Determination of Fitted Size Distribution for Atmospheric Aerosols

Kaitlin M. DuPaul*, Adam T. Whitten

Physics Department, College of Saint Benedict/Saint John's University, Collegeville, MN

Student: kmdupaul@csbsju.edu* Mentor: awhitten@csbsju.edu

ABSTRACT

A synthetic set of aerosol optical depths (AODs) generated from a standard set of aerosol size distributions was analyzed by a parameter based particle swarm optimization (PBPSO) routine in order to test the reproducibility of the results. Junge and lognormal size distributions were consistently reproduced. Gamma and bimodal distributions showed large variability in solutions. $\bar{\chi}^2$ values were used to determine the best subset of possible solutions allowing quantification of parameters with uncertainties when using PBPSO. AODs measured by a sun photometer on a clear day (20160413) and a foggy day (20160508) were then processed by the PBPSO program for both bimodal and lognormal distributions. Results showed that in general the foggy day has smaller $\bar{\chi}^2$ values indicating that the PBPSO algorithm is better able to match AODs when there is a larger aerosol load in the atmosphere. The bimodal distribution from the clear day best describes the aerosol size distribution since the $\bar{\chi}^2$ values are lower. The lognormal distribution best describes the aerosol size distribution on the foggy day (20160508).

KEYWORDS

Atmospheric Aerosols; Size Distributions; Junge; Bimodal; Gamma; Lognormal; Particle Swarm Optimization; Inverse Problem; Aerosol Optical Depth

INTRODUCTION

Atmospheric aerosols are of particular interest due to their environmental and public health impact.^{1, 2} For these reasons it is important to monitor atmospheric aerosol concentrations and size distributions. Atmospheric aerosols are tiny particles ranging in diameter from 0.001-10 µm suspended in the atmosphere. Natural aerosols are commonly made up of soil, minerals, salts, and various other chemicals mixed with water. Anthropogenic aerosol sources include industrial pollution and vehicle emissions. Aerosols are non-uniformly distributed throughout the troposphere and stratosphere with highly variable size distributions. Atmospheric aerosols can provide a surface on which chemical reactions occur, serve as condensation and ice nuclei, scatter and absorb light, and can influence the electrical properties of the atmosphere.³

Aerosol optical depth (AOD) is a common method for measurement of air quality and aerosol research. AOD can be inferred by comparing solar irradiance values to the standard E-490 solar spectrum created by the American Society for Testing and Materials⁴ and subtracting off Rayleigh scattering terms. Higher AODs correspond to more scattering and/or absorption of sunlight, indicating that one or more of the aerosol size distribution properties has increased (i.e., particle number, particle size, particle shape, and/or index of refraction). Given a set of AODs, the aerosol size distribution may be inferred by inverting a set of Fredholm equations,

$$\tau(\lambda) = \int \pi r^2 Q_{ext}(r, \lambda, \eta) n(r) dr$$
 Equation 1.

where Q_{ext} is the Mie extinction coefficient and n(r) is the aerosol size distribution. Matrix inversion,⁵⁻⁷ iterative,⁸⁻¹⁰ and maximum entropy spectral¹¹ methods have been used to retrieve aerosol size distributions. However, inversion of **Equation 1** is an ill-posed problem classified as a NP (nondeterministic polynomial) problem meaning that the inversion can be performed using the twostep process for solving any NP problem: (1) guess a solution in a non-deterministic way and (2) use a deterministic algorithm to verify or reject the guess.¹² It is important to note that accepting a guess does not guarantee that the best possible solution has been generated; it merely indicates that the solution is one of many possible valid solutions that fit the deterministic criteria.

Recently, particle swarm optimization (PSO) algorithms that implement the two-step NP problem solving method have been used to retrieve size distributions from **Equation 1**. PSO is an iterative algorithm that uses a swarm of particles randomly generated in the solution space to identify size distribution parameters that yield calculated AODs close to those inferred from solar irradiance measurements. Closeness is quantified by a fitness function and the iterations continue until a stop criteria that depends on the fitness function is reached or until a predefined maximum number of iterations have been performed. The goal of PSO is to

minimize the fitness function by moving each particle in the swarm simultaneously toward its own and the swarm's previous best positions.^{13, 14} Each particle's position is used to calculate an AOD and fitness value; the particle with the smallest fitness value is presumed to represent the best inversion of the AODs inferred from irradiance measurements and corresponds to the best position. However, given the nature of NP problems, multiple runs of PSO algorithms will in general converge to different solutions corresponding to local minima or the global minimum in the essentially infinite solution space.^{15, 16}

Examples of PSO algorithms in use include Yuan *et al.*'s (2010, 2011) Stochastic Particle Swarm Optimization (SPSO)^{17, 18} and Mao & Li's (2015) Improved Particle Swarm Optimization (IPSO).¹⁹ Both use a fitness function based on the difference between measured and inferred AODs: SPSO uses the square root of the average variance; IPSO uses the inverse of half the sum of the variances. SPSO's stop criteria are (1) when changes in the best fitness function are less than 10^{-10} and (2) when 1000 iterations have been performed. IPSO stops after 30 iterations without placing constraints on fitness function convergence. The remaining difference between these two algorithms is that SPSO generates a random particle with each iteration in order to explore more of the solution space while IPSO uses a variable velocity to speed up convergence. Neither method specifies what type of boundary conditions are employed. Neither paper discusses the reproducibility of results when the algorithms are run on the same set of data more than once. There is no measure of how often their algorithms may focus in on a local minimum of the fitness function instead of the global minimum. Furthermore, Yuan *et al.*'s (2011) stop criterion of the fitness function changing by less than 10^{-10} restricts the uncertainty in AODs inferred from spectral measurements (typically 0.1-2.5%^{20, 21}) by many orders of magnitude.

To address these deficiencies, a parameter-based particle swarm optimization (PBPSO) algorithm was written. Boundary conditions are such that particles moving outside the solution space are destroyed and a new random particle is generated. This boundary condition allows more exploration of the solution space. During each iteration, ten additional particles are generated to test for better solutions and then are destroyed. PBPSO's fitness function is the reduced chi-squared of calculated and inferred AODs. Like previous algorithms, the fitness function for PBPSO is used to guide evolution of the swarm towards a best solution that is, ideally, the global minimum. Unlike previous algorithms, the stop criterion is based on the relative changes in size distribution parameters, not on the fitness function or some fixed number of iterations. Using a size parameter convergence criterion ensures that a swarm has converged to a solution. The PBPSO stop criteria are (1) the size distribution parameters have a relative change of less than $10^{-3\%}$ for 100 consecutive iterations or (2) the algorithm reaches 5000 iterations. After multiple runs of PBPSO on the same data set, the quality of each solution can be assessed using the corresponding fitness function values.

METHODS AND PROCEDURES

This research tests the reproducibility of the PBPSO algorithm for four different size distributions. This algorithm implements the following steps to solve the inverse radiation problem.

<u>Step 1</u>: Input aerosol optical depths at 412, 441, 463, 479, 500, 520, 556, 610, 675, 750, 778, 870, and 1020 nm. These include the World Meteorological Organization (WMO) standard wavelengths (412, 500, 610, 675, 778 nm), WMO standard wavelengths shifted out of absorption bands (463, 870, 1020 nm), and five additional wavelengths (441, 479, 520, 556, 750 nm).²² Input the Mie extinction coefficient, Q_{ext}, for each of these wavelengths spaced at radius intervals $\delta r = 0.001 \,\mu$ m. Choose bounds for the parameters depending on distribution type. Initialize a particle swarm and calculate AOD, $\tau(\lambda)$, using **Equation 2** at each wavelength, λ , for each particle. Q_{ext} was calculated by using the Bohren-Huffman Mie scattering subroutine²³ from $r_{min} = 0.04 \,\mu$ m to $r_{max} = 10 \,\mu$ m with a complex index of refraction $\eta = 1.5+0i$ giving a 2×9961 array. r_{min} was chosen as 0.04 μ m since for $r \le 0.04 \,\mu$ m, Q_{ext} < 0.0012 and the contribution to the AOD is negligible. The real part of η was chosen to be between water (1.33) and silica (1.54).

$$\tau(\lambda) = \int_{r_{min}}^{r_{max}} \pi r^2 Q_{ext}(r,\lambda,\eta) n(r) dr \approx \pi \delta r \sum_{i=1}^{9961} r_i^2 Q_{ext}(r_i,\lambda,\eta) n(r_i) \quad \text{where} \quad n(r) = \frac{dN(r)}{dr}$$
Equation 2.

<u>Step 2</u>: Calculate reduced chi-squared ($\bar{\chi}^2$) value (**Equation 3**) for each particle where d = 13-# of independent parameters.

$$\bar{\chi}^2 = \frac{1}{d} \sum_{i=1}^{13} \left(\frac{\tau_{i,obs} - \tau_{i,calc}}{\tau_{i,calc}} \right)^2$$
Equation 3.

<u>Step 3</u>: Compare current $\bar{\chi}^2$ with the previous $\bar{\chi}^2$ for each particle to determine if the new position is lower; if so it becomes a new local minimum, P_i, otherwise retain the previous local minimum.

<u>Step 4</u>: From the list of all P_i choose the one with the lowest $\bar{\chi}^2$ and set it as the global minimum, P_g.

<u>Step 5</u>: Introduce ten randomly generated particles and see if their $\overline{\chi}^2$ s are less than P_g; if so replace P_g with the lower value.

<u>Step 6</u>: Check the stop criteria. (1) If the size distribution parameters have changed by less than 10^{-30} % for 100 consecutive iterations then the program is terminated (**Equation 4**). (2) If the program has gone through 5000 iterations without reaching the stop criterion (1) terminate the program. Otherwise evolve the swarm and loop back to step 2.¹³

$$\frac{1}{j} \sum_{i=1}^{j} \frac{\left| P_g[i](t+1) - P_g[i](t) \right|}{P_g[i](t)} \le 10^{-5}$$
Equation 4.

<u>Step 7</u>: Repeat steps 1-6 ten times. Examine the final $\overline{\chi}^2$ s for each run's P_g: the lowest $\overline{\chi}^2$ is corresponds to the solution that is most likely to be the global minimum while the others correspond to solutions that are local minima. Since there is still no guarantee that the global minimum has been found, use a subset of the ten solutions defined by $\overline{\chi}^2 \leq 100 \overline{\chi}_{min}^2$ to find size distribution parameter averages and standard deviations of the mean (SDOM). This process eliminates solutions that are deemed valid by step 2 of the NP selection process, but which represent outliers corresponding to local minima.

A synthetic set of AODs was generated from a standard set of size distribution parameters (N, α , and β) for each distribution (**Table 1**) in order to test the PBPSO algorithm. Aerosol optical depths at the 13 wavelengths were calculated for the four distributions listed in **Table 1** for an assumed set of values for the parameters (N₀, α , and β for the single mode distributions and α_0 , β_0 , N₁, α_1 , and β_1 for the bimodal distribution) using the Bohren-Huffman Mie scattering subroutine.²³ Parameter values were chosen to give AODs that might typically be obtained from a Kipp-Zonen PGS-100 sun photometer. These AODs were then used as input to the PBPSO program to evaluate how well the algorithm could reproduce the distribution parameters. The resulting $\bar{\chi}^2$ values are a measure of the difference between the algorithm's calculated AODs and the inputted AODs. Since multiple runs might produce different parameter values, it was hypothesized that choosing a subset of the results with the smallest $\bar{\chi}^2$'s and then finding the mean and standard deviation of the mean (SDOM) of this subset of parameters would give a reasonable representation of the actual size distribution. The PBPSO algorithm was run ten times for each distribution type and the results were plotted against the known distribution. These results were evaluated as a standard for identifying the best way to analyze atmospheric data.

Distribution	Equation	Parameters	Solution Space
Junge	$n_J(r) = N_0 \beta r^{-(\alpha+1)}$	N ₀ =2.5×10 ⁶ , α=3, β=0.3	$\begin{array}{l} 1.0{\times}10^3 \leq N_0 \leq 1.0{\times}10^9 \\ 2.0 \leq \alpha \leq 5.0 \\ 0.1 \leq \beta \leq 0.5 \end{array}$
Gamma	$n_G(r) = N_0 r^\beta e^{-\alpha r}$	$N_0=2.5\times10^9, \alpha=0.1, \beta=0.35$	$\begin{array}{l} 1.0 \times 10^6 \leq N_0 \leq 1.0 \times 10^{12} \\ 0.05 \leq \alpha \leq 0.50 \\ 0.1 \leq \beta \leq 0.5 \end{array}$
Lognormal	$n_L(r) = rac{N_0}{\sqrt{2\pi}} rac{1}{eta r} e^{5 \left[rac{ \ln(r/lpha) }{eta} ight]^2}$	N ₀ =1.0×10 ⁷ , α=0.5, β=0.307	$\begin{array}{l} 1.0 \times 10^4 \leq N_0 \leq 1.0 \times 10^{10} \\ 0.05 \leq \alpha \leq 1.0 \\ 0.1 \leq \beta \leq 0.5 \end{array}$
Bimodal	$n_B(r) = \frac{N_0}{\sqrt{2\pi}} \frac{1}{\beta_0 r} e^{-5\left[\frac{\ln(r/\alpha_1)}{\beta_1}\right]^2} + \frac{N_1}{\sqrt{2\pi}} \frac{1}{\beta_1 r} e^{-5\left[\frac{\ln(r/\alpha_1)}{\beta_1}\right]^2}$	$\begin{array}{l} N_0 {=}1.0 {\times} 10^7, \alpha_0 {=} 0.1, \beta_0 {=} 0.307 \\ N_1 {=} 2.5 {\times} 10^5, \alpha_1 {=} 2, \beta_1 {=} 0.307 \end{array}$	$\begin{array}{l} 1.0{\times}10^5 \leq N_0 \leq 1.0{\times}10^{11} \\ 0.05 \leq \alpha_0 \leq 0.50 \\ 0.1 \leq \beta_0 \leq 0.5 \\ 1.0{\times}10^3 \leq N_1 \leq 1.0{\times}10^9 \\ 0.50 \leq \alpha_1 \leq 5.0 \\ 0.1 \leq \beta_1 \leq 0.5 \end{array}$

Table 1. Typical distributions and solution space. To access synthetic data sets used and results, see http://www.physics.csbsju.edu/~awhitten/pso.html.

RESULTS AND DISCUSSION

Synthetic Data

For Junge and lognormal distributions there was little to no variation in the results, but for the gamma and bimodal distributions the results seemed to vary so additional runs of the algorithm were made (1) holding α constant while allowing β to vary and (2) holding β constant while allowing α to vary.

Junge Distribution

After analyzing the results for the Junge distribution it was noticed that the plots of the resulting calculated number density distributions were identical. For all ten runs of the PBPSO program the $\bar{\chi}^2$ values are equal (2.05× 10⁻¹³) and although there is variance in N₀ and β , α is consistently 3.0. The variation in N₀ and β is explained by the fact that they are not independent

parameters and their variability is constrained only by the choice of solution space given in Table 1. Considered separately the ten values of N_0 and β have distributions that are narrower than a normal distribution (all values fall within the 95th percentile), but because of their dependent nature it is their product, not the product of their individual averages, that is important. The PBPSO program gave $N_0 \times \beta = 7.5 \times 10^5$ for all ten runs. Furthermore, the PBPSO program was able to exactly reproduce the AOD's for each wavelength every time.

Lognormal Distribution

The lognormal distribution of the sample data shows a generally uniform distribution with two small outliers (**Figure 2**). These two outliers prove to be runs two and eight which correspond to the results with the highest $\bar{\chi}^2$ values (~10⁻⁷) and shouldn't be taken into consideration as the lowest $\bar{\chi}^2$ value for this data set is 3.3×10^{-13} . The distribution parameter values were N₀ = 2.5×10^6 cm⁻³, $\alpha = 0.5 \mu$ m, and $\beta = 0.307$. As shown in **Table 2**, the output means for each variable were in agreement with the distribution parameter values when taking the standard deviation of the mean (SDOM) into account.



Figure 2. Size distribution functions output from the PBPSO algorithm reveal a tight grouping. The actual distribution function is hidden behind the red line.

		N_0 (cm ⁻³)	α (µm)	β	
	Mean	2.500×106	0.5004	0.30698	
	SDOM	3×103	0.0003	0.00001	
Table 2	. Lognorm	al results with	neither α r	or β held c	onstant

Gamma Distribution

The gamma distributions showed some small, but significant variations. Therefore these distributions have been plotted on a smaller scale for the $\frac{dN}{dr}$ (cm⁻³) axis in order to highlight the difference between runs. Despite these apparent differences, the PBPSO routine was able to reproduce the target AODs with $\bar{\chi}^2$ values ranging from 10⁻¹¹ to 10⁻⁹.



Figure 3. The actual distribution function is represented as a dark blue line in each of the graphs. The PBPSO program has more trouble reproducing the gamma distribution as shown in Figure 3.1. When α is held constant the PBPSO program can zero in on a wide range of solutions with smaller N₀ and β (a wider distribution) or higher N₀ and β (a narrower distribution) as shown in Figure 3.2. When β is held constant the PBPSO program can produce a more consistent set of solutions as shown in Figure 3.3. This suggests that falling in to a local minimum for β can greatly influence the results (see text).

Gamma-no constant

When no parameters are held constant, a wide range of solutions are seen (**Figure 3.1**). To understand why this might be the case, additional optimizations were performed holding α or β constant. When α is held constant, β and N₀ vary significantly resulting in little to no overlap of the values (**Figure 3.2**). With β held constant, α varies from [0.06, 0.22], but N₀ stays relatively constant (**Figure 3.3**). This behavior is the result of the analytic forms of the Mie extinction coefficient and the gamma distribution. For visible and near infrared wavelengths Q_{ext} peaks at about 4.4 for $r < 1 \,\mu$ m and then exhibits decreasing oscillations as shown in **Figure 4**.



Figure 4. Mie extinction coefficient, Q_{ext} , calculated using the Bohren-Huffman subroutine for three wavelengths of light and complex index of refraction $\eta = 1.5+0i$. The values at large radii decrease to approximately one half of the peak values that occur at radii less than 1 μ m.

This variation in Q_{ext} means that small changes in the contribution to AODs from small radii particles can be compensated by larger changes from large radii particles to yield the same AOD value. As β increases in the gamma distribution the r^{β} term reduces contributions to the AOD for $r < 1 \mu m$ and increases contributions to the AOD for $r > 1 \mu m$. The exponential term reduces contributions to the AOD for $r > 1/\alpha$, but when combined with the r^{β} term it results in a peak in the size distribution at $r_0 = \beta/\alpha$. When α is held constant at 0.1, then $r_0 = 10\beta$ constraining the peak of the distribution to the range $1.0 \mu m \le r \le 5.0 \mu m$

which is greater than the radii of the peak values of Q_{ext} . Larger values of β give larger values of N_0 because the decrease from r_0 down to $r < 1 \mu m$ (where Q_{ext} is highest) is more rapid. When β is held constant at 0.35, $r_0 = 0.35/\alpha$ constraining the peak of the distribution to 0.7 $\mu m \le r \le 7.0 \mu m$. This range for r_0 is also greater than the radii of the peak values of Q_{ext} , but since β is fixed, the contribution to the AOD from small radii particles is well defined. Therefore, the variability in α and N_0 with β fixed is smaller than the variability in β and N_0 with α fixed.

For the case when neither α nor β are held constant the combined effect of the dependence on radius of both Q_{ext} and the gamma distribution makes it difficult for PBPSO to zero in on the correct solution because multiple solutions meet the fitness criterion. Therefore, to use this PBPSO routine to retrieve gamma size distribution parameters it is recommended to run the optimization multiple times and extract the best retrievals based on $\bar{\chi}^2$ values. Using solutions with $\bar{\chi}^2 \leq 100 \bar{\chi}_{min}^2$ eliminates outliers corresponding to local minima and the remaining solutions are used to calculate the means and SDOMs of the size distribution parameters (**Tables 3.1-3.3**). The set parameter values were N₀ = 2.5×10⁹ cm⁻³, $\alpha = 0.1 \mu m^{-1}$, and $\beta = 0.35$. Not shown are the results when α and β are both held constant. This constraint consistently reproduces the correct N₀ in agreement with the results of Yuan *et al.* 2010.

	N_0 (cm ⁻³)	α (μm ⁻¹)	β
Mean	2.50×109	0.19	0.348
SDOM	1.5×10 ⁸	0.03	0.009
			0.1.1.1

Table 3.1. Gamma results with neither α nor β held constant.

	N ₀ (cm ⁻³)	α (µm-1)	β						
Mean	3.1×10 ⁹	0.1	0.37						
SDOM	5×10^{8}	constant	0.02						
able 3.2. Gamma results with α held constant									

	N_0 (cm ⁻³)	α (µm-1)	β					
Mean	2.5000×10^{9}	0.11	0.35					
SDOM 3×10 ⁵ 0.02 constant								
Table 3.3. Gamma results with β held constant.								

Bimodal Distribution

The bimodal distributions shown in **Figure 5** indicate large variations in the retrieved size distribution parameters despite being able to reproduce the target AODs with small $\bar{\chi}^2$ values. **Figure 5.1** shows that when no parameters are held constant there is variability in solutions that give similar aerosol optical depths $(10^{-9} < \bar{\chi}^2 < 10^{-5})$. To examine this variability and understand which parameters PBPSO has trouble determining, α and/or β were held constant. When α is held constant (**Figure 5.2**), N₀ and β_0 stay relatively constant, but N₁ and β_1 do not $(10^{-11} < \bar{\chi}^2 < 10^{-5})$. When β is held constant (**Figure 5.3**) there is also more variability in N₁ and α_1 than in N₀ and α_0 $(10^{-9} < \bar{\chi}^2 < 10^{-7})$. PBPSO is less consistent when determining the distribution at larger radii. The source of this inconsistency is most likely the relatively flat Mie extinction profiles at larger radii for visible wavelengths of light making it difficult for the algorithm to pick out α_1 . It should be noted that when α_0 , α_1 , β_0 and β_1 are all held constant, N₀ and N₁ values are exactly reproduced which is consistent with the findings in Yuan *et al.* 2011.

To quantify the inconsistencies, subsets of results based on $\bar{\chi}^2 \leq 100 \bar{\chi}_{min}^2$ values were analyzed to find the means and SDOMs of the size distribution parameters. For no constants (**Figure 5.1**) run seven (light yellow) is the standard and solutions with a $\bar{\chi}^2 > 3.43 \times 10^{-7}$ were ruled out, resulting in five outliers for this set. In **Figure 5.2** run seven (light yellow) is the standard as it has the lowest $\bar{\chi}^2$ value for the data set (2.7×10^{-11}). After applying the $\bar{\chi}^2$ criterion three of the ten solutions are left. In **Figure 5.3** run six (light blue) is the standard with $\bar{\chi}^2 = 1.0 \times 10^{-9}$. It is much easier to identify the outliers in this plot as they all lie within the second lognormal distribution and after applying the $\bar{\chi}^2$ criterion five of the ten solutions are left. In order to have significant averages for the six size distribution parameters it is necessary to eliminate outliers.

The distribution parameters values were $N_0 = 1.0 \times 10^7$ cm⁻³, $\alpha_0 = 0.1 \mu m$, $\beta_0 = 0.307$, $N_1 = 2.5 \times 10^5$ cm⁻³, $\alpha_1 = 2.0 \mu m$, and $\beta_1 = 0.307$. Following the $\bar{\chi}^2$ rule stated above, the resulting size distribution parameters for no constants, both α s held constant, and both β s held constant are shown in **Tables 4.1-4.3**.



Figure 5. The actual distribution function is indicated in dark blue. Figure 5.1 shows that when all parameters are allowed to vary there are multiple solutions that reproduce the aerosol optical depths. Figure 5.2 shows that when α is held constant there is more variability in the width of larger aerosols. In Figure 5.2 the actual distribution is hidden behind the dark green line. Figure 5.3 shows that when β is held constant there is more variability in the radius of larger aerosols.

	N_0 (cm ⁻³)	α₀ (µm)	β0	N_1 (cm ⁻³)	α1 (μm)	β1					
Mean	9.20×107	0.11	0.29	2.6×10 ⁵	2.5	0.24					
SDOM	DOM 1.5×10^6 0.01 0.02 5×10^4 0.4										
Tal	Table 4.1. Bimodal results with neither α nor β held constant.										

	N_0 (cm ⁻³)	α ₀ (μm)	β0	N_1 (cm ⁻³)	α ₁ (μm)	β_1					
Mean	1.030×107	0.1	0.3056	2.8×10^{5}	2	0.26					
SDOM 1.0×10 ⁵ constant 0.0004 7×10 ⁴ constant 0.02											
	Table 4.2 Bimodal results with a held constant										

	N_0 (cm ⁻³)	α ₀ (μm)	β0	N_1 (cm ⁻³)	α ₁ (μm)	β1				
Mean	9.10×106	0.09990	0.307	1.62×10 ⁵	2.8	0.307				
SDOM 1.0×10 ⁵ 0.00004 constant 8×10 ³ 0.1 constant										
Table 4.3. Bimodal results with β held constant.										

Atmospheric Data

The PBPSO procedure is now applied to data taken by a Kipp & Zonen PGS-100 sun photometer on a clear day (2016-04-13) and a foggy day (2016-05-08). Aerosol optical depths inferred from irradiance measurements at the St. John's University observatory were processed ten times with the PBPSO algorithm. Solutions with $\bar{\chi}^2 > 100\bar{\chi}_{min}^2$ were eliminated and the remaining solutions were used to find the averages and SDOMs for size distribution parameters. Data from the two days was processed using a bimodal size distribution and a lognormal size distribution.

Bimodal Distribution

Aerosol optical depths at 13 wavelengths inferred from the PGS-100 and calculated for each of the tens runs of PBPSO for the clear day are shown in **Table 5**. Inferred AODs are reproduced by the PBPSO algorithm to within 3–7% for all wavelengths except for 610 nm, 750nm, and 778 nm where they are within 12–14%.

AOD						Way	elength (nm)						
from:	412	441	463	479	500	520	556	610	675	750	778	870	1020	$\bar{\chi}^2$
PGS-100	0.241	0.207	0.201	0.200	0.191	0.195	0.187	0.181	0.131	0.103	0.095	0.091	0.070	-
Run 1	0.227	0.219	0.211	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.067	1.4E-3
Run 2	0.230	0.220	0.211	0.205	0.197	0.189	0.176	0.156	0.136	0.115	0.108	0.089	0.065	1.5E-3
Run 3	0.229	0.220	0.212	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3
Run 4	0.230	0.219	0.211	0.205	0.197	0.190	0.176	0.157	0.136	0.115	0.108	0.089	0.065	1.4E-3
Run 5	0.229	0.220	0.212	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.067	1.4E-3
Run 6	0.230	0.220	0.211	0.205	0.197	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3
Run 7	0.229	0.220	0.211	0.205	0.197	0.190	0.176	0.156	0.135	0.114	0.108	0.088	0.066	1.4E-3
Run 8	0.229	0.220	0.211	0.205	0.197	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3
Run 9	0.230	0.220	0.211	0.205	0.197	0.189	0.176	0.156	0.135	0.115	0.108	0.089	0.066	1.4E-3
Run 10	0.228	0.220	0.212	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3

 Table 5. Aerosol optical depths for a clear day (20160413) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a bimodal distribution.

Table 6 shows the AODs inferred from the PGS-100 and calculated for each of the tens runs of PBPSO for the foggy day. Inferred AODs are reproduced to within 0.4–7% for all wavelengths.

AOD						Way	velength (nm)						
from:	412	441	463	479	500	520	556	610	675	750	778	870	1020	$\overline{\chi}^2$
PGS-100	0.704	0.633	0.611	0.590	0.564	0.540	0.503	0.452	0.347	0.274	0.250	0.205	0.146	_
Run 1	0.691	0.652	0.620	0.597	0.566	0.539	0.490	0.422	0.352	0.285	0.263	0.205	0.141	7.0E-4
Run 2	0.691	0.652	0.619	0.596	0.566	0.539	0.490	0.422	0.352	0.285	0.264	0.206	0.140	7.1E-4
Run 3	0.691	0.652	0.619	0.597	0.566	0.539	0.490	0.422	0.352	0.285	0.263	0.206	0.141	7.1E-4
Run 4	0.693	0.652	0.619	0.596	0.565	0.537	0.488	0.420	0.351	0.285	0.264	0.207	0.143	7.2E-4
Run 5	0.691	0.652	0.620	0.597	0.566	0.539	0.490	0.421	0.352	0.285	0.263	0.205	0.142	7.0E-4
Run 6	0.689	0.651	0.618	0.596	0.566	0.538	0.490	0.422	0.352	0.286	0.264	0.206	0.140	7.2E-4
Run 7	0.691	0.652	0.620	0.597	0.566	0.539	0.490	0.421	0.352	0.284	0.263	0.205	0.142	6.8E-4
Run 8	0.687	0.650	0.618	0.596	0.567	0.539	0.491	0.424	0.354	0.287	0.265	0.207	0.140	7.2E-4
Run 9	0.690	0.651	0.619	0.597	0.566	0.539	0.490	0.422	0.352	0.285	0.264	0.205	0.140	7.0E-4
Run 10	0.691	0.651	0.619	0.596	0.566	0.538	0.490	0.422	0.353	0.286	0.264	0.206	0.138	7.4E-4

Table 6. Aerosol optical depths for a foggy day (201600508) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a bimodal distribution.



Figure 6. For the given days 20160413 in Figure 6.1 and 20160508 in Figure 6.2, all parameters are allowed to vary and there are multiple solutions that reproduce the aerosol optical depths. The reduced chi-squared values in the bimodal distribution are closer with little to no variance.

The bimodal results (**Figures 6.1-6.2**) show that the $\bar{\chi}^2$ values are all within 10⁻⁴ of each other with little to no variance so following the earlier stated rule all values must be considered. For the 20160413 results there are two groups of solutions roughly 40% of which are narrower with higher α_1 values and 60% that are wider with lower α_1 values. Results show a 4% uncertainty in the median radius α_0 of the small aerosol mode for the clear day and a 2% uncertainty in determining the median radius α_0 of the small aerosol mode for the large mode aerosols the uncertainties in the median radius α_1 are 18% and 16% for the clear and foggy days, which confirms the assessment for synthetic data that the median radius of the small aerosol mode is easier to pick out than the large aerosol mode (**Tables 7.1-7.2**).

	N_0 (cm ⁻³)	α₀ (µm)	β0	N_1 (cm ⁻³)	α ₁ (μm)	β1				
Mean	2.78×107	0.167	0.147	3.5×104	2.2	0.212				
SDOM 2.7×10 ⁶ 0.007 0.009 1.2×10 ⁴ 0.4 0.02										
Table 7.1. Bimodal results 20160413 (clear day).										

	N_0 (cm ⁻³)	α₀ (µm)	β0	N_1 (cm ⁻³)	α ₁ (μm)	β1
Mean	8.6×107	0.171	0.124	7.3×104	1.12	0.214
SDOM	4×10 ⁶	0.003	0.005	1.6×10^{4}	0.18	0.022

Table 7.2. Bimodal results 20160508 (foggy day).

Lognormal Distribution

Aerosol optical depths for the 13 wavelengths inferred from PGS-100 measurements and calculated for 10 runs of the PBPSO algorithm are shown in **Table 8** for the clear day and **Table 9** for the foggy day. AODs are reproduced to within 4–19% for the clear day and to within 0.2–7% for the foggy day.

AOD	Wavelength (nm)									2				
from:	412	441	463	479	500	520	556	610	675	750	778	870	1020	$\bar{\chi}^2$
PGS-100	0.230	0.164	0.162	0.159	0.154	0.165	0.156	0.156	0.100	0.086	0.083	0.099	0.084	-
Run 1	0.206	0.189	0.178	0.170	0.162	0.154	0.142	0.127	0.113	0.102	0.098	0.089	0.081	3.5E-3
Run 2	0.206	0.189	0.178	0.170	0.162	0.154	0.142	0.127	0.113	0.102	0.098	0.089	0.081	3.5E-3
Run 3	0.203	0.189	0.178	0.171	0.163	0.155	0.143	0.127	0.113	0.101	0.097	0.089	0.081	3.5E-3
Run 4	0.205	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 5	0.205	0.189	0.178	0.171	0.162	0.155	0.143	0.127	0.112	0.099	0.095	0.088	0.087	3.3E-3
Run 6	0.204	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 7	0.206	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 8	0.205	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 9	0.203	0.189	0.178	0.171	0.163	0.155	0.143	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 10	0.204	0.189	0.178	0.171	0.163	0.156	0.143	0.127	0.112	0.099	0.095	0.087	0.086	3.3E-3

Table 8. Aerosol optical depths for a clear day (20160413) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a lognormal distribution.

AOD	Wavelength (nm)													
from:	412	441	463	479	500	520	556	610	675	750	778	870	1020	$\bar{\chi}^2$
PGS-100	0.704	0.633	0.611	0.590	0.564	0.540	0.503	0.452	0.347	0.274	0.250	0.205	0.146	-
Run 1	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 2	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 3	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 4	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 5	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 6	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 7	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 8	0.695	0.652	0.618	0.594	0.563	0.536	0.487	0.421	0.353	0.288	0.266	0.209	0.142	5.5E-4
Run 9	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 10	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4

Table 9. Aerosol optical depths for a foggy day (201600508) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a lognormal distribution.



Figure 7. For the given days 20160413 in Figure 7.1 and 20160508 in Figure 7.2, all parameters are allowed to vary and there are multiple solutions that reproduce the aerosol optical depths. The solutions for the lognormal distribution are almost perfect and extremely uniform.

The lognormal results (**Figures 7.1-7.2**) show that once again the $\bar{\chi}^2$ values are almost identical; all being equal or off by 10⁻⁴ of each other for the clear day and 10⁻⁵ for the foggy day. Therefore all solutions are used to find the averages and SDOMs of the parameter values. Results for N₀, α , and β are shown in **Table 10.1** for the clear day (20160413) and **Table 10.2** for the foggy day (20160508).

		N ₀ (cm ⁻³)	α (μm)	β						
	Mean	4.98×107	0.122	0.211						
	SDOM	2.7×10^{6}	0.003	0.004						
Table 10.1. Lognormal results 20160413 (clear day										

		N_0 (cm ⁻³)	α (µm)	β				
	Mean	1.20×10^{8}	0.1453	0.159				
	SDOM	6×106	0.0026	0.003				
Table 10.2 Logrammal results 20160508 (forget day								

Table 10.2. Lognormal results 20160508 (foggy day).

In general, the foggy day has smaller $\bar{\chi}^2$ values indicating that the PBPSO algorithm is better able to match AODs when there is a larger aerosol load in the atmosphere. **Tables 5** and **8** indicate that a bimodal distribution best describes the aerosol size distribution on the clear day (20160413) since the $\bar{\chi}^2$ values are lower in **Table 8**. **Tables 6** and **9** indicate that a lognormal distribution best describes the aerosol size distribution on the foggy day (20160508).

CONCLUSIONS

A PBPSO algorithm was used to analyze a synthetic set of atmospheric aerosol data in order to determine the reproducibility of results for four types of assumed size distributions. For Junge and lognormal distributions, there was little to no variation in the results from multiple runs of the algorithm, but the gamma and bimodal distributions the results indicated a large variability in size distribution parameters that would yield the same set of optical depths at 13 wavelengths. Running additional tests holding the α and/or β parameter(s) constant gave an indication of the sources of this variability. Inherent to the gamma distribution function, the β parameter directly relates to the rate of particle number increase as the radius increases to the peak radius $n_0 = \beta/\alpha$, while the α parameter determines the rate of particle number decrease for radii greater than the peak radius. Furthermore, the distribution is not symmetric in log space about n_0 . Hence, the retrieved size distributions are more sensitive to variations in β because the contribution to AODs from small radii particles (where Q_{ext} has a maximum) becomes highly variable leading to large variations in N_0 and α . When retrieving bimodal distributions, the relatively flat Mie extinction profiles at larger radii for visible wavelengths of light make it difficult for the algorithm to pick out α_1 , leading multiple valid solutions in which the relative contributions to the AODs from the small radius mode and the large radius mode are different. It should be noted that when α_0 , α_1 , β_0 and β_1 are all held constant, N_0 and N_1 values are exactly reproduced which is consistent with the findings in Yuan *et al.* 2011. In order to retrieve size distributions from AODs using PBPSO, multiple retrievals are performed and then a subset of solutions based on the criterion $\bar{\chi}^2 \leq 100 \bar{\chi}^2_{min}$ is chosen to find averages and SDOMs of size distribution parameters.

Generally when analyzing atmospheric aerosols, a bimodal or lognormal distribution is preferred. When using the bimodal distribution for inferred AODs from clear and foggy days, the uncertainty in the radius for the small aerosols is lower than for the large aerosols (**Tables 7.1** and **7.2**). In addition, the foggy day has smaller $\bar{\chi}^2$ values indicating that the PBPSO algorithm is better able to match AODs when there is a larger aerosol load in the atmosphere. To choose whether a bimodal or lognormal distribution best characterizes the atmospheric aerosols, both were used with the PBPSO algorithm and the distribution with the lowest $\bar{\chi}^2$ was identified. A bimodal distribution best describes the aerosols on the clear day (20160413) and a lognormal distribution best describes the aerosols on the foggy day (20160508).

ACKNOWLEDGEMENTS

The author thanks the College of Saint Benedict/Saint John's University for excellent work facilities and general support of undergraduate research.

REFERENCES

- 1. Tao, W.-K., Chen, J.-P., Li, Z., Wang, C., and Zhang, C. (2012) Impact of Aerosols on Convective Clouds and Precipitation, *Rev. Geophys.* 50, RG2001, doi:10.1029/2011RG000369.
- 2. Schwartz, J. and Neas, L. M. (2000) Fine Particles Are More Strongly Associated than Coarse Particles with Acute Respiratory Health Effects in Schoolchildren, *Epidemiology* 11(1), 6–10.
- 3. Fleagle, R. G. and Businger, J. A. (1980) An Introduction to Atmospheric Physics, 2nd ed., Academic Press, Orlando.
- 4. U.S. Department of Energy (DOE)/National Renewable Energy Laboratory (NREL)/Alliance for Sustainable Energy, LLC (ALLIANCE). "Solar Spectra: Standard Air Mass Zero." E490_00a_AM0.xls spreadsheet downloaded 9 May 2015. http://rredc.nrel.gov/solar/spectra/am0/ASTM2000.html

- 5. King, M. D., Byrne, D. M., Herman, B. M., Reagan, J. A. (1978) Aerosol size distributions obtained by inversion of spectral optical depth measurements, *J Atmos Sci* 35, 2153–2167.
- 6. Kocifaj, M. and Gueymard, C. (2012) Aerosol size distribution retrievals from sunphotometer measurements: Theoretical evaluation of errors due to circumsolar and related effects, *Atmos Env* 51, 131–139.
- 7. He, Z., Qi, H., Lew, Z., Ruan, L., Tan, H., and Luo, K. (2016) Application of the LSQR algorithm in non-parametric estimation of aerosol size distribution, *Opt Comms* 366, 154–162.
- 8. Yamamoto, G. and Tanaka, M. (1969) Determination of aerosol size distribution from spectral attenuation measurements, *App Opt* 8(2), 447–453.
- **9.** Grassl, H. (1971) Determination of aerosol size distributions from spectral attenuation measurements, *App Opt* 10(11), 2534–2538.
- 10. Wanf, Y., Fan, S., and Feng, X. (2007) Retrieval of the aerosol particle size distribution function by incorporating a priori information, *J Aer Sci* 38, 885–901.
- 11. Wang, Y. (2008) An efficient gradient method for maximum entropy regularizing retrieval of atmospheric aerosol particle size distribution function, *J Aer Sci* 39(4), 305–322.
- 12. Garey, M. R., and Johnson, D. S. (1979). Computers and intractability: A guide to the theory of NP-completeness, San Francisco, W.H. Freeman.
- 13. Eberhart, R. C. and Shi, Y. H. (2001) Particle swarm optimization: Developments, applications and resources, Proceedings of the 2001 Congress on Evolutionary Computation, Vols 1 and 2, 81–86 IEEE, New York.
- 14. Poli, R., Kennedy, J., and Blackwell, T. (2007) Particle swarm optimization: An overview, Swarm Intell 1, 33-57.
- **15.** Settles, M. (2005) An Introduction to Particle Swarm Optimization. Research Gate. Web. 25 May 2016, <u>https://www.researchgate.net/publication/242463151_An_Introduction_to_Particle_Swarm_Optimization</u>.
- **16.** Trelea, I. C., (2003) The particle swarm optimization algorithm: convergence analysis and parameter selection, *Inf Process Let* 85, 317–325.
- 17. Yuan, Y., Yi, H.-L., Shuai, Y., Wang, F.-Q., and Tan, H.-P. (2010) Inverse Problem for Particle Size Distributions of Atmospheric Aerosols Using Stochastic Particle Swarm Optimization, *JQSRT* 111(14), 2106–2114.
- Yuan, Y., Yi, H.-L., Shuai, Y., Liu, B., and Tan, H.-P. (2011) Inverse Problem for Aerosol Particle Size Distribution Using SPSO Associated with Multi-lognormal Distribution Model, *Atmos Env* 45(28), 4892–4897.
- Mao, J. and Li, J. (2015) Retrieval of Particle Size Distribution from Aerosol Optical Thickness Using an Improved Particle Swarm Optimization Algorithm, Opt Rev 22(5), 809–818.
- 20. Cimel Electronique. "Sun sky lunar multiband photometer CE318T." www.cimel.fr, downloaded 30 March 2017.
- 21. Che, H., Uchiyama, A., Yamazaki, A., Chen, H., Goloub, P., and Zhang, X. (2008) Intercomparison between aerosol optical properties by a PREDE skyradiometer and a CIMEL sunphotometer over Beijing, China, *Atmos Chem Phys* 8, 3199–3214.
- GAW Report No. 227 (2016), WMO/GAW Aerosol Measurement Procedures, Guidelines and Recommendations, 2nd Edition, World Meteorological Organization, Switzerland.
- 23. Bohren, C. F. and Huffman, D. R. (2007) Appendix A: Homogeneous Spheres in *Absorption and Scattering* of *Light by Small Particles*, Wiley-VCH Verlag GmbH.

ABOUT THE STUDENT AUTHOR

Kaitlin DuPaul graduated in May 2017 with a B.S. in Mathematics. After spending her summer researching atmospheric aerosols, she hopes that this exceptional research experience will give her an advantage as she plans to apply to graduate programs in Material Science this fall. In the long term she hopes to develop new ways to implement renewable energy resources.

PRESS SUMMARY

The parameter based particle swarm optimization (PBPSO) algorithm is introduced in order to retrieve aerosol size distributions from aerosol optical depth calculations without holding any size distribution parameters constant. PBSPO is tested against standard sets of aerosol optical depths to determine the reproducibility of results for Junge, gamma, lognormal, and bimodal size distributions. The PBPSO algorithm is applied to aerosol optical depth calculated from irradiance measurements from a Kipp-Zonen PGS-100 solar spectrometer. Results indicate that a lognormal size distribution best describes the aerosols on a foggy day (May 8, 2016) and a bimodal distribution best describes the aerosols on a clear day (April 13, 2016).