Estimation of Aerosol Extinction-to-Backscatter Ratios using AERONET Measurements and Cluster Analysis

Ali H. Omar, Jae-Gwang Won[†], Sun-Chang Yoon[†], M. Patrick McCormick
Hampton University, Center for Atmospheric Sciences
23 Tyler Street, Hampton, VA 23668
Tel: 757-7275127 ● Fax 757-7275090
email: ali.omar@hamptonu.edu

ABSTRACT

Radiance measurements and inversions of the AErosol RObotic NETwork (AERONET)[1] are used to characterize global atmospheric aerosols and key parameters such as the phase functions and column extinction-to-backscatter ratio (lidar ratio, S_a) for the analysis of lidar measurements. This study uses more than 10^5 records of aerosol size distributions and complex refractive indices to generate several optical properties of the aerosol at more 200 sites worldwide. These properties together with the radiance measurements are then classified using classical clustering methods which group the sites according to the type of aerosol with the greatest frequency of occurrence at that site. Six significant clusters are identified: two types of desert dust, biomass burning, polluted continental, clean continental, and a dust+smoke aerosol. In addition to categorizing the aerosol types, all the column properties that are important for the analysis and validation of lidar measurements such as the lidar ratio, optical depth and phase functions are determined for each of the seven aerosol types. The variances of the aerosol determined using principle component analysis (PCA). Aerosol lidar ratios at 532 nm of the six categories of aerosol are calculated.

INTRODUCTION

The extinction/backscatter ratio (S_a) is an important parameter used in the determination of the aerosol extinction and subsequently the optical depth from lidar backscatter measurements. Historically, the use of S_a came into common use because the single scattering lidar equation is underdetermined. A unique solution is only possible if the two unknowns, extinction and backscatter are combined into one variable (S_a). There is therefore a need to determine S_a at various locations worldwide especially in light of the advent of satellite-based lidar measurements such as the Cloud and Aerosol Lidar and Infrared imager Pathfinder Spaceborne Observations (CALIPSO)[2]. This study shows that assigning a lidar ratio to a given location based on a long-term average is unrealistic. This is because at any given location, the aerosol type is highly variable on time scales as short as a few hours. These variations result from transport of distinct airmasses to a site and non-systematic events such as fires, wind gusts, hurricanes, tornadoes, and land clearing and development activities. These variations lead to diverse aerosol characteristics at each site on time scales of a few hours and preclude the long term averaging of aerosols to develop a representative set of characteristics for a site or region. Aerosol optical measurements must therefore be made at short time scales (about 30 minutes) to develop a large data base which can be used to derive statistically significant correlations. The AErosol RObotic NETwork (AERONET) measurements are likely to provide such a data base albeit for column rather than vertically-resolved measurements. AERONET is an automatic robotic Sun and sky scanning measurement network that has grown rapidly to over 200 sites worldwide. AERONET uses multi-angle radiance measurements to retrieve the discrete aerosol size distributions in 22 size bins ranging from 0.05 to 15 μm and the complex refractive index[3,4]. This study uses the AERONET refractive indices and size distributions for Mie scattering calculations from which several parameters of aerosols are determined. The calculated single scattering albedos determined using Mie calculations are compared with those of AERONET retrievals for code validation. On average , the relative error between these two results is less than 1%. Cluster analysis is used for categorization of atmospheric aerosol types. Six significant types: two types of desert dust, biomass burning, polluted continental, clean continental, and a dust+smoke aerosol are suggested by the cluster analysis.

[†] Seoul National University, Seoul, S. Korea

METHOD

Cloud screening using radiance measurements and optical properties

The sun/sky radiometer measurements are frequently contaminated by clouds and depending on the cloud reflectivities, can have a significant effect on the sun-sky radiance measurements. This study used the AERONET Level 1 size distribution data and applies a two-part cloud screening scheme. The first part checks the symmetry of the almucantar measurements and the second part is a statistical screening procedure. The almucantar measurement is made at several azimuthal angles with the same elevation angle of the direct sun. For the aureole measurement, the degree of angle change is set to be quite small near the direction of sun. In order to ensure that the sky is clear at the time of the measurement, we calculate the relative error between the seven pairs of data measured on either side of the direct sun. The azimuthal angles of these measurements are 2.0°, 2.5°, 3.0°, 3.5°, 4.0°, 5.0°, and 6.0°. Using this method ensures that the sky is clear near the measurements because the presence of cloud cover would result in disparities between symmetrical pairs of measurements. The statistical screening procedure uses the basic idea of the cloud screening method for the direct solar measurement of AERONET sun/sky radiometer[5]. In this procedure, records with optical depths(τ) and Angstrom coefficients (α) that exceed a fixed number of standard deviations (σ) on either side of the mean of the distribution are not included in the analysis. The range of acceptable optical depths is (τ_{mean}–σ, τ_{mean}+3σ). The acceptable Angstrom coefficients are greater than α_{mean}-3σ. These conditions assume that unrealistically large aerosol extinction values are due to fugitive cloud contamination that escaped the cloud screening procedure. A very small Angstrom coefficient is also indicative of cloud contamination. The cloud screening scheme reduced the total number of data points from 181316 to 104442 (58% reduction).

Several aerosol optical properties were calculated using the screened AERONET data. Of greatest interest to this study are the lidar ratios. Since the AERONET data set is retrieved from the radiance at four wavelengths, the aerosol optical properties at these four wavelengths were calculated using the AERONET-derived complex index of refraction and size distribution. Three pairs of optical depths at sequential wavelengths (i.e., 440/670 nm, 670/870 nm, 870/1020 nm) were used to calculate three optical depth-based Angstrom exponents. The AERONET algorithm for size distribution retrieval provides the volume distribution data of 22 size bins (dV/dlnr, V is volume, and r is radius) from 0.05 µm to 15 µm. The best fit for the AERONET size distribution data is a two-mode log-normal size distribution described by equation (1),

$$\frac{dV}{d\ln r}=\frac{C_f}{\sqrt{2\pi}\ln\sigma_f}\exp\left(-\frac{\left[\ln r-\ln\overline{r_f}\right]^2}{2(\ln\sigma_f)^2}\right)+\frac{C_c}{\sqrt{2\pi}\ln\sigma_c}\exp\left(-\frac{\left[\ln r-\ln\overline{r_c}\right]^2}{2(\ln\sigma_c)^2}\right) \tag{1}$$

where the Cs are mode constants and the subscripts f and c denote fine and coarse modes, respectively. σ is the geometric standard deviation and $\bar{r}$ is the geometric mean radius. This partition of the size distribution into fine and coarse modes yields six parameters by which a particular size distribution can be described. The extinction and backscatter coefficients are calculated at four AERONET wavelengths and subsequently interpolated to the Nd:YAG laser wavelengths of 1064 nm and 532 nm using Angstrom exponents and polynomial fits. Fig 2 shows the sample results of polynomial fitting with the dataset of AERONET measurements at the Goddard site.

Cluster analysis and aerosol categorization

Cluster analysis is a statistical tool used for grouping the large data sets into several categories using predefined variables. In this study, cluster analysis is used to categorize the AERONET data set based on several optical and physical characteristics of the aerosol. The first step is to determine which variables should be included in the cluster analysis. Each record of the AERONET data set has up to 33 optical and physical variables associated with it. Some of these are direct AERONET measurements while the rest are generated from AERONET observations using Mie calculations. Atmospheric aerosols are generally estimated to have at least five distinct classes: marine aerosol, desert-dust, biomass burning aerosol, urban aerosol and rural -background aerosol. This forms the minimum number of aerosol clusters. The actual number of clusters is determined by calculating the clusters using successively larger cluster numbers until the calculation does not yield any new significant clusters. Having defined the number of categories, a random function is used to determine the initial conditions of each variable. These initial coordinates of the variables are located within one standard deviation away from the mean assuming a normal distribution of the variables.

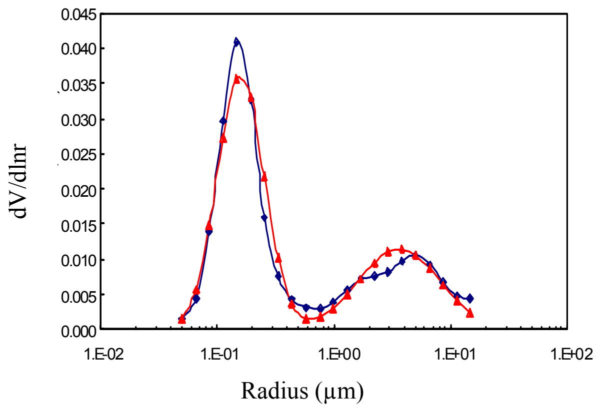

Figure 1. The discrete AERONET size distribution(•) and the size distribution obtained by fitting the data to equation 1 (▲).

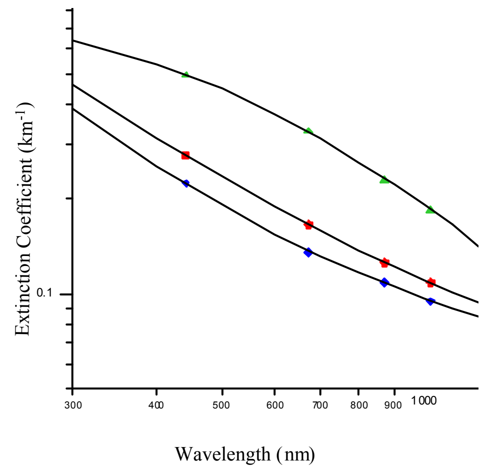

Figure 2. Interpolation of σ_{ext} to the Nd:YAG wavelengths for three AERONET measurements

Each record is grouped in the cluster whose center is nearest to the record. The distance is not a physical distance but the mathematical distance between two points in 26-dimensional space. The distance, normalized by the standard deviation to eliminate bias resulting from the different magnitudes of the variables, between a record and the center of category j is calculated using equation 2,

$$\text{dis tan ce}(j) = \sum_{\substack{i=1 \\ j=1}}^{\substack{j=8 \\ i=26}} \frac{\text{data record}(i) - \text{center}(j,i)}{\{\text{s tan dard deviation}(i)\}^2} \quad . \tag{2}$$

The eight distances are compared and each record is grouped in the category with the shortest distance to the cluster center. After grouping the records into the eight groups, a new center for each cluster is determined by averaging the variables in each cluster. The process is repeated with the new cluster centers until the relative error between the centers of new and old clusters is less than a prescribed value. In this study the criterion for stopping the iteration is set as 0.1%. Convergence is achieved within 25 iterations for the various initial conditions. Principle component analysis was used to determine the variables that are most important in the variation of the data set by examining the respective variances. The six largest eigenvalues account for 70% of total variance. Complex refractive indices are the most significant variables because the 1st and 2nd components explain more than 43% of the total variance.

The categories of aerosol

Table 1 shows the cluster analysis results including the size distributions and the 532 nm lidar ratios of the eight clusters. Fig 3 shows the locations of the sites that show a predominance of desert dust. Cluster 1 has relatively large amount of coarse size aerosol indicating that these could be desert dust. The size distribution of Cluster 2 suggests that it might be biomass burning possibly mixed with marine aerosol. Clusters 3 is located in regions where the predominant aerosol is clean continental. Cluster 4 aerosols have relatively large fine fraction (found by comparing C_f and C_c) of fine particles and the sites of this cluster are mainly located the eastern United States and western Europe. Cluster 4 is most likely urban/ industrial aerosol. Cluster 5 and 7 have very few members and are not likely to represent any real aerosol type and were therefore dicarded. Cluster 6 is similar to Cluster 1 and is located in the same regions and is therefore likely to be a dust aerosol. Cluster 8 is located in regions where biomass burning is common, such as areas of southern Africa and South America including the Amazon. Cluster 2 also has a significant coarse mode particle size distribution but the total number of sites of this category is small with most members of this category located near coastal regions or islands. Cluster 3 has the largest number of sites and the complex refractive indices and size distributions are consistent with rural aerosols. This categorization compares favorably with the other findings[6,7] from similar data sets.

Conclusion

The AERONET data set has been used to categorize global aerosol into six distinct clusters described by 26 column optical and physical properties. These properties include the aerosol lidar ratios at 532 nm which are critical for global lidar data analysis especially following the launch of earth-orbiting satellite-based lidars. The categorization relies on telltale properties such as the fine and coarse fractions, particle size, optical depth, geographic location and in some cases seasonal variation, and therefore requires additional validation using long term observations. The continued growth of the AERONET data base is likely to improve and reduce the uncertainty in the classification. The effects of non-sphericity on the cluster analysis are unknown but not negligible. These need to be addressed in future studies.

Table 1. Central values of the size distributions and lidar ratios of the six clusters.

Category / Parameter	Cat 1	Cat 2	Cat 3	Cat 4	Cat 6	Cat 8
$\bar{r}_f (\mu m)$	0.1184	0.1388	0.1443	0.1872	0.1649	0.1597
σ_f	1.9316	1.5783	1.5236	1.7234	1.7604	1.7337
C_f	0.1294	0.0392	0.0525	0.0465	0.0274	0.0216
$\bar{r}_c (\mu m)$	2.3645	3.793	3.4869	3.08	3.2156	3.4824
σ_c	1.7656	2.0583	2.018	1.9796	2.0169	2.0728
C_c	0.2944	0.0309	0.0323	0.065	0.0633	0.0293
S_a (532 nm)(sr)	32	71	57	64	33	108

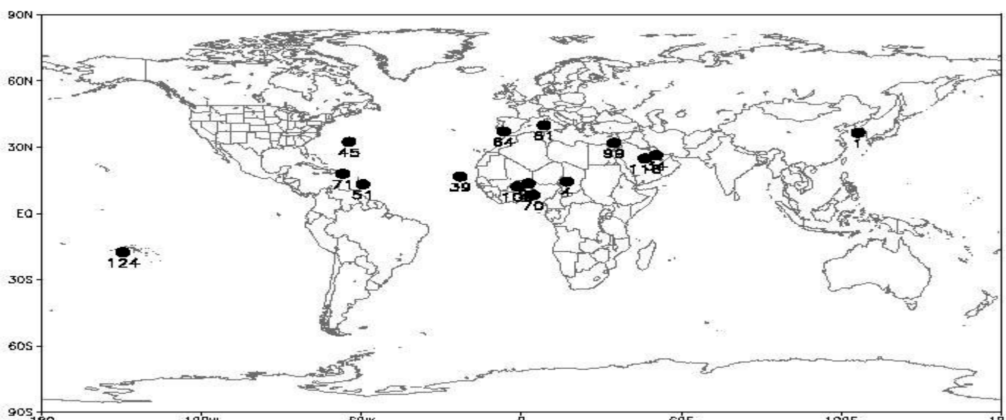

Figure 3. Location of category 1 aerosol types with (S_a = 32 sr, coarse mode dominant).

REFERENCES

1. Holben, B.N., D.Tanre, A.Smirnov, T.F.Eck, I.Slutsker, N.Abuhassan, W.W.Newcomb, J.Schafer, B.Chatenet, F.Lavenue, Y.J.Kaufman, J.Vande Castle, A.Setzer, B.Markham, D.Clark, R.Frouin, R.Halthore, A.Karnieli, N.T.O'Neill, C.Pietras, R.T.Pinker, K.Voss, and G.Zibordi, 2001: An emerging ground-based aerosol climatology: Aerosol Optical Depth from AERONET, *J. Geophys. Res.*, **106**, 12 067-12 097.

2. Winker, D. M., 2002, The CALIPSO mission, Proceedings of the 21st ILRC, July 2002, Quebec, Canada.

3. Dubovik, O. and King, M. D., 2000, A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, *J. Geophys. Res.*, Vol. 105, pp 20673-20696

4. Dubovik, O., A. Smirnov, B. N. Holben, M. D. King, Y.J. Kaufman, T. F. Eck, and I. Slutsker, 2000: Accuracy assessments of aerosol optical properties retrieved from AERONET sun and sky-radiance measurements, *J. Geophys. Res.*, **105**, 9791-9806

5. Smirnov A., B.N.Holben, T.F.Eck, O.Dubovik, and I.Slutsker, 2000: Cloud screening and quality control algorithms for the AERONET data base, *Rem.Sens.Env.*, **73**, 337-349.

6. Eck, T.F., B.N. Holben, D.E. Ward, O. Dubovik, J.S. Reid, A. Smirnov, M.M. Mukelabai, N.C. Hsu, N.T. O'Neill, and I. Slutsker, 2001: Characterization of the optical properties of biomass burning aerosols in Zambia during the 1997 ZIBBEE field campaign, *J. Geophys. Res.*, **106**, 3425-3448.

7. Remer, L.A., Y.J.Kaufman, B.N.Holben, A.M.Thompson and D.McNamara, 1998: Biomass burning aerosol size distribution and modeled optical properties, *J.Geophys.Res.*, **103**, 31 879-31 891.