Multi-modal fitting of AERONET size distributions during atypical aerosol conditions

Taylor M., Kazadzis S., Gerasopoulos E.

To date, aerosol size distributions obtained from AERONET have been fit with bi-lognormal distributions that separate fine and coarse modes at a local minimum in the radial interval 0.439μm to 0.992μm. In the present work, a new technique for calculating the separation point is proposed to optimize the goodness of fit to the size distribution. The six microphysical parameters derived from this process (the volume concentration, effective radius, and the variance of fine and coarse particle modes) are compared with those provided by AERONET for characteristic aerosol types including: desert dust, biomass burning, urban sulfate, and marine sea salt. Following on from this, a Gaussian mixture model (in logarithmic space) is used to fit the dataset. The optimal number of Gaussians is calculated based on the assumption of whether or not the addition of extra modes statistically improves the square of the correlation coefficient or passes the Fisher Test. The main results are summarized as follows: 1) the new method for calculating the separation point shows an improvement in the case of desert dust events, 2) the fit to the size distribution of marine aerosol improves with 3 Gaussians, and 3) the multi-modal fit to the AERONET data is the most accurate, automatically-detects the optimal number of modes regardless of the aerosol type and provides the microphysical parameters of additional modes which are currently not available from the methodology used by AERONET.

Taylor M.*, Kazadzis S., Gerasopoulos E.

Institute for Environmental Research and Sustainable Development (IERSD), National Observatory of Athens (NOA), Metaxa & Vas. Pavlou, Penteli, 15236, Athens, Greece.

***corresponding author e-mail: patternizer@gmail.com**

1 Introduction

The retrieval of parameters such as the aerosol optical depth (AOD) from satellite measurements is accomplished by algorithms that model the optical characteristics of columnar aerosol via parameters of microphysical structure including the aerosol volume size distribution (AVSD) and the spectral complex refractive index. Retrievals are therefore rather sensitive to the choice of model of particle size and composition. Furthermore, difficulties are compounded by the fact that the complete set of required parameters cannot presently be obtained unambiguously (Hasekamp and Landgraf 2007) especially when the spectral and directional behavior of the surface reflectance is unknown (Kokhanovsky et al. 2010). Since remote sensing of the AVSD is exceedingly difficult (King et al. 2009), satellite retrievals are usually validated (e.g. Remer et al. 2005) against co-located and synchronous retrievals provided by ground-based sun photometer and sky radiometer systems like those in the aerosol robotic network (AERONET). The heavy reliance on the AVSDs provided by AERONET, combined with the fact that the AVSD plays a pivotal role in the relation of the radiation field to the microphysics of aerosol particles (Hansen and Travis 1974) and the determination of aerosol type and composition, motivated us to develop two new methods to improve the fit to retrieved AVSDs (described in detail in Taylor et al. 2013). Here we briefly outline these new methods and present some results of their application to a couple of very atypical AVSDs.

The AERONET inversion code presently approximates the AVSD by optimization of the shape of each radial size bin (Dubovik and King 2000). In particular, lognormal-shaped bins provide notable improvements over the trapezoidal approximation suggesting that the AVSD can better be modeled as a superposition of n-lognormals with the modal volume concentrations V_i geometric mean radii r_i and standard deviations σ_i as fixed secondary microphysical parameters,

$$\frac{dV(r)}{d\ln r} = \sum_{i=1..n} \frac{V_i}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln r - \ln r_i}{\sigma_i}\right)^2} \quad (1)$$

For several decades now, the AVSD of tropospheric aerosols has been known to contain several distinct modes, each most commonly being modeled by a lognormal function (Whitby, 1978). The statistical properties of the lognormal and bi-lognormal distribution are well-known (O'Neill et al. 2000) and are applied in this paper to test the feasibility of modeling the AVSD of distinct aerosol cases with super-positions of several ($n \geq 2$) lognormals. Furthermore, since many available radiative-transfer codes are now able to take as input lognormal distribution parameters (Sayer et al. 2012), the results of applying the methods we have developed, can be readily applied and implemented.

2 Data and Methodology

2.1 Data

A search for AVSDs corresponding to interesting and atypical aerosol conditions must presently be done manually or with reference to targeted field campaigns. We have developed a simple approach for isolating candidate AERONET sites as well as those days which are most dominated by for example desert dust, biomass burning products, urban sulphate, and marine (sea salt) aerosol. The Georgia Institute of Technology–Goddard Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) model (Chin et al. 2000, Ginoux et al. 2001) used by NASA's GEOS-5, simulates the contribution to the AOD from major types of tropospheric aerosols including: sulphate (SU), black carbon (BC), organic carbon (OC), desert (mineral) dust (DU) and sea salt (SS). While a comparative analysis of GOCART AOD at co-located AERONET sites is beyond the scope of this work, global distributions of GOCART AOD and its multi-decadal variations have been reported to agree with different satellite observations within a factor of 2 (Chin et al. 2010). Here, we use GOCART data for a model-driven aerosol classification. GOCART data spanning the years 2001-2005 (inclusive) was downloaded from the global grid (spatial resolution: 2.5 x 2 degrees) co-located with 155 AERONET sites (75% of all *Level 2.0 Version 2 Inversion Product* records *N*) which were ranked by the size of their daily-averaged data record. The ratio of the contribution of individual aerosol types to the total AOD was then calculated (as percentages) and used to sort the list of ranked sites again but by dominant ("nearly-pure") aerosol type. As a result, it was possible to automate the identification of sites most dominated by DU, OC+BC (i.e. biomass burning aerosol), urban SU and SS: Banizoumbou (Niger), Mongu (Zambia), GSFC-Washington (USA) and Lanai (Hawaii) respectively. Then, daily-averaged AERONET Level 2.0 Version 2 quality-assured and cloud-screened data was aligned (i.e. filtered for synchronous values) with the daily-averaged GOCART AOD data at each site and the day corresponding to the peak %DU at Banizoumbou, peak %(BC+OC) at Mongu, peak %SU at GSFC-Washington, and peak %SU at Lanai was found (see Table 1 for their composition) and the associated AVSD was extracted.

Table 1. Days of dominant aerosol type as derived from ranking of the GOCART global chemical data source contributions to the AOD.

Site	Peak	%SU	%OC	%BC	%DU	%SS	%OC+BC
Banizoumbou	16/3/05	1.0	0.7	0.3	97.9	0.0	1.0
Mongu	14/8/03	5.6	77.4	16.8	0.2	0.1	94.1
GSFC	17/8/05	87.5	8.3	2.7	1.4	0.1	11.0
Lanai	21/2/02	28.9	5.3	2.3	3.3	60.1	7.6

2.2 Methodology

The AVSD is usually assumed to be bi-modal so as to distinguish between a "fine" (accumulation) mode containing small particles ($< \approx 0.6\mu m$) and a "coarse" mode containing larger aerosol particles. With reference to Eq. (1), the partition of the AVSD into $n = 2$ lognormal modes with a fine (f) mode and a coarse (c) mode therefore requires the calculation of 6 secondary parameters: V_f, V_c, r_f, r_c, σ_f and σ_c. These parameters are estimated in the AERONET retrieval algorithm by dividing the total volume size distribution dV/dlnr into two parts at a radial mode separation point r_s. Values of r_f, r_c, σ_f and σ_c are then calculated from the AVSD by integrating to and from the separation point. The equations necessary for their calculation is presented in (Taylor et al. 2013). The AERONET Level 2.0 Version 2 inversion algorithm provides the values of all of these parameters including the value of r_s upon which they all depend by finding the local minimum within the size interval $0.439 \le r \le 0.992\mu m$ (Dubovik et al. 2000). By interpolating the AERONET AVSD (comprising 22 equidistant logarithmic radial size bins spanning the range of particle radii $0.05 \le r \le 15\mu m$) on a higher resolution radial grid containing 2200 points, we varied the separation point r_s across the grid while keeping the total volume concentration constant. For each different r_s we calculated the standard error s and the coefficient of determination R^2 arising from the difference between the interpolated AERONET AVSD and the bi-lognormal re-constructed from the derived secondary parameters V_f, V_c, r_f, r_c, σ_f and σ_c in Eq. (1) for $n = 2$ modes. The optimal separation point arising from this sensitivity analysis was found from the location of r_s that i) maximizes R^2 and ii) minimizes s. We call this first method the optimized equivalent volume (OEV) method (see Taylor et al. 2013 for details).

In our second method, we drew upon the knowledge that the AVSD is normally-distributed in the ln$r$ domain (i.e. it is lognormal in r) and performed nested nonlinear fitting that involved consecutively adding Gaussian distributions and re-calculating the value of R^2.

On addition of each additional mode in this Gaussian mixture model (GMM) method, we performed nested hypothesis testing with Fisher transforms (Fisher 1921) and the Welch statistic (Welch 1947) to test that supplementary modes produce a statistically-significant improvement in R^2 (i.e. genuinely improved the fit). This stopping condition allowed us: 1) to fit the AVSD with n > 2 modes, 2) to fit the AVSD without resorting to the determination of mode separation points (there are n-1 of them for n modes), and 3) to lift the constant volume constraint applied in the OEV method. Details of the fitting procedure are described in Taylor et al. (2013) and an example of fitting the clearly 3-modal marine aerosol-dominated AVSD at Lanai is shown in Fig. 1.

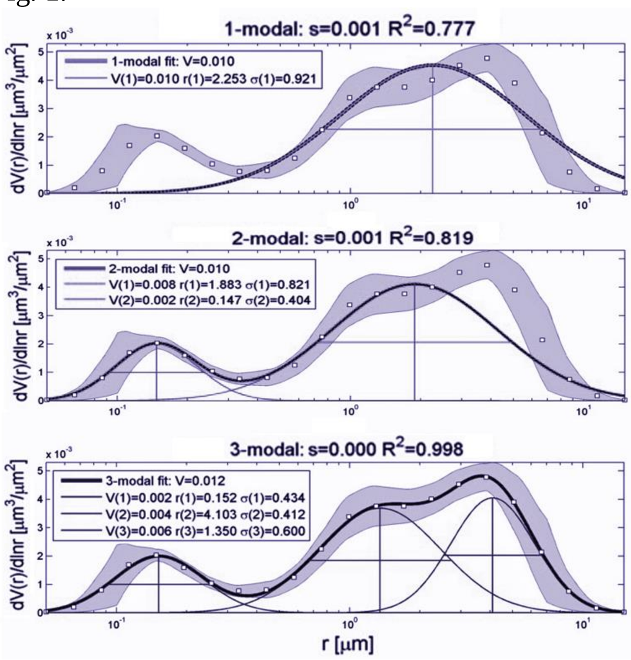

Fig. 1. GMM fitting of the interpolated AVSD of maritime-dominated aerosol at Lanai on the 21st of February 2001 which shows a novel double-peak in the coarse mode region > 1μm. Note the gradual improvement on addition of each new mode peaking at R^2 = 0.998 for n = 3 modes. The grey band is the uncertainty on the AERONET AVSD as per the assessment of Dubovik et al. (2002).

3. Results

The AVSD of marine-dominated aerosol at Lanai presents an atypical double-peak in the coarse mode region. A double-peak in the fine mode region (≤ 0.6μm) has been observed at Fresno (USA) and associated with fog- and cloud-induced modification of ambient aerosol Eck et al. (2012). In Fig. 2 we compare the fit of this AVSD with the standard AERONET bi-lognormal and with the OEV and GMM fits. In Fig. 3, the fits to the AVSD of volcanic ash impact at a coastal site in the UK in presented.

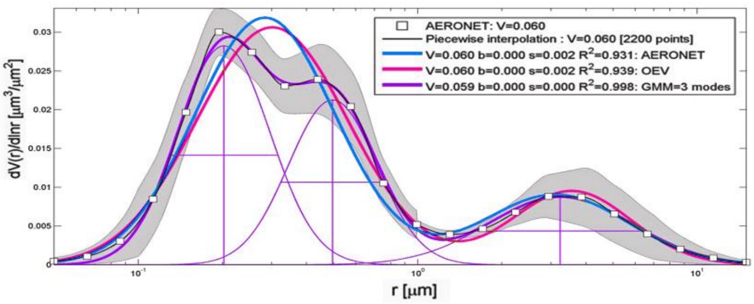

Fig. 2. Comparison of the interpolated AVSD, the AERONET bi-lognormal fit (blue), the OEV bi-lognormal fit (pink) and the GMM optimal fit (purple) to the interpolated daily-averaged AERONET AVSD at Fresno on the 11th of February, 2006 which displays a clear double-hump in the fine mode region. The GMM fit is clearly better than the OEV and AERONET fits. Both the OEV and AERONET fits successfully capture the coarse mode but fail to capture the double-peak fine mode feature.

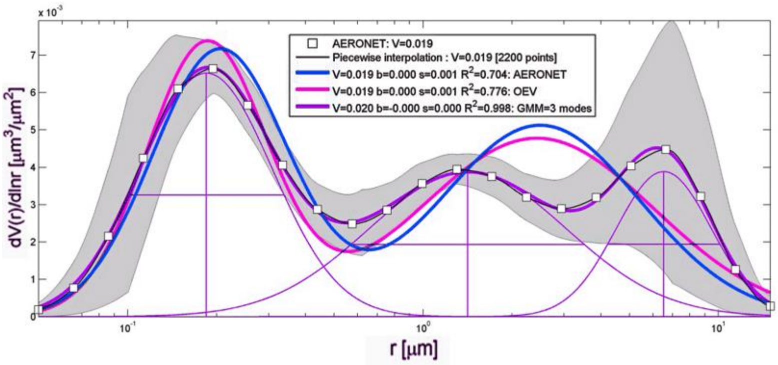

Fig. 3. Comparison of the fits to the interpolated daily-averaged AERONET AVSD at Chilbolton (UK) on the 23rd of May 2010 – 2 days after the eruption of volcanic ash from the Icelandic Eyjafjallajökull Volcano. The GMM is clearly better than the OEV and AERONET fits. The OEV fit is able to capture the location of the fine mode peak better than the AERONET fit, but its restriction to a bi-lognormal means that it fails to fit the remaining two peaks in the coarse mode region.

For the results of fitting dust-, biomass burning- and urban SU-dominated AVSDs with the OEV and GMM methods we refer the reader to Taylor et al. (2013).

4 Conclusions

The GMM method is able to fit very well even the very atypical AVSDs presented here. It is an automated procedure that can easily be incorporated into operational retrieval algorithms. In contrast, the OEV method, while more elastic in its fitting and slightly better than AERONET, is limited to a bi-lognormal and is unable to fit AVSDs where the number of modes $n > 2$.

Acknowledgments MT was supported by a Marie-Curie IEF funded project "AEROMAP: Global mapping of aerosol properties using neural network inversions of ground and satellite based data", CN: 300515.

References

Chin M, Rood RB, Lin SJ, Müller JF, Thompson AM (2000) Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties. J. Geophys. Res.: Atmospheres (1984–2012), 105(D20):24671-24687.

Chin M, Diehl T, Streets D, Qian Y, Yu H, Bian H, Tan Q, Kucsera T, Wang W (2012) Multi-decadal variations of atmospheric aerosols and their effects on surface radiation trends. EGU 2010, Available at: http://gmi.gsfc.nasa.gov/mtgs_rpts/2010_09/chin_decadal_aerosols.pdf.

Dubovik O, King MD (2000) A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. J Geophys. Res. 105(D16): 20,673-20,696.

Dubovik O, Smirnov A, Holben BN, King MD, Kaufman YJ, Eck TF, Slutsker I (2000) Accuracy assessment of aerosol optical properties retrieval from AERONET sun and sky radiance measurements. J. Geophys. Res., 105: 9791–9806.

Eck TF, Holben BN, Reid JS, Giles DM, Rivas MA, Singh et al. (2012) Fog- and cloud-induced aerosol modification observed by the Aerosol Robotic Network (AERONET). J Geophys. Res. 117(D7), D07206.

Fisher RA (1921) On the "probable error" of a coefficient of correlation deduced from a small sample. Metron 1:3–32.

Ginoux P, Chin M, Tegen I, Prospero JM, Holben B, Dubovik O, Lin SJ (2001) Sources and distributions of dust aerosols simulated with the GOCART model. Journal of Geophysical Research, 106(D17):20255-20273.

Hasekamp OP, Landgraf J (2007) Retrieval of aerosol properties over land surfaces: capabilities of multi-viewing-angle intensity and polarization measurements. Appl. Opt., 46(16): 3332–3344.

King MD, Kaufman YJ, Tanré D, Nakajima T (2009) Remote Sensing of Tropospheric Aerosols from Space: Past, Present, and Future. Bull. Am. Met. Soc. 2229-2259.

Kokhanovsky AA et al (2010) The determination of spectral aerosol optical thickness from satellites: An inter-comparison of algorithms using synthetic backscattered solar light characteristics, Atmos. Meas. Tech., 3: 909–932.

O'Neill NT, Ignatov A, Holben BN, Eck TF (2000) The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET sunphotometer data. Geophys. Res. Lett. 27(20): 3333-3336.

Remer LA, Kaufman YJ, Tanre D, Mattoo S, Chu DA, Martins J et al (2005) The MODIS Aerosol Algorithm, Products, and Validation. J. Atmos. Sci., 62: 947–973.

Sayer AM, Smirnov A, Hsu NC, Holben BN (2012) A pure marine aerosol model, for use in remote sensing applications. J Geophys. Res. 117: 1-25.
Taylor M, Kazadzis S, Gerosopoulos E (2013). Multi-modal analysis of aerosol robotic network size distributions for remote sensing applications: dominant aerosol type cases. Atmospheric Measurement Techniques Discussions 6, 10571–10615, 2013.
Welch BL (1947) The generalization of "Student's" problem when several different population variances are involved. Biometrika 34 (1–2): 28–35.
Whitby KT (1978) The physical characteristics of sulfur aerosols, Atmos. Environ., 12: 135-159.