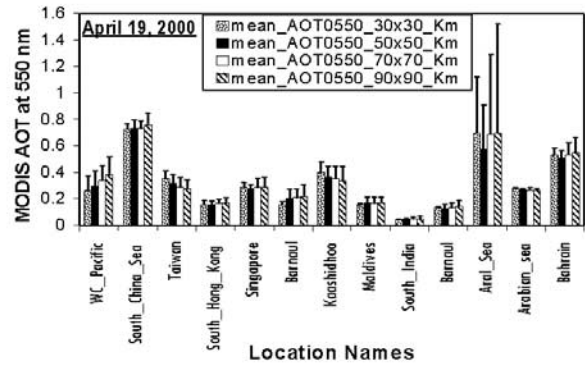


Figure 1. Averaging results of MODIS AOT at 550 nm, for 19 April 2000, over different locations using different window sizes (30×30 Km, 50×50 Km, 70×70 Km, and 90×90 Km). The means are represented by the heights of the wide bars while the standard deviations are plotted as the top spikes.

fitting. Moreover, the average travel speed of an aerosol front is of the order of 50 Km/h. This was visually estimated from animated daily sequences of TOMS aerosol index images (<http://jwocky.gsfc.nasa.gov/aerosols/aermovie.html>) for July to September 1988, where aerosol fronts are seen crossing the Atlantic from the west coast of Africa to the East coast of America (approximately 6000 Km) in about five or six days. Therefore, the 50×50 Km window would match a 1-hour sunphotometer data segment. All references to MODIS spatial statistics in the rest of this paper imply those based on the 50×50 Km (5×5 pixel) subset grid boxes.

4. AERONET Data Subsetting and Statistics

[11] For each location identified in the MODIS data, the AERONET sunphotometer data segment acquired within the 1-hour period centered on the MODIS overpass time are extracted. Statistics of the sunphotometer data subsets, mean (μ_t) and standard deviation (σ_t) are computed. However, since the data are not spatially distributed, it is not possible to fit a plane as for MODIS. Instead, the slope (θ_t) and the linear correlation coefficient (r_t) are computed in the temporal domain (as designated by the subscript ‘t’).

5. Result Analysis and Applications

[12] The derived spatio-temporal statistics allow us not only to identify uncertainties in the retrievals, but also to analyze the local spatial behavior of the aerosol parameters at individual stations and between different locations. Because of possible gaps in the data subsets used in deriving the statistics, only those computed from a certain minimum number of values (5 pixels for MODIS and 2 data points for AERONET) will be considered in the following analysis. Examples of the symbol representation in the following analysis are: $\mu_s(\tau_{a660})$, which stands for ‘spatial mean of MODIS AOT at 660 nm’; and $\sigma_t(\tau_{a670})$ denoting ‘temporal standard deviation of AERONET AOT at 670 nm’.

[13] Figure 2 shows scatterplots of $\mu_s(\tau_{a660})$ against $\mu_t(\tau_{a670})$ corresponding to MODIS retrievals for October 2000 over (a) land, and (b) ocean. The standard deviations $\sigma_s(\tau_{a660})$ and $\sigma_t(\tau_{a670})$ are plotted as error bars. The corresponding 1-1 lines (broken) as well as the least squares lines (solid) and the associated equations and correlation coefficients, R , are shown. The correlation and regression coefficients show excellent agreement with AERONET meas-

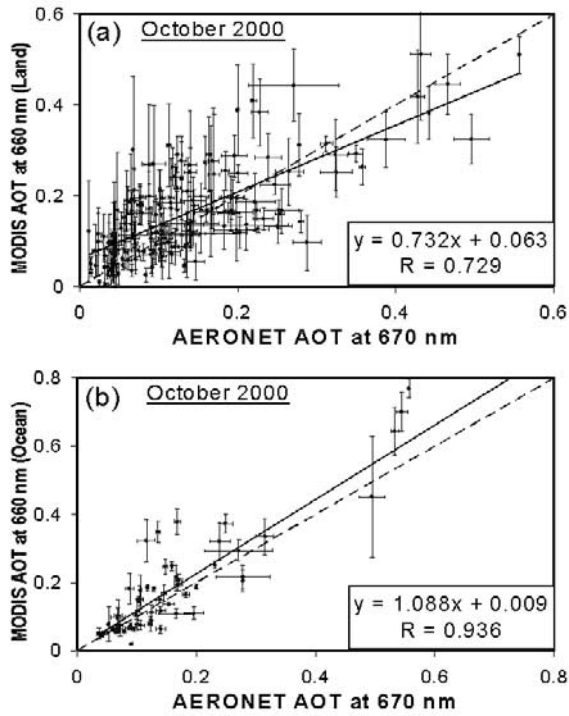


Figure 2. Scatterplots of means of τ_{a660} (MODIS 50×50 Km subsets) against τ_{a670} (AERONET 1-hour subsets) for all concurrent retrievals around the world during October 2000 (a) over land (b) over ocean. The error bars represent the respective standard deviations. The 1-1 lines are shown as the diagonal broken lines, while the solid lines are the linear regression fits, with their corresponding equations and correlation coefficients, R , shown at the lower right corners.

urements over the ocean (Figure 2b). Over land (Figure 2a), these coefficients reveal a relatively poorer performance, and at low AOT values ($\mu_r(\tau_{a670}) < 0.2$) the standard deviations appear to be larger for MODIS than for AERONET. This is probably due to the effects of surface variability in the MODIS retrievals over land. Figure 3a shows $\sigma_r(\tau_{a670})$ (AERONET) and $\sigma_s(\tau_{a660})$ (MODIS land and ocean) plotted against $\mu_r(\tau_{a670})$ for October 2000. The least squares line and R , as well as the root mean square (rms) value for each data set, are shown. Obviously, the values for MODIS over land show the least correlation and the most scatter, especially at low AOT ($\mu_r(\tau_{a670}) < 0.2$). Reflectance at 2100 nm (ρ_{2100}) enables a good view of the surface variability since most aerosols are transparent at this wavelength. Figure 3b shows $\sigma_s(\tau_{a470})$ and $\sigma_s(\tau_{a660})$ plotted against $\sigma_s(\rho_{2100})$ for $\mu_r(\tau_{a670}) < 0.2$. Amazingly, the correlations, R , practically exceed those of Figure 3a, showing indeed that when AOT is low ($\tau_{a670} < 0.2$), the variability of MODIS AOT over land is highly influenced by the variability of the land surface properties.

[14] The local spatial distribution of the aerosols can be visualized from the spatial slope (θ_s) and slope azimuth (α_s) computed from MODIS. Figure 4a shows serial plots of $\mu_s(\tau_{a470})$ and $\mu_s(\tau_{a660})$, depicting $\theta_s(\tau_{a470})$ and $\theta_s(\tau_{a660})$ by the circle sizes, and $\alpha_s(\tau_{a470})$ and $\alpha_s(\tau_{a660})$ by the arrow directions. They represent MODIS AOT over NASA Goddard Space Flight Center (GSFC) in Greenbelt, Maryland, USA, for the period of September to December 2000. The slope (θ_s) is an indicator for the relative aerosol loading gradient, while the slope azimuth (α_s) points to the

direction of lowest aerosol concentration, assumed to point away from the source direction. For most days, the spatial distribution is approximately the same at both wavelengths. Because of the directional ambiguity associated with very shallow slopes, directions are not shown for $\theta_s < 1^\circ$ (equivalent to an AOT change of less than 0.018 over a distance of 1 degree (or about 110 Km). Figure 4b shows similar plots only for τ_{a660} comparing two sites: GSFC and the Maryland Science Center (MSC) located at about 50 Km NNE of GSFC. These two locations (GSFC and MSC) almost always seem to have approximately the same direction, suggesting that they share the same aerosol source(s), except perhaps on October 4. However, except for September 17 and October 8, MSC almost always has a relatively higher average aerosol loading $\mu_s(\tau_{a660})$ and steeper gradient $\theta_s(\tau_{a660})$. The low value of $\theta_s(\tau_{a660})$ on September 17 suggests that this could not be caused by any permanent surface artifact. Rather, it is probably due to the impact of the source strength of urban aerosols on most days, since MSC is in the city of Baltimore.

6. Conclusions

[15] The spatio-temporal approach developed here has enabled the objective and rapid validation of satellite aerosol retrievals from MODIS with ground aerosol retrievals from AERONET. This is in spite of the differences in the characteristics of the two data sources. Furthermore, it has been possible to condense the local spatial

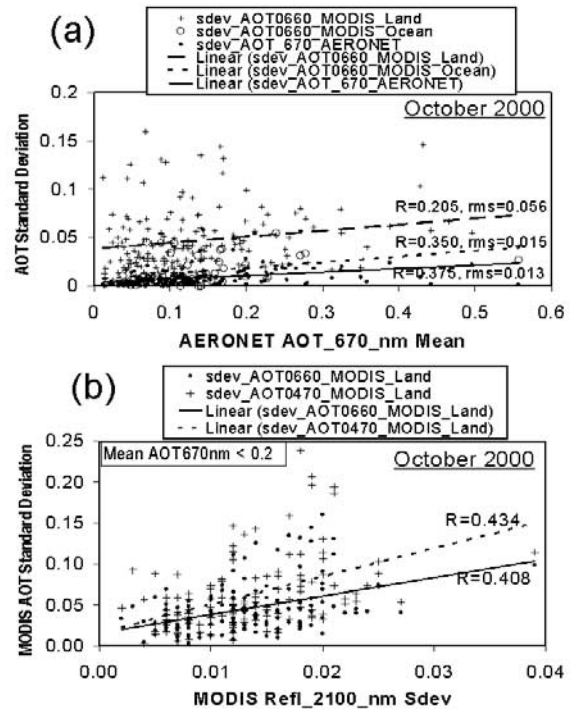


Figure 3. Scatterplots of (a) standard deviations (sdev) of τ_{a660} over land and ocean (MODIS 50×50 Km subsets) and τ_{a670} (AERONET 1-hour subsets) against mean τ_{a670} (AERONET 1-hour subsets) (b) sdev of τ_{a470} and τ_{a660} over land against the sdev of surface reflectance (Refl) at 2100 nm wavelength, only for cases where the AERONET mean $\tau_{a670} < 0.2$. The plots represent all concurrent retrievals around the world during October 2000. In all cases, linear regression lines are fitted and their correlation coefficients, R , are shown.

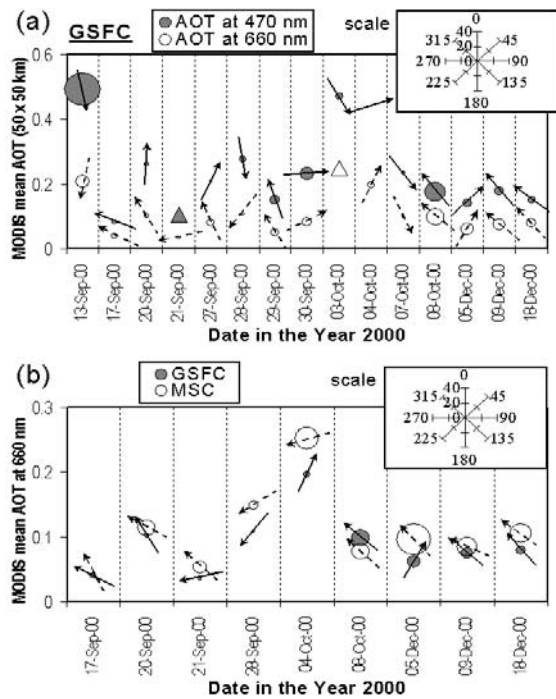


Figure 4. MODIS mean AOT over 50×50 Km subsets, showing corresponding spatial slopes (circle diameters) and spatial slope directions (arrows). The scale at top right corner of each plot shows the angular relationships expressed in degrees. For slope, a value of θ_s represents a change of $\tan \theta_s$ per geographical degree (about 110 Km) of horizontal distance. The arrow point designates the principal direction of lower AOT values. Arrows are not shown for very shallow slopes ($\theta_s < 1^\circ$) because of associated directional ambiguity, but triangles are used to designate such points. The plots represent: (a) τ_{a470} and τ_{a660} over the NASA Goddard Space Flight Center (GSFC, Greenbelt, MD) validation site (b) only τ_{a660} over two validation sites, GSFC and the Maryland Science Center (MSC, Baltimore, MD), which are about 50 km apart.

distribution of each MODIS aerosol parameter for every site into just a few parameters (μ_s , σ_s , θ_s , α_s , and R_s). This offers great opportunities to use an easy approach, such as spreadsheet analysis, to study the local spatial distribution of aerosols at various locations from satellite data. The methodology developed here will allow the

integration of satellite-retrieved aerosol parameters (and even other satellite products, such as water vapor) with data from various other sources to accomplish a wide variety of studies almost seamlessly.

[16] **Acknowledgments.** We would like to thank the various MODIS software development and support teams for the production and distribution of the MODIS data, and the AERONET teams for collecting, processing, and making available ground-based aerosol observations around the world. We would also like to thank Eric Moody (of the MODIS Atmosphere Group) for developing a MODIS data subsetting code, some of which routines were used in the MAPSS software.

References

Chu, D. A., Y. J. Kaufman, C. Ichoku, L. A. Remer, D. Tanré, and B. N. Holben, Validation of MODIS aerosol optical depth retrieval over land., *Geophys. Res. Lett.*, 29(12), mod 2, 1–4, 2001.

Fox, J., *Applied Regression Analysis, Linear Models, and Related Methods*, SAGE Publications, Thousand Oaks, 596 pp, 1997.

Holben, B. N., T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, AERONET—A federated instrument network and data archive for aerosol characterization, *Rem. Sens. Environ.*, 66, 1–16, 1998.

Kaufman, Y. J., D. Tanre, L. A. Remer, E. F. Vermote, A. Chu, and B. N. Holben, Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer, *J. Geophys. Res.*, 102, 17,051–17,067, 1997.

Remer, L. A., D. Tanré, Y. J. Kaufman, C. Ichoku, S. Mattoo, R. Levy, D. A. Chu, B. N. Holben, O. Dubovik, Z. Ahmad, A. Smirnov, J. V. Martins, and R.-R. Li, Validation of MODIS Aerosol Retrieval Over Ocean, *Geophys. Res. Lett.*, 29(12), mod 3, 1–4, 2001.

Smirnov, A., B. N. Holben, T. F. Eck, O. Dubovik, and I. Slutsker, Cloud-screening and quality control algorithms for the AERONET database, *Rem. Sens. Environ.*, 73, 337–349, 2000.

Tanré, D., Y. J. Kaufman, M. Herman, and S. Mattoo, Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances, *J. Geophys. Res.*, 102, 16,971–16,988, 1997.

Zhao T. X.-P., L. L. Stowe, A. Smirnov, D. Crosby, J. Sapper, and C. R. McClain, Development of a global validation package for satellite oceanic aerosol optical thickness retrieval based on AERONET observations and its application to NOAA/NESDIS operational aerosol retrievals, *J. Atmos. Sci.*, 59, 294–312, 2002.

C. Ichoku, D. A. Chu, S. Mattoo, and I. Slutsker, Science Systems and Applications Inc., NASA/GSFC, Greenbelt, MD 20771, USA. (ichoku@climate.gsfc.nasa.gov; achu@climate.gsfc.nasa.gov; mattoo@climate.gsfc.nasa.gov)

Y. J. Kaufman and L. A. Remer, Laboratory for Atmospheres, NASA/GSFC, code 913, Greenbelt, MD 20771, USA. (kaufman or remer@climate.gsfc.nasa.gov)

D. Tanré, Laboratoire d’Optique Atmosphérique, Centre National de la Recherche Scientifique et Université des Sciences et Technologies de Lille, Villeneuve d’Ascq, France. (Didier.Tanre@univ-lille1.fr)

B. N. Holben, Laboratory for Terrestrial Physics, NASA/GSFC, code 923, Greenbelt, MD 20771, USA. (brent@aeronet.gsfc.nasa.gov)