

This document details some of the calculations of the items below:

$Name := ".\Data Fitting Example Data.xlsx"$

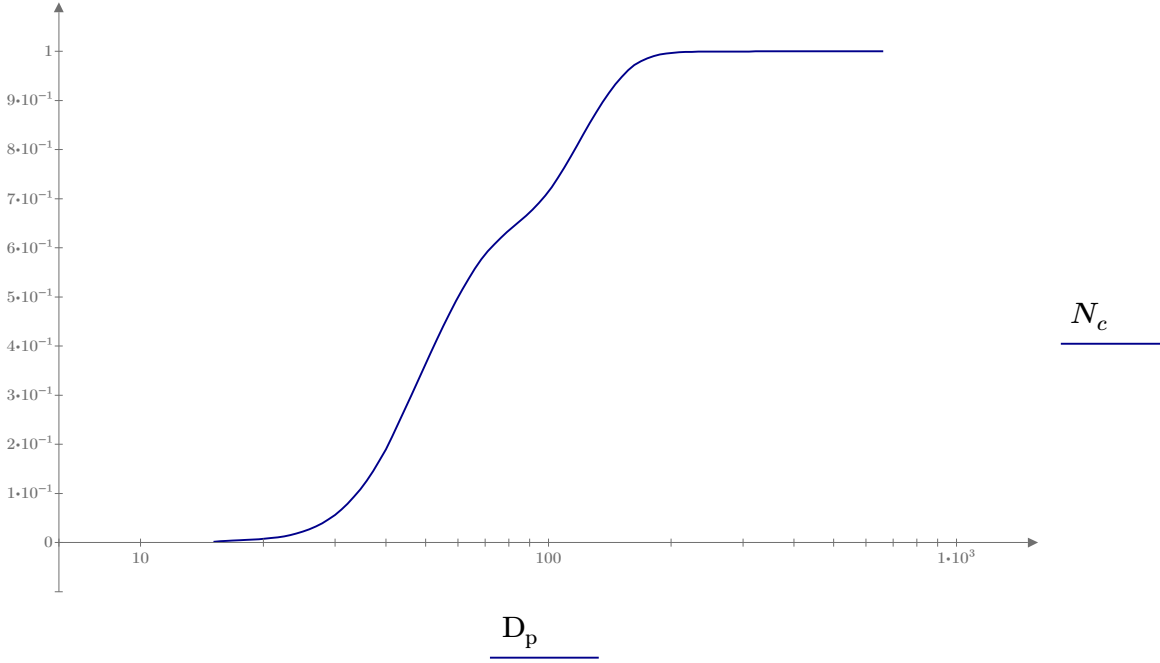
$N_c := \text{READExcel}(Name, "Sheet1!U23:U128")$

$D_p := \text{READExcel}(Name, "Sheet1!A23:A128")$

DMA Inne	0.00937			
DMA Oute	0.01961			
DMA Char	0.44369			

Outputs

$D_{pm} := \text{excel}_{"A23:A128"} \quad N_{cm} := \text{excel}_{"U23:U128"}$



$$\text{range}(D_p) := \begin{array}{|l} \parallel \\ \parallel \text{Vec} \leftarrow \frac{0 \cdot (\max(D_p) - \min(D_p))}{\text{length}(D_p)} + \min(D_p) \\ \parallel \\ \parallel \text{for } i \in 1 \dots \text{length}(D_p) \\ \parallel \parallel \\ \parallel \parallel V \leftarrow \frac{i \cdot (\max(D_p) - \min(D_p))}{\text{length}(D_p)} + \min(D_p) \\ \parallel \parallel \\ \parallel \parallel \text{Vec} \leftarrow \text{stack}(\text{Vec}, V) \\ \parallel \end{array}$$

Defining the B-Spline: One using the Midpoint Diameter, Other using a linear range

$b := \text{Spline2}(D_p, N_c, 3) \quad \text{DWS}(b) = 2.271 \quad b_m := \text{Spline2}(D_{pm}, N_{cm}, 3)$

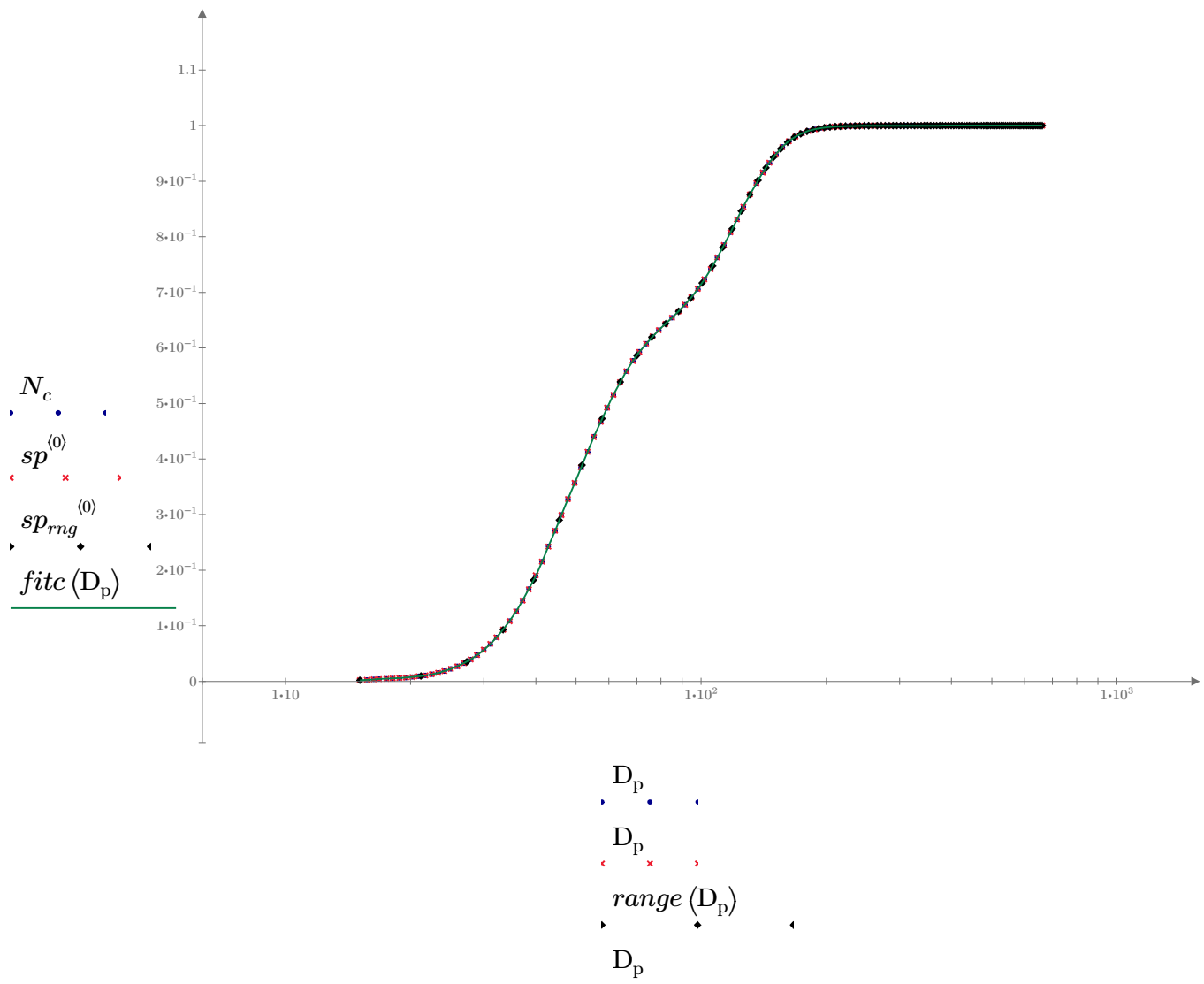
$sp := \text{Binterp}(D_p, b)^T \quad sp_{rng} := \text{Binterp}(\text{range}(D_p), b)^T \quad sp_m := \text{Binterp}(D_{pm}, b_m)^T$

Defining a cubic spline fit as below:

$c := \text{cspline}(D_p, N_c)$

$\text{fitc}(x) := \text{interp}(c, D_p, N_c, x) \quad x \text{ is the independant variable of evaluation}$

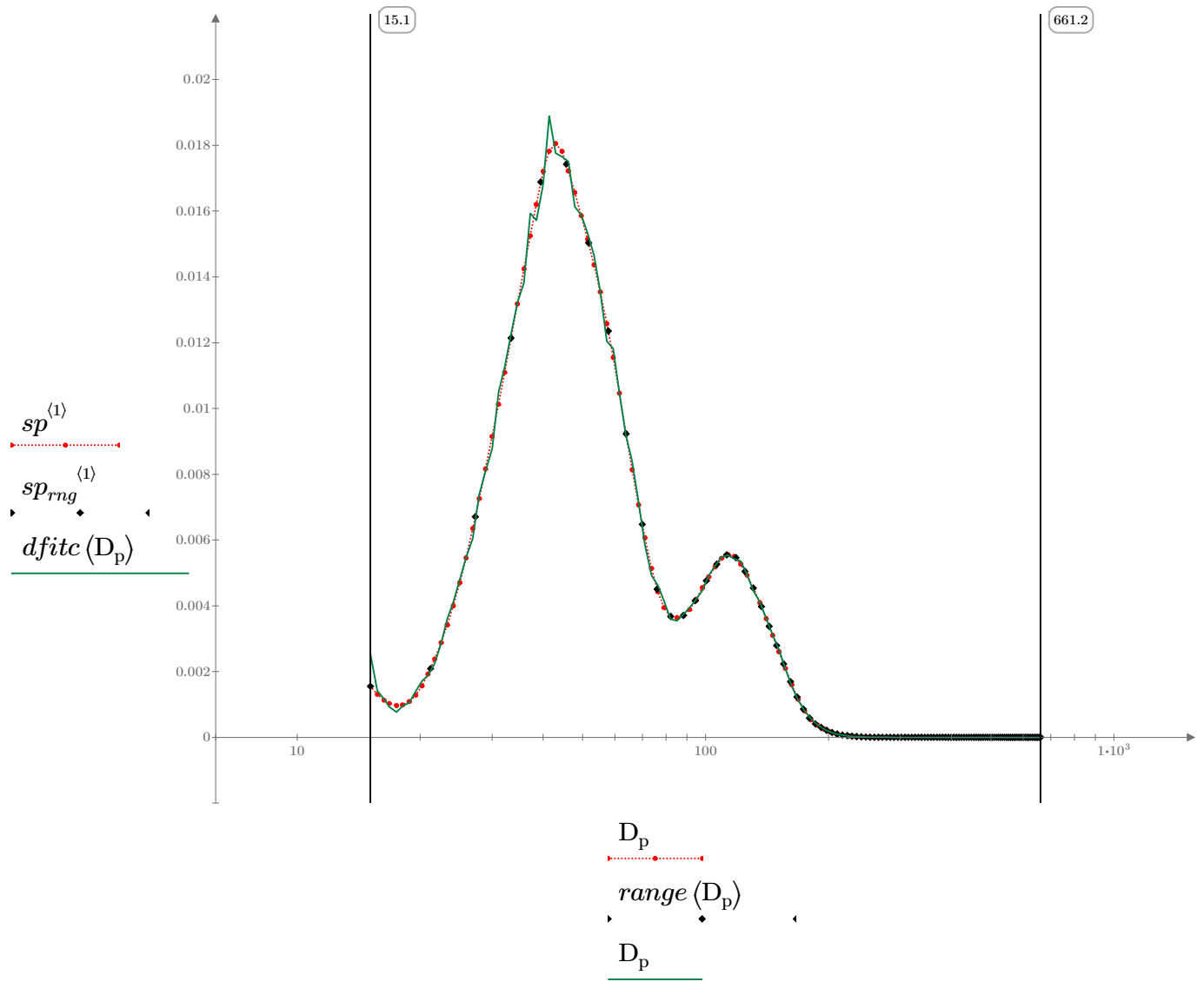
$\text{dfitc}(r) := \frac{d}{dr} \text{fitc}(r)$



As we can see, our 3 curves fit the data very nicely. The sp line has the logarithmic spacing, as given by the diameter midpoint vector of TSI's equipment. The sp_{rng} line has the linear spacing which has divisions equal to the number of elements given. The last line is a cubic line fit.

The entire reason for expressing the data this way is due to the fact that the bins are spaced logarithmically. This is due to the observed data that most of the particles occur as smaller ranges and aerosols generally tend to have a logarithmic distribution. As such however, when looking at the data, it can be misrepresented due to this bin spacing. The example is if we had a constant distribution of particles, the bins with larger bounds will pick up more particles and represent that we have more at the higher range, when in fact this is not correct. It is also not beneficial to express the values as a logarithm, as that would also skew the data as $\ln(x)$ is not a constant. As such, we are normalizing the data by removing the bin bounds by taking the derivative.

The area under the derivative curve should be equal to one, as the cumulative concentration is equal to 1.



As we can see, the B-Spline derivative is continuous through our entire range, where as the Cubic Spline is not. The B-Spline was developed to be continuous, and as such, is our best way to represent the data in a normalized manner.