Supplement to: An empirically derived adjustable model for particle size distributions in advection fog

M. Kolářová[1], L. Lachiver[1] and A. Wilkie[1]

[1] Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

1. Developing the general model of advection fog

1.1. Raw Data Preparation

The measurement paper by Zak [Zak94] provides two datasets of measured particle sizes for each altitude, obtained using two different sensor types, each covering a different range of particle sizes. One type of sensor measured particles in the range of 2 - 47 μm, while the other covered the range of 10 - 310 μm. Each bin of the sensor was able to detect particles in a certain range, the obtained particle count was placed to the centre of the bin range for the purpose of creating a scatter plot graph. To generate a single distribution curve per altitude that covered the entire measurement range, we merged the data in the overlapping region (see Figure 1). This resulted in a slight increase in the number of particles in the tail of the distribution curve, as the particle counts obtained from the second sensor were slightly higher than those obtained from the first detector, likely due to differences in detection sensitivity.

The particle counts were significant in the region of 3.5 - 60 μm, data for particle diameters higher than 60 μm were therefore discarded. Zero point [0,0] was added to the distributions data in order to ensure the correct fitting of the curves. The observed distributions showed a bimodal pattern.

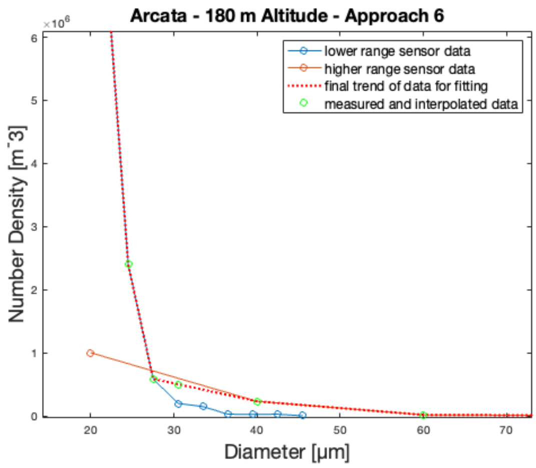

Figure 1: *An example of the distribution data collected in approach 6 where the fog was 180 m high. It demonstrates the merging of data obtained from different types of sensor. The plotted data are not fitted, the lines are included only to guide the eye. Blue colour depicts data from the first sensor that detected the lower range of diameters, while the orange colour represents data from the second sensor. The red dotted curve shows the final trend in data, while the green points represent either the measured values, or the interpolated values in the overlapping region.*

1.2. Data Fitting

Two lognormal distributions were used to fit the data as a first hint. Neverthless, due to the observed symmetry of the second peak, we opted for a normal (Gaussian) distribution in order to achieve better shape conformity. Moreover, Gaussian peak proved to be much better to use in our fitting algorithm. In contrast to the lognormal distribution, Gaussian parameters directly specify both the position and width of the peak, facilitating a more accurate initial estimate of the fitting parameters. This improved initialization of parameters effectively mitigates the risk of converging to local minima during the fitting process, consequently enhancing the efficiency of the fitting procedure. The final choice of function used for the fitting is therefore a linear combination of lognormal and Gaussian function.

Curve fitting was performed using the `OriginPro 2022b` software which uses the Levenberg-Marquardt algorithm to achieve the best fit. As a result of the fitting procedure, 6 parameters in total were obtained per each altitude of a particular dataset. The quality of the fit was relatively good, even though the data was sparse. The range of relative regression standard errors (RRSE) was from 2 to 30%. This relatively high range in fitting errors can be assigned to occasionally encountered outlier values, which were observed in bin counts due to both the detector's error ($\pm$ 1 bin) and the variable nature of fog.

All three datasets of Arcata fog exhibit similar patterns in the development of curve shape with increasing altitude. An example of the evolving curve shape for the dataset where the fog reached 260 m is shown in Figure 2. The curves at lower altitudes are characterised by a more prominent lognormal peak, which is located at lower diameter values of approximately 4 μm. The Gaussian peak becomes more prominent at higher altitudes, and its maxi-

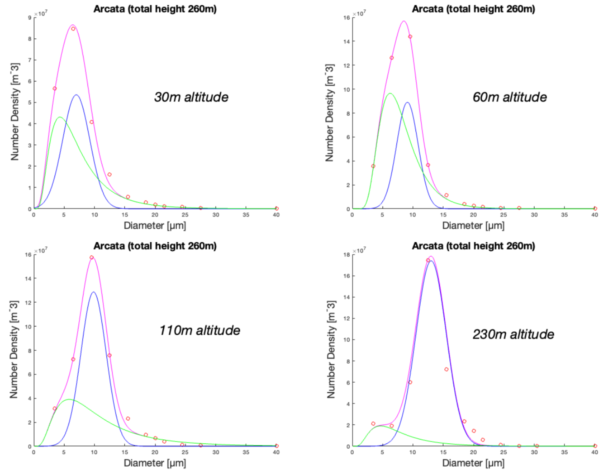

Figure 2: *Four examples of fitted curves are presented for the dataset corresponding to the fog height of 260 m, at altitudes of 30, 60, 110, and 230 m, respectively. The increase in Gaussian peak magnitude towards the lognormal peak as well as shift of its maximum towards higher diameter can be observed. Regression errors for 30, 60, 110 and 230 m were 8, 6, 13 and 22%.*

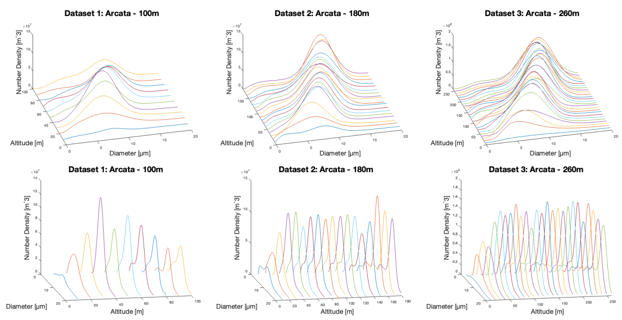

Figure 3: *3D plots of Arcata fog particle size distributions for altitudes ranging from 10 m altitude to the appropriate maximum heights of the fog. The side view shows the development of the total particle count with increasing fog height and it better depicts the bimodal character of these curves.*

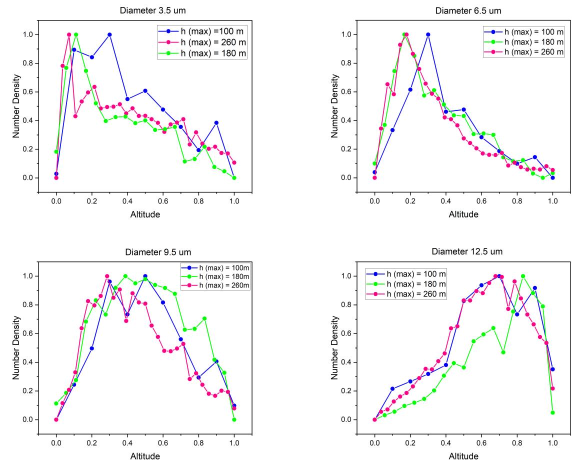

Figure 4: *Normalised plots showing the dependence of number density of a constant diameter throughout the fog layer: 3.5, 6.5, 9.5 and 12.5 μm. These graphs show similarity of Arcata fog size distribution properties in its different stages of height. The data points represent the actual raw data not the number density obtained from the fitted curve. Note that these plots correspond to cuts along the altitude axis in Figure 3.*

mum shifts towards higher diameters (from 8 to 10 μm). The total area of the fitted curve, which corresponds to the number density in each layer, rapidly increases from the lowest level of fog and then remains more or less constant until the top layer of fog. This is followed by a rapid decrease, but this data is not included in the analysis. For each dataset corresponding to different stages of Arcata fog, the 3D plots have been created in order to see the distribution shape changes with increasing altitude, see Figure 3.

In order to investigate the similarity of Arcata fog datasets the normalised plots showing number density developments were created for specific particle sizes, see Figure 4. The results indicate that Arcata fog exhibits similar distribution properties at different stages of development. Lower diameter particles tend to have an increase in concentration at low altitudes but their amount decreases afterwards. For particles of higher diameter there is an increase in their amount towards $50-60\%$ of height followed by decrease towards the top of the fog.

1.3. Curve Shape Parameters

Based on the observed similarity in fog behaviour (depicted in Figure 4), we have created plots of the curve shape parameters (σ_L, σ_G, μ_L and μ_G). In these plots values of these parameters are plotted as dependencies on normalised altitude. The purpose was to see whether they also follow similar trend so they can be subsequently fitted. By replacing the scattered data with one single function, the shape properties of the distributions can be derived for any requested fog height within the range of 100 - 260 m. For each of these parameters, data from all 3 datasets were gathered to create one plot against the normalised altitude and then `Matlab 2022b` software was used to choose the best curve fit type. An example is shown in Figure 5, which shows a fit for the μ_G parameter by a fourth-degree polynomial curve. This parameter corresponds to the maximum of the Gaussian peak, and it shows a similar distribution behaviour of Arcata fog across all of its 3 measured thicknesses.

1.4. The Proportionality of Lognormal to Gaussian peak

The two remaining parameters of the fitting process (L and G) correspond to areas under lognormal and gaussian peak. Their ratio (G:L) can be also viewed as a shape determining parameter, since it gives an information on how much of the total particle count per unit volume is distributed in each parts of the fitting function. Ratio values obtained from all three datasets were plotted as a function of the normalized altitude and fitted in similar manner as the shape

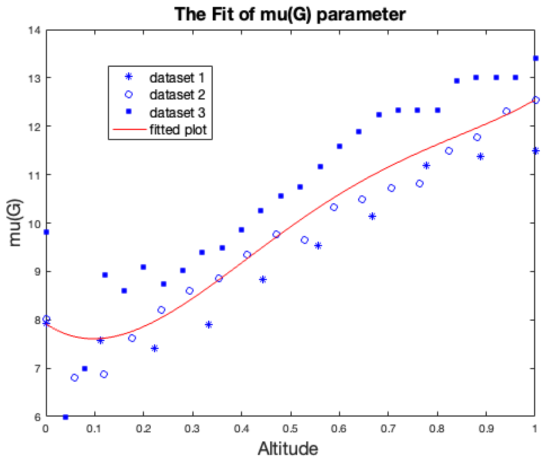

Figure 5: *The fit of evolution of parameter μ_G through the fog layer. This parameter characterises the maximum of Gaussian peak. Blue datapoints represent the gathered values from fitting all three datasets and red dashed line represents the fit of the data. In this case fourth-degree polynomial fit was chosen. Five polynomial parameters were obtained: p1 = 20.3, p2 = -50.85, p3 = 42.01, p4 = -6.82 and p5 = 7.92. Relative standard error of this plot was* 8.3%.

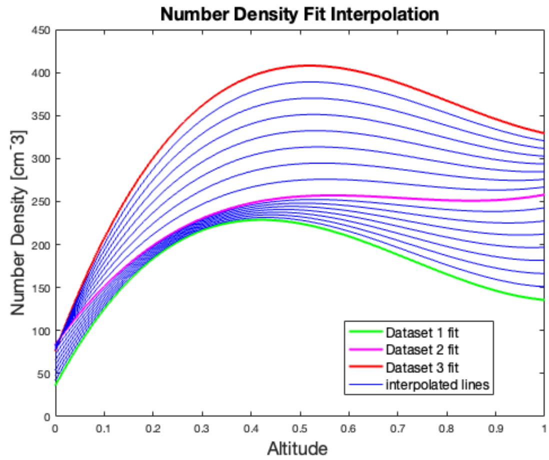

Figure 6: *Introducing the scaling factor for number density evolution due to the observation that overall number density of fog rises linearly with its height. The lower bound is derived from the 100 m dataset fit (green line), middle line corresponds to the fit of 180 m dataset (magenta line) and the upper bound arises from 260 m dataset fit (red line). Our approach allows users to input any fog height within this range and it will automatically interpolate to the appropriate curve shape. Individual example interpolations are shown as blue lines.*

parameters in order to obtain a general dependence. The values of number densities for each layer in $[cm^{-3}]$ have been provided by the authors of the source paper and are listed in [Zak94]. The determined area ratio is then used to obtain values of L and G in units of $[cm^{-3}]$ from the value of provided number density, i.e. sum of L and G must give the actually measured number density value for a particular layer.

1.5. Number Densities

When plotting the number density dependence on normalised fog height one can see the similarity in the curve shape, see Figure 6. However, the magnitude of these 3 curves is significantly different. It shows that fogs reaching higher altitudes are denser. In addition to this reasonable feature, we can see that the maximum of reached density is moving towards higher altitude for higher total height reached. In order to keep this dependence, we introduced a scaling factor that adjusts the density values based on a fog height input. The graphical representation of this interpolation can be found in Figure 6. To create lower and upper bounds of these dependencies, we used fitted curves of all three datasets, including the lower bound derived from the 100 m dataset fit, the upper bound derived from the 260 m dataset fit, and the middle range curve shape derived from the 180 m dataset fit. All three curves were fitted using a third-degree polynomial. For other fog heights within this range, we conducted interpolation either between 180 m and 260 m in the case of higher fog heights or between 100 m and 180 m in the case of lower fog heights. The model can be also to a certain degree extrapolated to heights that are above 260 m and below 100 m, but caution must be taken, whether these values make physical sense. Too high values might indicate the fog is already merging into a low cloud layer, too low values would probably correspond to other fog types, or to advection fog in the beginning of its formation, for which we do not have reliable data yet.

1.6. Extension to the Ground Level

In order to have a complete model, we include the calculation between the 10 m altitude and the ground level for number density curves. These data were not determined in the reference paper [Zak94] where all measurements end at 10 m altitude. Given the steep rise in number density from the 10 m layer upwards, linear extrapolation from 10 m to 0 m yielded negative values for the ground level. The extrapolation is simply performed by connecting the value of number density for 10 m altitude to the [0,0] point because no droplets are expected at 0 m altitude due to their dissipation by the touch with the ground. Regarding the choice of the curve shape parameters within 0-10 m range, we have chosen to keep the same values as for the appropriate generated 10 m layer.

1.7. 3-D Plot Generation

By fitting all the necessary properties, we obtained 6 sets of polynomial coefficients corresponding to 6 properties we tried to generalise (total particle count, area ratio and 4 parameters describing the curve shape). In addition to particle count polynomial parameters, we provide the scaling matrix that recalculates the overall fog density for an input fog height. The procedure of the 3D distribution generation is then quite simple. The total height of the advection fog (H) is an input parameter defined by user. Based on the value of H, the particle density curve is created, providing particle concentrations (N) for each of the appropriate altitudes. Afterwards, the value of N is split based on the area ratio dependence into particular scaling factors L and G for each desired altitude layer within defined H. The remaining four parameters, describing the shape of lognormal and Gaussian peaks, are obtained from their particular dependencies. Finally, once all parameters have been computed, the two final curves are generated and summed up for each altitude within the requested range which gives the desired particle distribution curve.

2. Reproducibility of the original data

The resulting 3D plots from our general model in comparison to the actual datasets are shown in Figure 7 for the fog height of 100 m, Figure 8 for fog height of 180 m and Figure 9 for the height of 260 m.

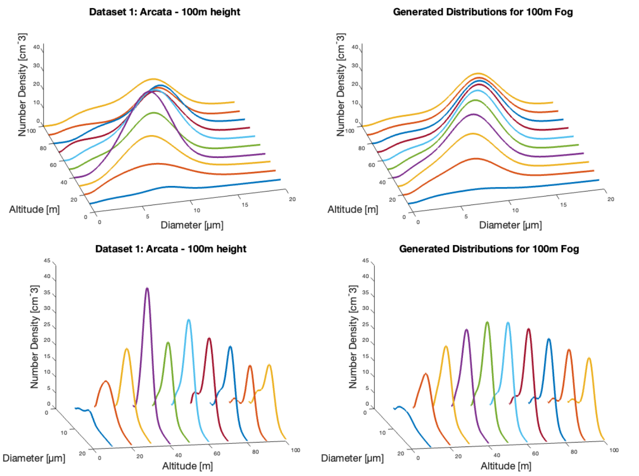

Figure 7: *Generated curves for the fog height of 100 m (right column) in comparison to the original data fit of 100 m dataset (left column). Note the caveat about the shape of the fitted distribution discussed in the caption of Figure 8.*

3. Sampling algorithm

The algorithm for our particle diameter sampling method can be seen in Figure 10. When the ray hits the fog medium, intersection position gives us a particular altitude within the fog. The input total fog height is already defined by the user and our model can give all parameters necessary to create a probability density function (pdf) for particle sizes corresponding to any altitude within the defined fog medium (see getDistributionParameters function).

Since our pdf is a sum of lognormal and Gaussian pdf, we decide which one to sample based on their appropriate magnitudes within our overall particle size pdf. A random number is chosen between 0 and 1. If the area ratio (G:L) is lower than this number, we sample Gaussian pdf. Otherwise, the lognormal pdf is sampled. In our method, the sampling is performed using normal and lognormal distributions from `std::random` library in C++.

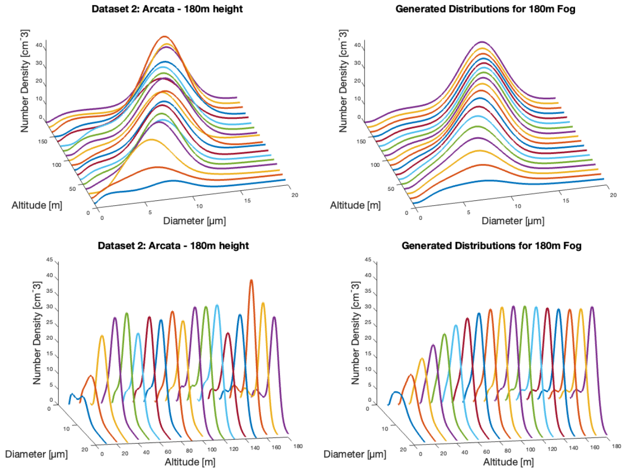

Figure 8: *Generated curves for the fog height of 180 m (right column) in comparison to the original data fit of 180 m dataset (left column). An important point to note is that in the bottom view, our fitted model looks like it increases the particle counts beyond what is found in the original data: the "ridgeline" of the plot is significantly higher. But the fit is reasonable insofar as it preserves the area under each measured distribution curve: the shape of the distribution (and in particular, its peak) is also affected by noise, which is smoothed out by the fit. So the overall transmission characteristics of the fog are still properly captured by the fit.*

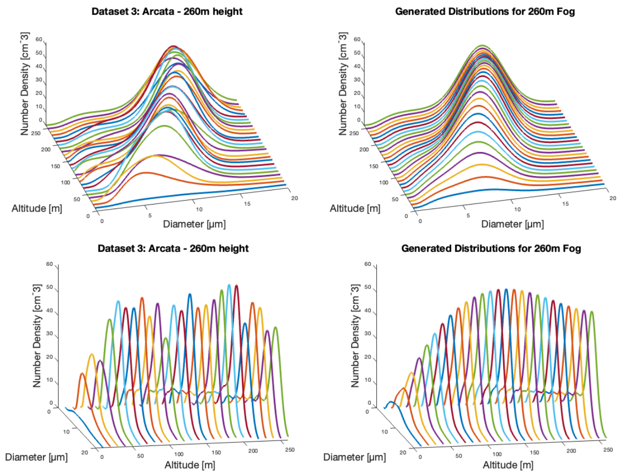

Figure 9: *Generated curves for the fog height of 260 m (right column) in comparison to the original data fit of 260 m dataset (left column). Note the caveat about the shape of the fitted distribution discussed in the caption of Figure 8.*

References

[Zak94] ZAK J. A.: Drop size distributions and related properties of fog for five locations measured from aircraft. NASA Contractor Report 4585, DOT/FAA/CT-94/02, April 1994. Prepared for Langley Research Center under Contract NAS1-19341. 1, 3

Algorithm 1 Sampling a fog particle diameter

Input: $altitude$
$distParams \leftarrow getDistributionParameters(altitude)$
$gaussDist \leftarrow gaussDist(distParams[''\mu_G''], distParams[''\sigma_G''])$
$logDist \leftarrow logDist(distParams[''\mu_L''], distParams[''\sigma_L''])$
$vGauss \leftarrow gaussDist.sample()$
$vLog \leftarrow logDist.sample()$
$dice \leftarrow random()$ ▷ Random float between 0 and 1
if $dice < distParams[''gAreaRatio'']$ **then**
    $diameter \leftarrow vGauss$
else
    $diameter \leftarrow vLog$
end if
$density \leftarrow getDensityForDiameter(diameter)$
Return $diameter, density$

Figure 10: *Our algorithm for sampling a fog particle diameter. Function $getDistributionParameters$ retrieves the curve shape parameters described in section 4.3 of the main paper. "gAreaRatio" parameter is then used when throwing the dice to decide which distribution we sample (normal or lognormal).*