

Atmospheric aerosol characterization using multi-wavelength, multi-static light scattering

Andrea M. Brown^{*a}, Michelle G. Snyder^b, Lydia Brouwer^d, and C. Russell Philbrick^{a,b,c}
^aDept. of Electrical Engineering, Penn State University, University Park, PA, USA 16802;
^bDept. of Physics, and ^cDept. of Marine, Earth, and Atmospheric Sciences,
North Carolina State University, Raleigh, NC USA 27695
^dAlion Science and Technology, Research Triangle Park, NC 27711

ABSTRACT

A sensor for measuring scattering at multiple wavelengths and multiple angles has been designed and is being tested for the characterization of atmospheric aerosols. Charge coupled device (CCD) imagers are used to record scattering measurements at two polarizations and as a function of angle relative to the co-aligned laser beams. A diffraction grating is used to spatially separate the wavelengths across the field-of-view of the CCD array, allowing simultaneous measurements at multiple wavelengths. Experiments are conducted to measure the scattering intensities for two polarizations at discrete wavelengths that span the visible spectrum. The data from the CCD images are inverted using a genetic algorithm and Mie scatter equations to determine aerosol properties of artificially generated fog. The results are compared with *in-situ* measurements of the aerosol size distribution and concentration using an aerodynamic particle sizer spectrometer and a condensation particle counter.

Keywords: aerosol characterization, aerosol size distribution, multi-static lidar, multi-wavelength lidar, polarization ratio, scattering phase function

1. INTRODUCTION

The intensity distribution of light scattered from a collection of particles is dependent on properties of the particles and the incident light. The wavelength and polarization of the incident light, the refractive index, the number density, the size distribution, and the shape of the particles determine the scatter intensity. The large number of variables makes it impossible to remotely characterize unknown airborne aerosols using only backscattering measurements at a single wavelength. Additional information can be gathered by examining the scattering intensity at multiple angles. Multi-static scattering has been researched at Penn State University for a number of years^[1-5]. The angular scattering from a horizontal path of 140 meters through coastal aerosols was measured using a single linear photo-diode array and a 532 nm Nd:YAG laser^[1]. The use of the polarization ratio to minimize error and uncertainties across the field-of-view of the receiver was described by Stevens and measurements for a case of radiation fog were obtained^[2]. Vertical profiles of atmospheric aerosols were examined using three CCD cameras and a 532 nm laser^[3]. Retrieval of aerosol size distribution in the regions of strong altitude dependent layers at lower altitudes in the atmosphere was a major finding of Novitsky's research^[4]. This variation as a function of height adds significant complexity to the inversion process by removing the constraint of a uniformly mixed atmosphere within the boundary layer. More information is required for a high resolution retrieval of the vertical profile of aerosol size distributions. The current work focuses on extending the multi-static approach to multiple wavelengths and validating the approach using a stable, well-characterized environment.

*amw351@psu.edu

Multi-static, multi-wavelength scattering measurements were performed at the U.S. Environmental Protection Agency's (EPA) Aerosol Test Facility (ATF) in Research Triangle Park, North Carolina. The ATF provides technologies that are at the cutting edge of aerosol and fluid measurement to increase understanding of aerosol particle and fluid properties and the behavior of contaminants in the environment. The aerosol wind tunnel in the ATF is unique in the world in size, fluid parameter (temperature, humidity, velocity, and turbulence) control, and integration of multiple state-of-the-art aerosol systems [6]. The aerosol generation system was designed specifically for this tunnel and is capable of producing large quantities of monodisperse aerosols while maintaining uniform particle size, shape, and density. The ability to stabilize and control the scattering environment while simultaneously monitoring the aerosol size and number concentration made this a great facility for testing the multi-wavelength, multi-static scattering technique and inversion algorithm.

2. THE INVERSE PROBLEM

Mie scatter analysis can be used to accurately describe the angular scattering intensity of collimated light from a volume of spherical particles provided that the scattering volume is optically thin, and the particles are widely spaced [7]. The ensemble-averaged angular scattered intensity, $I(\theta)$, is expressed as a Fredholm integral of the first kind:

$$I(\theta) = N \int i(\theta, \alpha, m') f(a) da \quad (1)$$

where N is the number of particles, $f(a)$ is the particle size probability density function and a is the particle radius, $\alpha = 2\pi a/\lambda$ is the size parameter, λ is the wavelength of the incident light, m' is the complex refractive index of the particle, and $i(\theta, a)$ is the intensity of the scattered light of a single spherical particle. Due to the oscillatory nature of the Mie scattering intensity $i(\theta, a)$, the solution to Eq. (1) lacks uniqueness, and is an ill-posed problem. Many techniques have been used over the years to invert particle size distributions from scattering data. Recently, stochastic search techniques have entered the scene as promising alternatives to the linear search methods [8, 9]. A genetic algorithm was used by Ye et al to invert particle size distributions from simulated angular scattering data with promising results [8]. The flexibility and simplicity of the genetic algorithm were strong influencing factors in selecting this technique for the inversion of the multi-wavelength, multi-static data.

2.1 Genetic algorithms

Genetic algorithms (GA) are global optimization techniques based on the Darwinian theory of natural selection and evolution [10]. Binary strings are used to represent each of the parameters in the optimization problem and are referred to as chromosomes. Each subsection of the chromosome that represents a design parameter is called a gene. The first generation of individuals (chromosomes) is randomly generated and a function is used to determine the fitness of each individual. A method of selection is used to determine which individuals to use as parents in the mating process. Crossover and mutation are used to produce a new generation that should contain the best characteristics from the parent generation. The process repeats until a maximum generation determined by the user has been reached.

2.2 Binary coding

Atmospheric aerosol size distributions are commonly modeled as lognormal distributions or a combination of lognormal distributions. Lognormal size distributions are fully characterized using two parameters, the geometric particle radius mean ($a_{p, mean}$) and the geometric standard deviation (σ_g).

$$f(a) = \frac{N(a)}{N_{total}} = \frac{1}{\sqrt{2\pi} a \ln(\sigma_g)} \cdot \exp\left(-\frac{(\log(a) - \log(a_{p, mean}))^2}{2 \ln^2(\sigma_g)}\right). \quad (2)$$

In the genetic algorithm the geometric mean and the geometric standard deviation are represented as a binary string of zeros and ones. The binary string is translated to the lognormal parameters using the generalized formula

$$n' = n * n'_{res} + n'_{min} \quad (3)$$

where n' is either the geometric mean or the geometric standard deviation, n is the decimal equivalent of the binary string, n'_{res} is the resolution of the variables, and n'_{min} is the minimum decimal value that the variable is allowed to assume. Selecting the correct range and resolution for the size distribution variables is essential to the success of the inversion algorithm. A ten bit binary number was used to represent the geometric mean radius with a minimum radius

value of 0.01 microns and a resolution of 0.01 microns. Utilizing all 10 bits of the binary number allows the mean radius to assume a maximum radius value of 10.32 microns. The geometric standard deviation was represented by an eight bit binary number with a minimum value of 1.01, a resolution of 0.01, and a utilization of all eight bits allowed a maximum value of 3.56.

2.3 Fitness

The fitness function is the most important part of a genetic algorithm, as it is part of the algorithm that forms the connection to the physical problem being optimized. The success of the algorithm is dependent on how well the fitness function evaluates each solution in relation to the overall objectives of the optimization problem. The measured polarization ratio, which will be used as the principle data for the aerosol characterization, is formed by dividing the received scattering intensity from parallel polarized incident light by the received scattering intensity from perpendicular polarized incident light. An example of the theoretical polarization ratio of 650 nm light scattering from water droplets ($m=1.33 - i0$) with a lognormal size distribution (mean diameter of 1 micron and a geometric standard deviation of 1.2) is shown in Fig. 1.

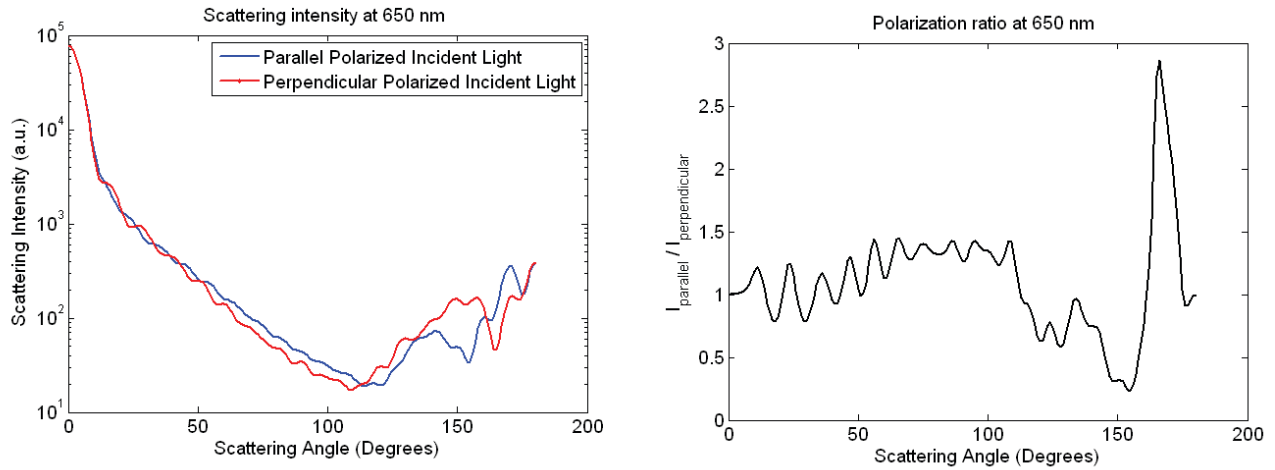


Figure 1: Theoretical scattering intensities and polarization ratio for lognormal size distribution of water droplets with a mean of 1 μm and a geometric standard deviation of 1.2

Using the polarization ratio, rather than the scattering phase function, eliminates many complexities introduced by path extinction between the scattering volume and the detector, and system parameters, such as the spectral response of the transmitter optics and the detector, and nonlinearities across the field-of-view of the detector^[1]. The theoretical polarization ratio used in the genetic algorithm as the measure of each size distributions' fitness is calculated using the following equation,

$$PR(\theta, \lambda) = \frac{\sum_{i=a_{\min}}^{a_{\max}} w_i \cdot Q_{scat}(a_i, \lambda)(S_{11}(a_i, \theta, \lambda) + S_{12}(a_i, \theta, \lambda))}{\sum_{i=a_{\min}}^{a_{\max}} w_i \cdot Q_{scat}(a_i, \lambda)(S_{11}(a_i, \theta, \lambda) - S_{12}(a_i, \theta, \lambda))} = \frac{I_{\parallel}(\theta, \lambda)}{I_{\perp}(\theta, \lambda)} \quad (4)$$

where w_i are the weights for each particle radius (a_i) determined by the probability density function generated by the genetic algorithm, Q_{scat} is the Mie scattering efficiency of each particle radius-wavelength combination, and S_{11} and S_{12} are the elements of the Mueller scattering matrix. S_{11} , and S_{12} are calculated using the Mie scattering amplitudes s_1 and s_2 as follows^[1],

$$S_{11} = \frac{1}{2} \left(|s_2|^2 + |s_1|^2 \right) \quad \text{and} \quad S_{12} = \frac{1}{2} \left(|s_2|^2 - |s_1|^2 \right) \quad (5)$$

The squared-error between the measured polarization ratios and the polarization ratios calculated using Mie theory will be used as the measure of each generated size distributions' fitness, F .

$$F = \sum_{k=1}^M \left(\sum_{j=1}^N \left(PR(\theta_j, \lambda_k) - \frac{I_{\parallel}(\theta_j, \lambda_k)}{I_{\perp}(\theta_j, \lambda_k)} \right)^2 \right) \quad (6)$$

2.4 Selection, Crossover, and Mutation

Stochastic universal sampling^[11] and sigma scaling^[12] were implemented in the algorithm. This fitness-proportionate selection method gives every individual a chance of being selected, while still ensuring that fitter candidates more likely to be chosen than weaker individuals. The individuals are mapped to a wheel, such that each individual's segment is proportionate in size to its fitness. A single random spin of the roulette wheel determines the first selected individual. The selection process then proceeds by advancing all the way around the wheel in equal sized steps, where the step size is determined by the number of individuals to be selected. Selection pressure is the probability of the best individual being selected compared to the average probability of selection of all individuals. Sigma scaling attempts to moderate selection pressure over time so that it is not too strong in early generations and not too weak once the population has stabilized and fitness differences are smaller. The standard deviation of the population fitness is used to scale the fitness values so that selection pressure is relatively constant over the lifetime of the evolutionary program. Uniform crossover was used to generate two new offspring from two selected individuals by exchanging a random number of bits. Uniform crossover is accomplished through the use of a randomly generated mask that contains the same number of binary bits as the parent chromosomes. The numbers in the mask indicate whether the bit from parent 1 or parents 2 should be translated to each child. For example, if the two chromosomes of the parents are '1 1 1 1 1' and '0 0 0 0 0', and the uniform crossover mask is '000101' then the two offspring will be '1 1 1 0 1 0' and '0 0 0 1 0 1'. Mutation is a random flipping of a bit after crossover has occurred. It is used to prevent premature convergence to a local optimum. The probability of mutation must be carefully selected to balance the risk of premature convergence and the possibility of the degrading the algorithm to a random-search method^[3]. The inversion algorithm used a one percent probability of mutation per bit.

3. EXPERIMENTAL SET-UP AND DATA PROCESSING

Three lasers at wavelengths of 407 nm, 532 nm, and 650 nm, are co-aligned into a single transmitted beam using dichroic mirrors as shown in Fig. 2. The beam is sent through a polarizing beamsplitter, which transmits only the polarization component that is parallel to the scattering plane. A broadband half-waveplate is used to rotate the polarization by 90 degrees and a turning mirror is used to direct the beams from below the floor and into the EPA wind tunnel. The laser beam traveled a distance of approximately eight meters before hitting an optical power head at the other end of the tunnel section as shown in Fig. 3. A transmission diffraction grating is used to vertically spread the three wavelengths on the CCD imager, as shown in Fig. 3, which allows separate analysis at each wavelength. The camera is a Meade Deepsky Pro II with a 49° field-of-view focusing lens. An image of the 1st order diffracted parallel polarized scattering intensity is taken as the beam propagates through the generated fog. The half-waveplate is rotated to switch the polarization from parallel to perpendicular and a second image is taken. An example of such an image is shown in Fig. 4.

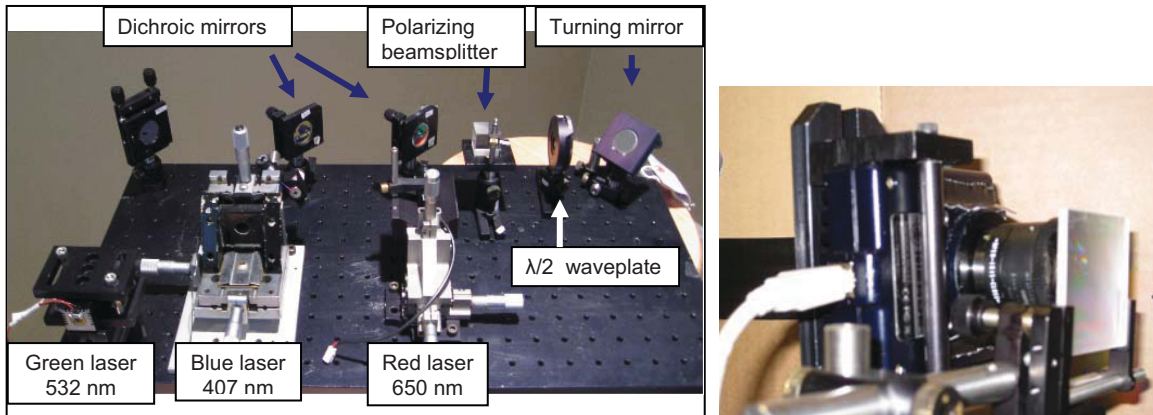


Figure 2: Picture of the (a) transmitter setup and (b) and the receiver with the diffraction grating

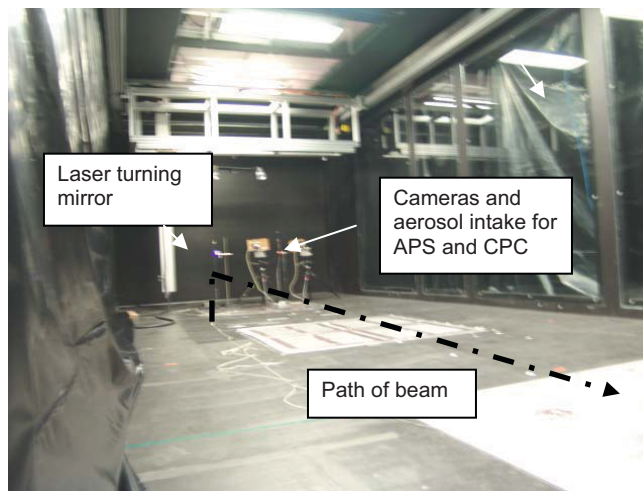


Figure 3: Picture of the experimental set up in the EPA aerosol wind tunnel

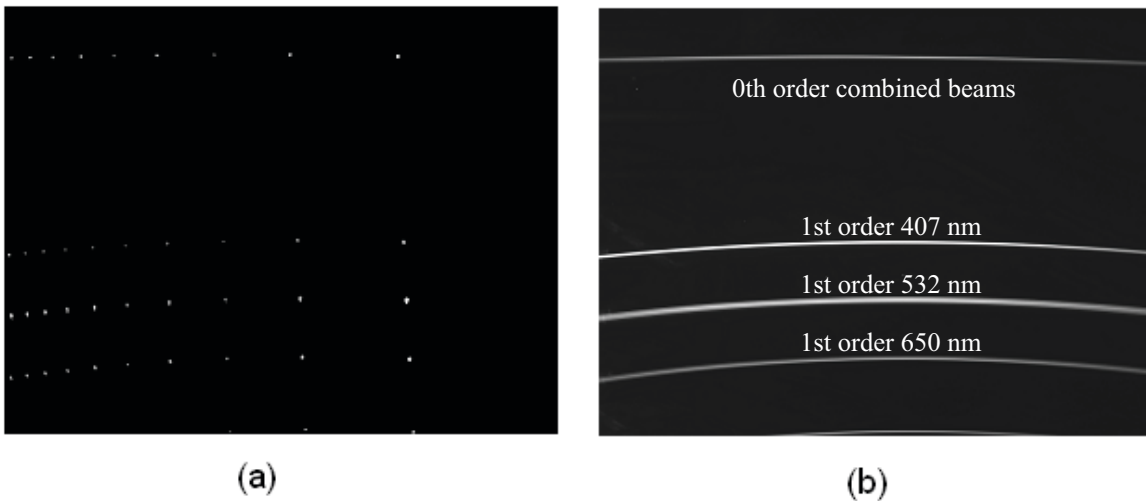


Figure 4: (a) Alignment image of the hanging wires to define measurement points and (b) CCD image of scattering from parallel polarized laser beam through the 300 lines/mm diffraction grating

The bend in the diffracted scattered beams is caused by imaging through the diffraction grating, and adds some complexity to the analysis of the image. An alignment image was used to accurately extract the spatial location and assign a scattering angle to the intensities from the image. Vertical wires were equally spaced at two feet intervals along the length of the wind tunnel section and an image is taken of the lasers striking the wires (Fig. 4a). A second order polynomial is fit to the center pixels for each wavelength and used to extract the scattering intensity from the images. The beam profile is examined for each laser wavelength as a function of column and the 10% intensity point on each side of the maximum intensity point is used to determine the number of rows to integrate across for each diffracted wavelength. The angle-dependent polarization ratio is calculated by dividing the integrated scattered intensity for incident parallel polarization by the integrated scattered intensity for incident perpendicular polarization on a column-by-column basis. The ‘raw’ polarization ratio, shown in Fig. 5a, is then smoothed using a 15 point moving-average, which corresponds to one degree, and re-sampled at one degree intervals, as shown in Fig. 5b. This ‘processed’ polarization ratio is then used by the genetic algorithm to determine the corresponding size distribution of the scatterers.

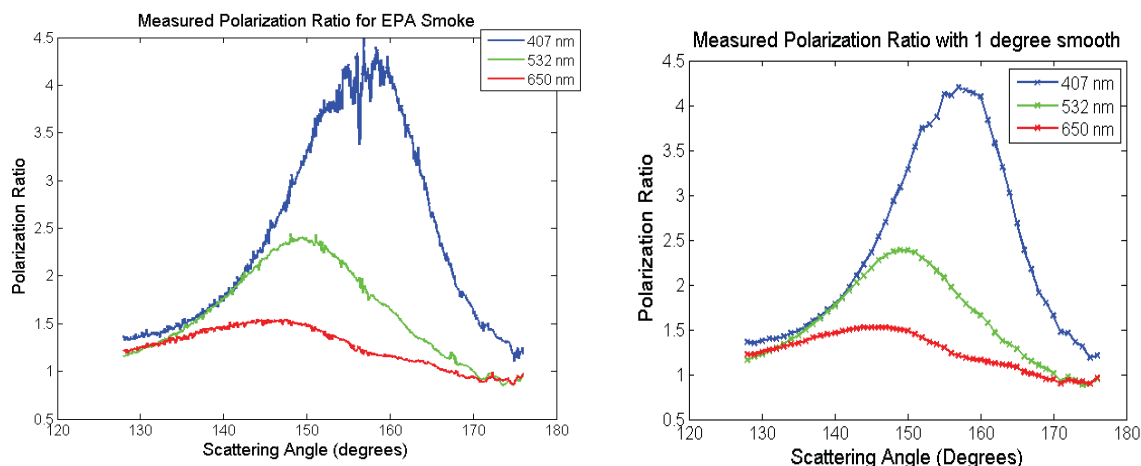


Figure 5: Example measured polarization ratio (a) ‘raw’ (b) a 1 degree moving-average smooth and re-sampled at 1 degree intervals

4. RESULTS AND COMPARISONS

Multiple sets of perpendicular and parallel scattering images were collected at backscattering angles between 120 and 175 degrees. The aerosols in the wind tunnel, produced by the MDG MAX 3000 fog generator, were simultaneously monitored by a TSI model 3321 aerodynamic particle sizer spectrometer (APS) and a TSI model 3007 condensation particle counter (CPC). The APS, using 52-channels that bins particles between ~0.5 to 20 microns, recorded averaged size distributions at one minute intervals. The CPC, which can detect particles in the size range of 0.01 to greater than 1.0 micron, recorded averaged concentration measurements every ten seconds. The point sensors were fed by a single tube that was inserted through the bottom of the wind tunnel next to the entry point for the co-aligned lasers (See Fig. 3). The total particle number concentration as measured by the two point sensors is shown in Fig. 6, along with the times of each of the collected datasets. The results from the circled datasets, labeled dataset 1 through dataset 4 in Fig. 6, are shown in this paper and the size distributions retrieved by the inversion algorithm are compared with the *in-situ* point sensor measurements. The genetic algorithm was run ten times for each dataset with a population size of 100 to the 30th generation with the resulting fitness values for each generation and the run-times averaged to produce the convergence graph in Fig. 7. The averaged run-time for the thirty generations was 57 seconds on a Pentium D 3.4 GHz processor.

A direct search algorithm was run for a truncated set of the solution space that the genetic algorithm explored to ensure that the algorithm was converging on the global solution. The direct search varied the lognormal geometric mean from 0.1 microns to 1.0 micron by 0.01 micron steps and varied the geometric standard deviation from 1.01 to 3.0 by 0.01 steps. The truncated solution space is shown in Fig. 8, and the geometric mean and standard deviation that was produced by the genetic algorithm inversion is verified by the direct search. Both algorithms arrived at the same

solutions for an assumed lognormal with a mean of 0.17 microns and a geometric standard deviation of 1.34 for data set 1, a mean of 0.19 microns and a geometric standard deviation of 1.27 for dataset 2, a mean of 0.16 microns and a geometric standard deviation of 1.30 for dataset 3, and a mean of 0.17 microns and a geometric standard deviation of 1.30 for dataset 4. The inverted lognormal probability density functions retrieved by the genetic algorithm for dataset 1 through dataset 4 are shown in Fig. 9. For comparison, the direct search run-time was ~384 seconds on the same Pentium D processor, and the genetic algorithm run-time was ~57 seconds and searched a much larger solution space.

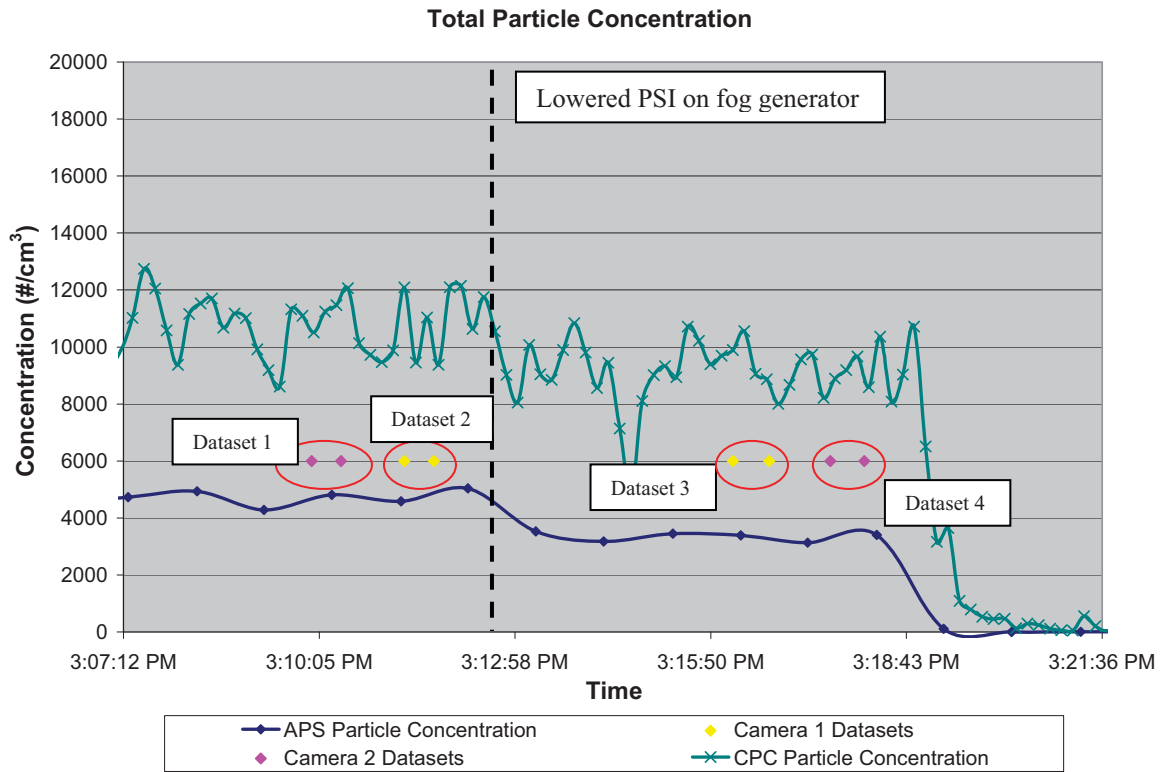


Figure 6: Total MDG fog concentration as measured by the TSI aerodynamic particle sizer spectrometer and the TSI condensation particle counter

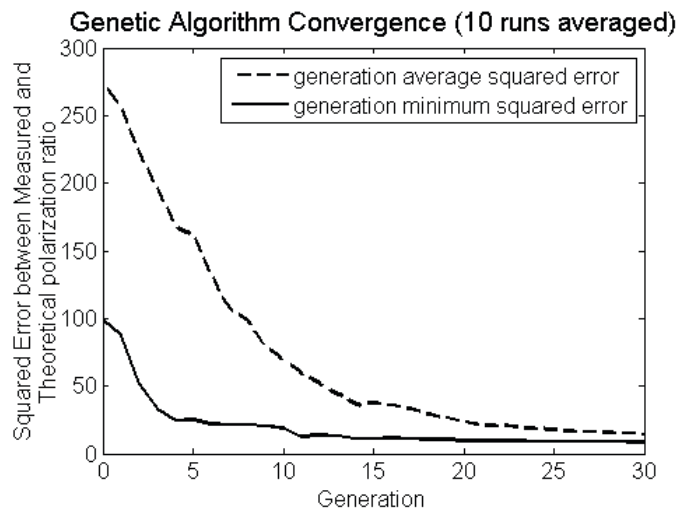


Figure 7: Averaged genetic algorithm performance for ten individual runs.

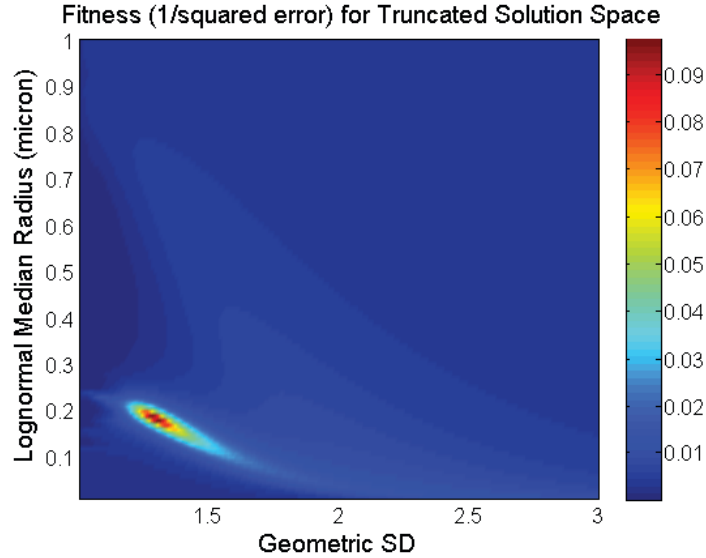


Figure 8: Truncated solution space for dataset 1.

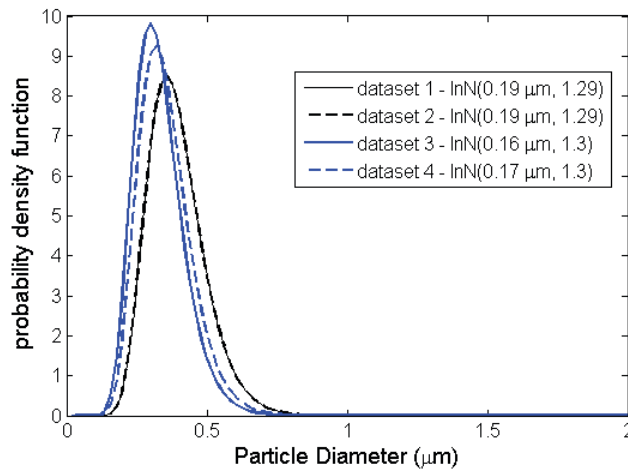


Figure 9: Probability density functions retrieved by the inversion algorithm for the four datasets

The APS normalized number concentrations ($dN/d\log(d_a)$) are converted to number concentrations (dN/dd_a) by multiplying each concentration value by the difference between the logarithm of the size bins lower and upper edge (Eqn. (5)).

$$\frac{\Delta N}{\Delta d_a} = \frac{\Delta N}{\Delta \log(d_a)} \times (\log(d_{a,upper}) - \log(d_{a,lower})) \quad (5)$$

Aerodynamic particle diameter (d_a) is then converted to geometric particle diameter (d_g) using the formula:

$$d_g = \frac{d_a}{\sqrt{SG}} \quad (6)$$

where SG is the specific gravity of the spherical particles. The specific gravity of the MDG neutral fog fluid is listed as 0.85 in the MSDS sheet supplied by the manufacture. Fig. 10 shows the inverted size distributions, which have been scaled by the average total particle concentrations measured by the CPC, and the corresponding APS size distribution

which had been measured around the approximate time that the datasets had been collected. The comparison between the measured polarization ratios and the theoretical polarization ratios when using the inverted lognormal size distribution is shown in Fig. 11 for dataset 2 and dataset 4.

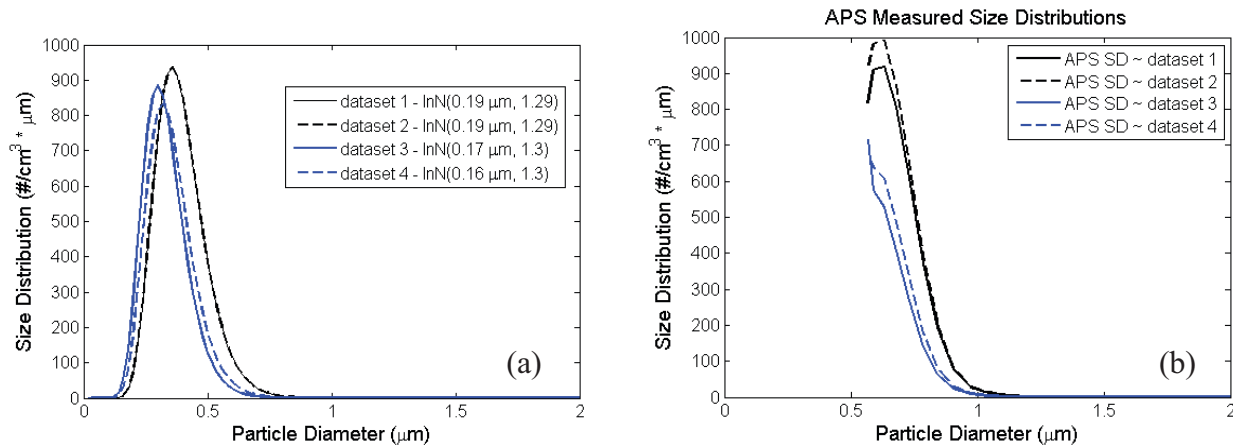


Figure 10: (a) Lognormal size distributions inverted from measured polarization ratio and (b) APS measured size distribution for datasets 1 through 4.

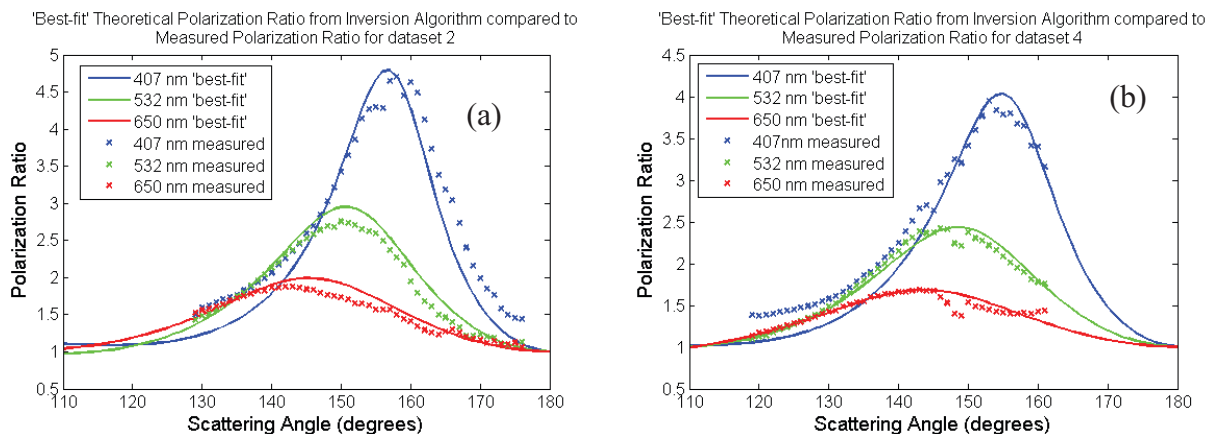


Figure 11: Comparison of measured polarization ratios and the best-fit produced from the inversion algorithm for (a) dataset 2 (b) and dataset 4.

4.1 Sources of discrepancies

The inverted size distributions predict slightly smaller particles than the APS, though in general the inverted size distributions show good agreement with the measured distributions. It is important to bear in mind that the APS and the CPC are point sensors, while the multi-wavelength scattering data will produce a path-averaged size distribution. Only a portion of the particles produced by the MDG fog generator fell within the size range of the APS, particle diameters from ~0.5 to 20 microns, making it difficult to quantify the total difference in size distribution between the APS and the inversion algorithm. The slight difference between lognormal parameters at particles diameters within the APS range is most likely due to a collection of factors. The refractive index of the fog fluid is not supplied by the manufacture, but the main ingredient listed on the MSDS sheet is mineral oil. A refractive index of 1.47-*i*0.001 was used in the inversion algorithm for all three wavelengths. In reality, the complex refractive index will vary as a function of wavelength, with the shorter wavelengths having a higher real and imaginary refractive index. Future versions of the inversion algorithm will allow the complex refractive index to vary around a reasonable ‘best-guess’ as a function of wavelength. This will hopefully improve the fit between the inverted size distribution and the APS measured size distribution.

An interesting observation came from analyzing the fit between the measured polarization ratios and the theoretical polarization ratios calculated from Mie scatter equations and the size distribution retrieved by the genetic algorithm (Fig. 11). It was noticed that the amplitudes of the polarization ratios never matched for all three wavelengths simultaneously. This led to an investigation of the initial assumptions, or that there was no wavelength-dependent polarization bias in the system. The polarization bias of the receiver was measured by imaging a white light line source onto a wall with flat white paint and placing a polarizer in front of the camera between the diffraction grating and the focusing lens, as shown on Fig. 12. Multiple sets of images were taken with various integration times with the polarizer allowing either parallel or perpendicular polarized light (with respect to the scattering plane) to pass through. A pixel-by-pixel ratio of the parallel to perpendicular intensities was formed after dark noise and readout noise was subtracted from the images. Figure 13 shows the polarization bias of the receive system when the 300 lines/mm grating and the 600 lines/mm was used to spread the diffracted light down across the CCD array. The polarization basis will affect the calculation of the polarization ratio differently for each wavelength and angle and helps explain the observed discrepancies between the measured and the theoretical polarization ratios. The errors introduced by the polarization bias would definitely contribute to the slight differences between the measured size distribution and the lognormal size distributions produced by the inversion algorithm.



Figure 12: Setup for polarization bias measurement using line source

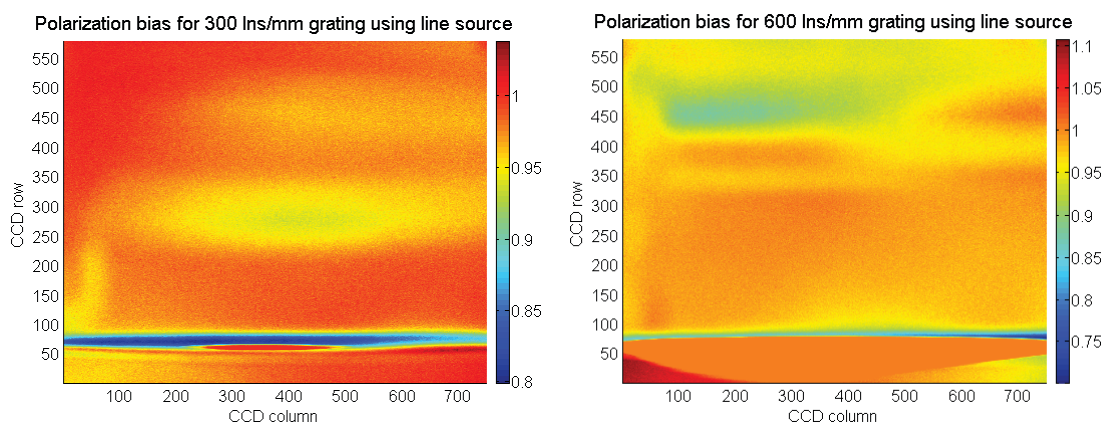


Figure 13: Polarization bias for 300 l/mm and 600 l/mm grating when using a line source

5. CONCLUSION

A multi-wavelength, multi-static system was used to collect scattering data from generated fog in the EPA Aerosol Test Facility wind tunnel. A genetic algorithm was developed to invert size distributions (probability density functions) assumed to be lognormal in shape from the polarization ratios calculated from the scattering data. The genetic algorithm converges quickly to lognormal size distributions that produced the smallest squared error between the measured polarization ratios and the polarization ratios calculated using Mie scatter equations. The inverted path-average size distributions match favorably to the *in-situ* size distributions recorded by a TSI aerodynamic particle sizer spectrometer. The simplicity of the experimental setup coupled with the ease of portability and low-cost of the necessary hardware, make the multi-wavelength, multi-static a promising technique for the characterization and monitoring of optical properties of atmospheric aerosols.

6. ACKNOWLEDGMENTS

The authors would like to thank Dr. Russell Weiner for the use of the EPA Aerosol Test Facility wind tunnel, which was an excellent test facility for our aerosol characterization research, and Zora Drake for her guidance in operating the tunnel's control and aerosol generation systems. We would also like to thank Timothy Wright from North Carolina State University for his assistance during the data collection process.

REFERENCES

- [1] Stevens, T. D. and Philbrick, C. R., "Particle size distributions and extinction determined by a unique bistatic lidar technique." Geoscience and Remote Sensing Symposium. IGARSS '96. 'Remote Sensing for a Sustainable Future', International 2: 1253-1256 (1996).
- [2] Stevens, T. D., "Bistatic Lidar Measurements of Lower Tropospheric Aerosols", Ph. D Thesis, Pennsylvania State University, (1996).
- [3] Novitsky, E. J., Multistatic Lidar Profile Measurements of Lower Tropospheric Aerosol and Particulate Matter. University Park, Pennsylvania State University. Ph. D thesis, (2002).
- [4] Novitsky, E. J. and Philbrick, C. R., "Multistatic lidar profiling of urban atmospheric aerosols." Journal of Geophysical Research, 110 (D07S11), 1-10 (2005).
- [5] Park, J. H., Multiple Scattering Measurements using Multistatic Lidar. University Park, Pennsylvania State University. Ph. D thesis, (2008).
- [6] Heist, K. D., Tolocka, M. P., Vanderpool, R. W., Peters, T. M., Chen, F.-L., and Weiner, R. W., "Changes in operating procedures for achieving aerosol concentration uniformity for PM_{2.5} and PM₁₀ sampler testing," Aerosol Science and Technology, 34, 430-432 (2001).
- [7] Bohren, C., and Huffman, D., [Absorption and Scattering of Light by Small Particles], John Wiley & Sons, New York, 83-129 (1983).
- [8] Ligon, D. A., Chen, T. W., and Gillespie, J. B., "Determination of aerosol parameters from light-scattering data using inverse Monte Carlo technique," Applied Optics, 35(21), 4297-4303 (1996).
- [9] Ye, M., Wang, S., Lu., Y., Hu, T., Zhu, Z., and Xu, Y., "Inversion of particle-size distribution from angular light-scattering data with genetic algorithms," Applied Optics, 38(12), 2677-2685 (1999).
- [10] Haupt, R. L., and Werner, D., [Genetic Algorithms in Electromagnetics], Wiley-IEEE, New Jersey, 29-43 (2007).
- [11] Baker, J., "Reducing bias and inefficiency in the selection algorithm," Proc. of the Second International Conference on Genetic Algorithms and their Application (Hillsdale), 14-21 (1987).
- [12] Goldberg, D. E., and Deb, K., "A Comparative Analysis of Selection Schemes Used in Genetic Algorithms," in: [Foundations of Genetic Algorithms], ed. by Rawlins, G.J.E., Morgan Kaufmann Publishers, San Mateo, 69-93 (1991).