Measurement methods of black carbon fractal parameters and size distribution parameters from SEM images

Sergejs Lukanihins
Oxford University

August 22, 2014

1 Introduction

The purpose of this report is to present the methods used for measuring the fractal parameter as well as size distribution parameters of black carbon fractal aggregates by analysing SEM images of fractal aggregates.

The following parameters are measured for each sample: the characteristic fractal dimension of a sample D_f, average radius of gyration of aggregates $\langle R_g \rangle$, z-average radius of gyration of aggregates $\langle R_g \rangle_z$, average number of constituent primary particles within aggregates $\langle N \rangle$, average volume of aggregates $\langle V \rangle$, the polydispersity factor C_p, as well as parameters of gamma distribution and log-normal distribution that we attempt to fit to observed distribution of N. Some of the parameters are given a definition below.

Radius of gyration as a root mean square distance from the centre of mass.

z-average radius of gyration is the "average" radius of gyration that is naturally measured in a dynamic light scattering experiment [Sorensen, 2001]. Sorensen [2001] defines z-average radius of gyration as an average weighted by the square of mass, which we can express as follows:

$$\langle R_g \rangle_z = \frac{\sum N^2 R_g(N)}{\sum N^2}, \tag{1}$$

where N is a number of constituent primary particles within an aggregate and $R_g(N)$ is radius of gyration of an aggregate with N constituent primary particles. The sum is over all aggregates in a sample.

Polydispersity factor is another useful parameter appearing in RDG (RayleighDebyeGans) theory of light scattering, whenever a sample has aggregates of different sizes [Sorensen, 2001]. It is defined as:

$$C_p = \frac{\langle N \rangle}{\langle N^2 \rangle} \left(\frac{\langle N^{2+\frac{2}{D_f}} \rangle}{\langle N^2 \rangle} \right)^{\frac{D_f}{2}}, \tag{2}$$

where $\langle N^i \rangle$ is defined as i$^{\text{th}}$ moment of distribution of N.

2 Theory underlying the fractal parameter measurements

The defining equation for fractal aggregates is:

$$N = k_f \left(\frac{R_g}{a} \right)^{D_f}, \tag{3}$$

where a is a mean radius of individual primary particles, D_f is a fractal dimension and k_f is a fractal prefactor, which is a constant of the order unity [Köylü and Faeth, 1992].

If we were able to measure the values of N and R_g of aggregates from the SEM image, we could easily calculate D_f from the linear fit of $\ln(N)$ versus $\ln(R_g)$. The difficulty lies in the fact that we only have access to 2D projections of real 3D structures. It might be possible to obtain those values by analysing stereopair images [Köylü et al., 1995], however, the individual primary particles are not distinguishable on the SEM images that were taken with significant aggregate sample size, which makes this method infeasible. Köylü et al. [1995] suggests several useful relationships that can be used to infer true 3D parameters from the parameters of 2D projected images.

First is a relationship between a projected area of the aggregate and the number of constituent particles in an aggregate [Köylü et al., 1995]:

$$N = k_a \left(\frac{A_a}{A_p} \right)^{\alpha}, \tag{4}$$

where A_a is a projected area of an aggregate, A_p is an area of a primary particle within an aggregate, α is an empirical scaling constant, and k_a is a constant of the order unity (different from k_f mentioned earlier)

Taking a logarithm of both sides of equation (4) provides a useful way for estimating parameters from a linear fit:

$$\ln(N) = \alpha \ln\left(\frac{A_a}{A_p}\right) + \ln(k_a) \tag{5}$$

Second is a relationship between 2D projected R_g and true 3D R_g, the ratio of which is suggested to be remain constant for different sizes and different fractal dimensions of aggregates [Köylü et al., 1995]:

$$\frac{R_g(2D)}{R_g(3D)} = \text{const}, \tag{6}$$

This relationship provides us with a simple way of calculating 3D radius of gyration by measuring 2D radius of gyration, which can be easily measured from the SEM image.

We can substitute equation (4) and equation (6) into equation (3) and take logarithm of both sides of equation to obtain:

$$\ln(A_a) = \frac{D_f}{\alpha}\ln(R_g(2D)) + c, \tag{7}$$

where c is constant, that has the same value for all aggregates in a sample, assuming that the primary particle radius remains constant for different aggregates in the same sample.

As long as we can determine the value of α, this provides us with a way of calculating D_f from a linear fit of equation (7), since we can rather easily measure both A_a and $R_g(2D)$ from the SEM image. Determining α is problematic; however it is possible to estimate it by analysing simulated fractal images. The details of simulation algorithm used and the results obtained from it are described in section 3.

In fact, we can use any linear lengthscale instead of $R_g(2D)$ in equation (7). We consider two important cases: the largest projected distance between two points of the aggregate F_{max} (also called the largest calliper diameter or Feret's diameter) and the geometric mean of the largest caliper diameter and the smallest calliper diameter $D_{gm} = \sqrt{F_{max}F_{min}}$ [Köylü et al., 1995].

We can rewrite equation (7):

$$\ln(A_a) = \frac{D_f}{\alpha}\ln(L) + c, \tag{8}$$

where L is some lenghtscale, with possible choices for L described above.

N is given by equation (4). However, in order to calculate N, we need to know the value of k_a. It is possible to estimate the value of k_a by analysing simulated fractal images. The details of simulation algorithm and the results obtained are given in section 3. Once we have calculated N, we can calculate any moment of N distribution we want, as well as the polydispersity factor C_p.

3 Analysis of simulated fractals

As established in section 2, we need to determine the best suited lengthscale for D_f calculation, the ratio of $R_g(3D)$ to $R_g(2D)$ and the possible values of α and k_a. In order to do that, we analyse computer generated images of fractal aggregates.

The algorithm used to generate the fractal aggregates is the Tunable Cluster-Cluster Aggregation Algorithm (CCA), which follows the description outlined in Filippov et al. [2000]. The algorithm takes the values of D_f, k_f, a and N as an input, and generates a 2D projected image of fractal (with an optional output of a true value of 3D R_g). An example of fractals generated by this algorithm can be seen in figure 1.

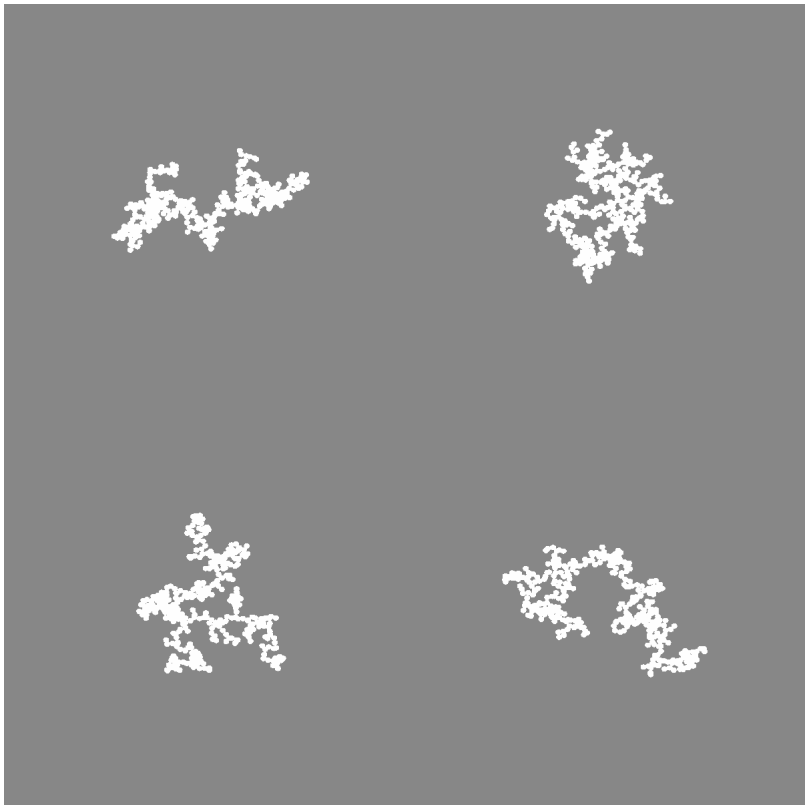

Figure 1: Fractal aggregates generated using CCA algorithm. The input parameters are: $D_f = 1.8$, $k_f = 1.2$, $N = 400$, $a = 4$ pixels

3.1 D_f model

In order to test the best suited lengthscale for D_f calculation, a significant sample of 1800 fractals was generated. Within this sample, N varied from 40 to 1240 with increment 50 (25 different values), D_f varied from 1.5 to 2.3 with increment 0.1(9 different values), and then this whole process was repeated 8 times (where the values of N and D_f were the same, but the generated fractal images were different due to the random nature of the algorithm). The value of k_f was kept constant and equal to 1.2. Since D_f is determined from a linear fit to equation (8), we can measure one value of D_f from a set of 25 fractals with different N values; so in total we get a sample of 72 independent measured D_f values.

In order to calculate D_f, we use a linear fit of $\ln(N)$ versus $\ln(R_g)$, as follows from equation (3). We know true values of N, since they were supplied as an input to algorithm.

In the figure 2 we can see true D_f values (which were supplied as an input to the algorithm) plotted against D_f values that were calculated from the fractal images, using one of the three lengthscale models. The data points are represented by the average measured D_f with standard deviation error bars, where the averaging was done over the D_f values obtained from 8 repeated samples with the same value of true D_f. The linear fit to the data is also present. We can see that the model using geometric mean of minimum and maximum calliper diameters provides the best prediction. The fact that geometric mean provides best prediction also agrees with Köylü et al. [1995].

However, this model seems to slightly underpredict D_f values. Therefore we introduce a correction to the measured values of D_f by dividing them by the correction coefficient $D_{f(corr)} \approx 0.962$.

Finally, we estimate the contribution of the model uncertainty to the error of measured D_f. This is done by calculating the standard deviation of measured D_f errors, where the standard deviation is taken over the entire sample of measured D_f values:

$$D_f(\text{model error}) = \text{stdev}\left(1 - \frac{D_f(measured)}{D_f(true)}\right) \tag{9}$$

We get $D_f(\text{model error}) \approx 2.535\%$

Another way of looking at D_f model error would be to calculate the standard deviation of $D_{f(corr)}$ for all values of D_f, which is equivalent to the calculation we used above.

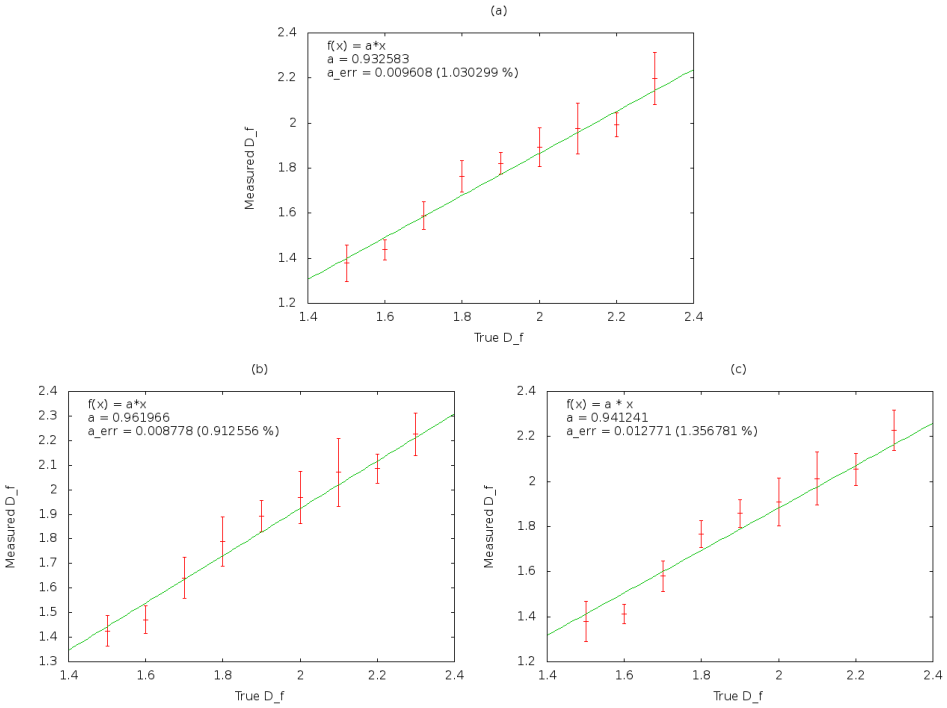

Figure 2: A relationship between true values of D_f and values of D_f measured using three different lengthscale models. (a) Model using maximum calliper diameter as a lengthscale. (b) Model using a geometric mean of minimum and maximum calliper diameter as a lengthscale. (c) Model using 2D projected radius of gyration as a lengthscale

3.2 R_g ratio

We pursue two goals when analysing the gyration radii of simulated fractals: calculate the ratio of true 3D R_g to projected 2D R_g, and verify that this ratio remains constant for different values of D_f, true R_g and N. For this purpose, we have generated 105 images of fractals, with D_f ranging from 1.7 to 2.1 with increment 0.1, N ranging from 100 to 1100 with increment 50. The algorithm also outputs the true values for 3D R_g.

The results are presented in figure 3, with 2D R_g plotted against true 3D R_g and against N with all 105 data points, and then 2D R_g plotted against D_f, where the data points for the same value of D_f are represented by the average of all the data points with the same value of D_f with standard deviation error bars. It's evident that there is no noticeable relationship between 2D R_g and either one of D_f, N or

3D R_g.

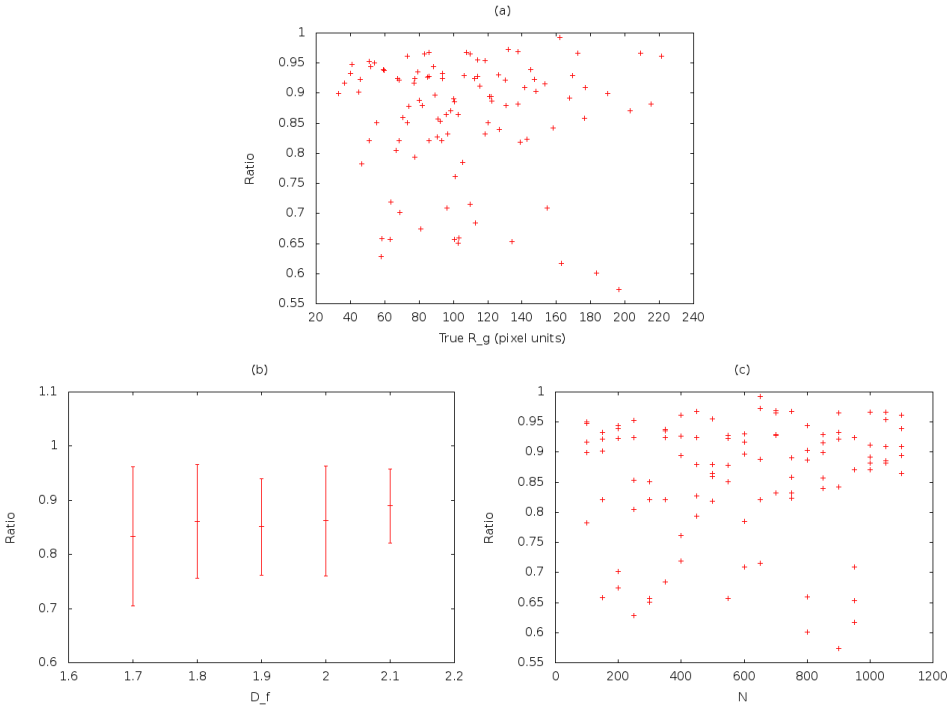

Figure 3: Ratio of 2D radius of gyration to true 3D radius of gyration as a function of : (a) true 3D radius of gyration, (b) number of constituent primary particles, (c) fractal dimension.

Finally, we calculate the average ratio of 2D R_g to 3D R_g, as well as the standard deivation of this ratio, in order to estimate the error we are going to get on the measured 3D R_g. We get ratio $r = (0.86 \pm 0.10)$.

3.3 α estimate

For the analysis of possible α values, we use the same set of generated fractals as in subsection 3.1. α was calculated as a slope from a linear fit of equation (5). For similar reasons as described in subsection 3.1, we get a sample of 72 values of α from this set.

In figure 4 below, we can see the relationship between true D_f and calculated α, where the data points for the same value of D_f are averaged over the values obtained from 8 repeated samples with standard deviation error bars. Unfortunately, the value of alpha does not stay constant, and has a weak dependence on

D_f, which does not seem to be strictly linear. The best we can do here is choose a reasonable range of D_f values, that will cover the values of D_f we expect to get from the SEM images, and then calculate mean and standard deviation of alpha for this range. Since the literature commonly reports values of D_F close to 1.8 [Sorensen, 2001], and motivated by our own D_f measurements from SEM images, we choose a range of D_f from 1.6 to 2.2. In this case, we get $\alpha = (1.069 \pm 0.47)$. This value of α is in agreement with value reported by Köylü et al. [1995]. If we happened to measure a value of D_f that fell outside of this range, we would have to re-evaluate the value of α and its error.

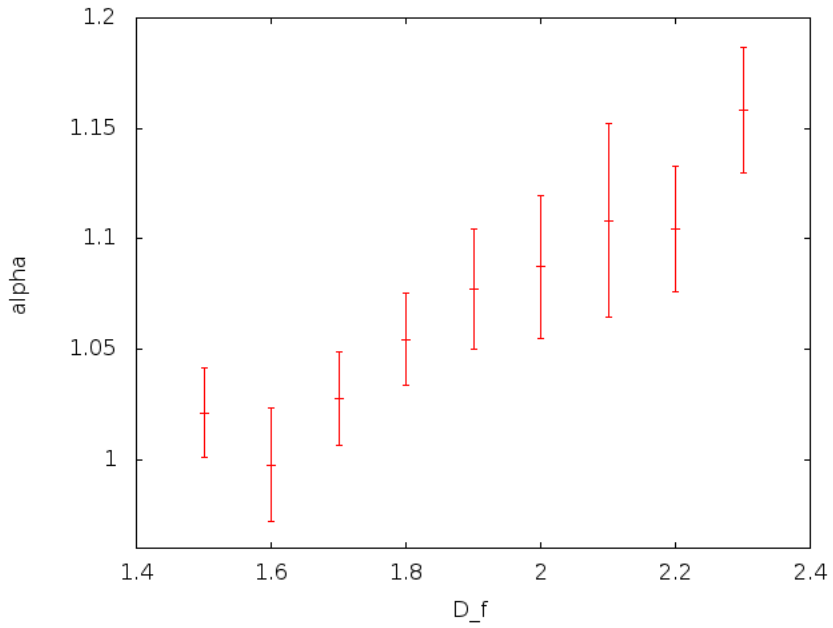

Figure 4: α plotted as a function of D_f

3.4 k_a estimate

In order to estimate value of k_a, we have generated a sample 3528 fractals. Within this sample, N ranged from 40 to 1000 with increment 20 (49 different values), k_f ranged from 1.0 to 1.8 with increment 0.1, and then the whole process was repeated 8 times(with the same parameters, but different randomized images). $\ln(k_a)$ was calculated as an intercept from the linear fit of equation (5), and then k_a was calculating by exponentiating $\ln(k_a)$. As a result, we get a sample of 72 different values of k_a (one result per 49 fractals with different values of N).

In the figure 5, the measured values of k_a are plotted against the true values of k_f. It remains inconclusive whether there actually is any relationship between those values. However, for our purpose, we should be fine taking the average value of values of k_a from 72 measurements and use this value in our future measurements of SEM images. We can also estimate an error on k_a as a standard deviation of all 72 measurements. The result we get is $k_a = (1.18 \pm 0.14)$. This result agrees with value reported by Köylü et al. [1995].

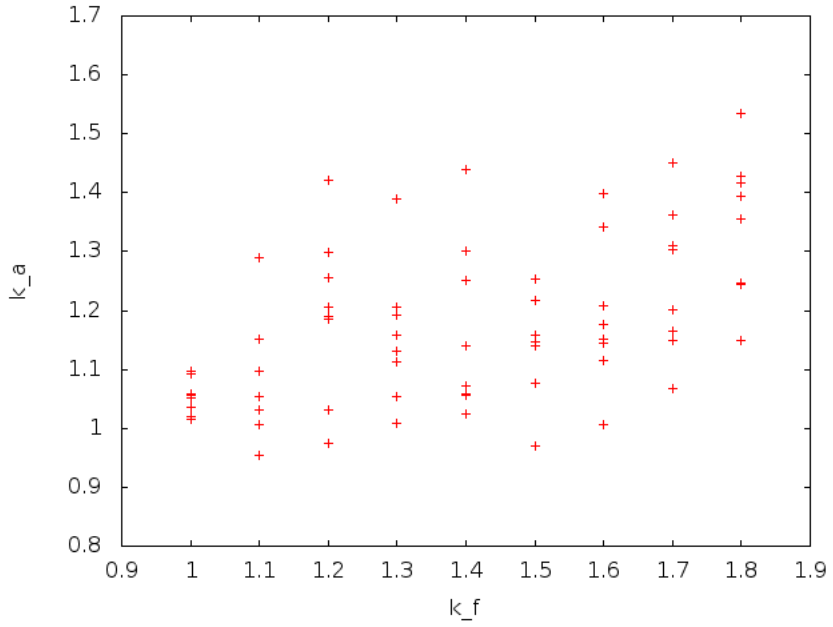

Figure 5: Area prefactor k_a plotted as a function of fractal prefactor k_f

4 Measurements

Two different SEM were used for taking images of BC samples. First microscope is capable of taking many images over a large area, and then montage them together in one big image, therefore creating an image containing a statistically significant number of aggregates. An example of image taken by this microscope can be seen in figure 6. However, the resolution it offers is not enough to be able to distinguish individual primary particles within the aggregates, and therefore not enough to measure primary particle radii. For that purpose, we used a different microscope and took images of several aggregates at high magnification. Unfortunately, the

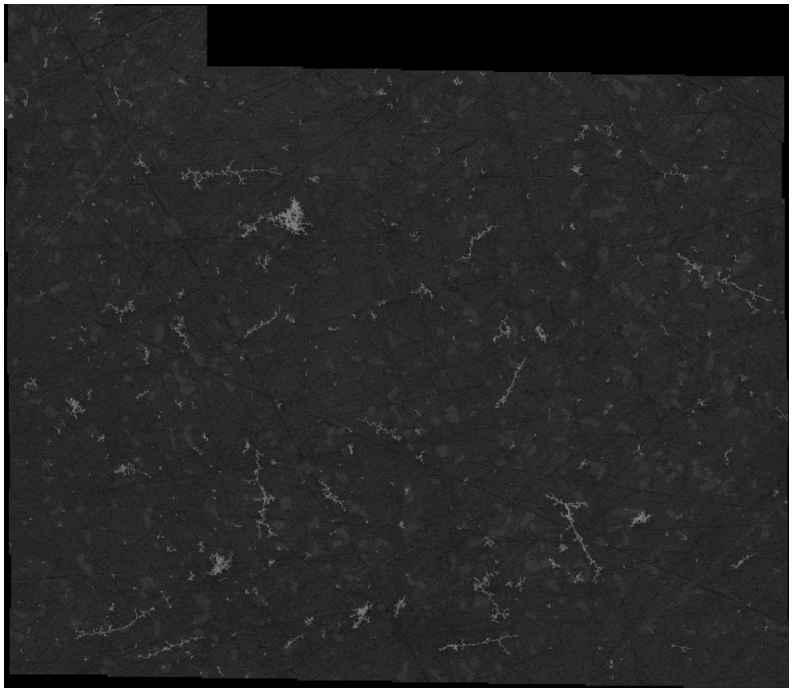

Figure 6: Large montaged image, taken using first SEM. There are about 600 aggregates in total in this image, including the small aggregates.

data gathered from second microscope is lacking, since we only took images of three aggregates for just one of the samples (sample PBCS25). An example of images taken by second microscope can be seen in figure 7.

First we analyse images obtained from the first SEM. For each aggregate, we measure directly a 2D projected area, maximum and minimum Feret's diameter (i.e. calliper diameter) and 2D projected radius of gyration.

Then we analyse images obtained from the second SEM to determine the primary particle radius. Due to lack of additional data, we are going to assume that primary particle radius will be constant for all other aggregates and all other samples, and equal to the radius calculated using the three SEM images we do have. We obtain primary particle radius $a = (0.0265 \pm 0.0028)\,\mu\mathrm{m}$.

5 Parameter calculation

This section provides a step-by-step process used to calculate all the parameters of interest.

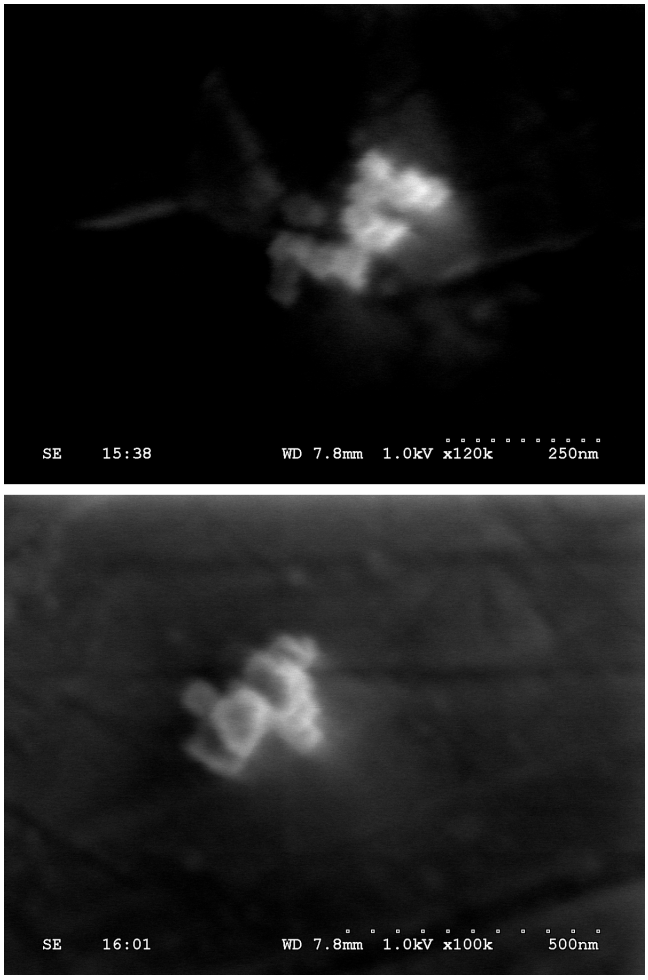

Figure 7: Images of individual aggregates taken by second SEM at high magnification. The primary particles within the aggregates can be barely distinguished.

The table below summarizes the values and standard deviation errors of all parameters we have calculated in the course of this report so far. These are the values we are going to use for analysis of every sample.

Parameter	Value	Stdev
α	1.069	0.046
k_a	1.176	0.137
r	0.859	0.100
a	0.0265 μm	0.0028 μm
$D_{f(corr)}$	0.962	0.025

To calculate D_f, we modify equation (7) by introducing a correction coefficient $D_{f(corr)}$ determined in section 3.1.

$$\ln(A_a) = D_f \frac{D_{f(corr)}}{\alpha} \ln(R_g(2D)) + c, \tag{10}$$

D_f is then calculated from a linear fit to equation (10)

The error on D_f comes from three independent sources: the error associated with model uncertainty (which can be seen as the error on $D_{f(corr)}$) , α error and the error on D_f from the linear fit. The total D_f error is then calculating by adding these three errors in quadrature

We can measure directly both a projected area for each aggregate in a sample and a characteristic primary particle radius. Therefore we can calculate N for each aggregate using equation (4). The calculation of any moment of distribution of N then follows trivially, once we know the values of N.

Once we know the values of N, C_p is calculated using equation (2).

2D R_g is measured directly for each fractal from the SEM image. 3D R_g is then estimated using the ratio $r = \frac{R_g(2D)}{R_g(3D)}$ calculated in section 3.2. We can then easily calculate $\langle R_g \rangle$

Once we know N and R_g values, $\langle R_g \rangle_z$ can be calculated using equation (1). However, in equation (1), $\langle R_g \rangle_z$ depends both on R_g and N, which increases the uncertainty of the result. We use the defining fractal relation, equation (3), to rewrite $\langle R_g \rangle_z$ in terms of R_g and D_f only:

$$\langle R_g \rangle_z = \frac{\sum R_g^{2D_f+1}}{\sum R_g^{2D_f}} = \frac{\langle R_g^{2D_f+1} \rangle}{\langle R_g^{2D_f} \rangle} \tag{11}$$

With the assumption of constant primary particle radius, the volume V is calculated by simply multiplying N by the volume of the primary particle.

$$V = \frac{4}{3}\pi a^3 N \tag{12}$$

5.1 Gamma and log-normal distributions

We use the method of maximum likelihood to estimate the parameters of both log-normal and gamma distributions.

A random variable N is distributed log-normally if its logarithm $\ln(N)$ is distributed normally. The probability distribution function (pdf) for N is then:

$$\mathrm{f}(N;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma N}\exp\left(-\frac{(\ln(N)-\mu)^2}{2\sigma^2}\right) \tag{13}$$

The maximum likelihood estimates for μ and σ are trivial, and are given by : $\mu = \langle\ln(N)\rangle$ and $\sigma = \mathrm{stdev}(\ln(N))$.

The pdf for gamma distribution can be expressed as :

$$\mathrm{f}(N;\lambda,\tau) = \frac{\lambda^\tau}{\Gamma(\tau)}N^{\tau-1}\exp(-\lambda); \qquad \tau,\lambda > 0, \tag{14}$$

where $\Gamma(x)$ is a gamma function of x.

For gamma distribution, there are no closed form expressions for maximum likelihood parameter estimates [Choi and Wette, 1969], so the solution needs to be found numerically. In order to estimate parameters τ and λ, we follow the maximum likelihood scoring method, exactly as described in Choi and Wette [1969].

5.2 Errors

This section deals with derivation of errors for all parameters except D_f, whose errors we have already calculated in section 5.

Since many parameters we want to calculate involve relatively complex relations to the parameters with known uncertainties, we opt for a numerical analysis of errors, rather than analytical analysis. The idea is to generate a large set of random parameter values, which are generated according to the uncertainties we do know, and then estimate the errors on parameters by analysing the resulting distribution.

First we describe an algorithm used for estimating the uncertainties of N moments, as well as C_p, which is directly related to N moments.

Values of N for each aggregate depend on values of k_a, a, α; the moment $\langle N^{2+\frac{2}{D_f}}\rangle$ additionally depends on D_f. We know the errors on those parameters, which are expressed as standard deviations. Let's assume that we can describe parameters k_a, a and α as independent normally distributed random variables.

First, generate a large set of random normally distributed random values of k_a, a, α (we use a set of 10000 values) . D_f random values should be derived directly from α, since D_f and α errors are strongly correlated. Then use those random values and equation (4) to calculate a set of random N values. The results of this

step may be conveniently presented in a 2D matrix containing values of N, where different rows correspond to different sets of random k_a, a and α, and columns correspond to different aggregates. We can then create a set of random values of $\langle N \rangle$, $\langle N^2 \rangle$ and $\langle N^{2+\frac{2}{D_f}} \rangle$ from the matrix by calculating corresponding averages of each row. We should take care to use the same N value matrix for calculation of different moments, in order to preserve the correlations (e.g. higher k_a value should result in higher values for all moments). For the sake of notation, we name this set of random values of N moments, "the original random set".

There is an additional contribution to error of moments, which would be non-zero even if we knew precisely the values of k_a, a and α, which arises from the fact that we looked only at one of many possible finite sets of aggregates. We may assume this error is normally distributed, and estimate it as a standard deviation of mean. For example, for $\langle N \rangle$, this error would be equal to $\sqrt{\frac{1}{m}(\langle N^2 \rangle - \langle N \rangle^2)}$, where m is a number of aggregates. Therefore, in order to have a complete picture, we need to additionally consider values of N moments in the original random set to be normally distributed random variables.

To simplify notation, define random variables X_1, X_2 and X_3, which are normally distributed with means equal to values of $\langle N \rangle$, $\langle N^2 \rangle$ and $\langle N^{2+\frac{2}{D_f}} \rangle$ in the original random set, and standard deviations equal to standard deviations of mean of those moments. Also define a random vector $\mathbf{X} = (X_1, X_2, X_3)$.

The equation for C_p involves ratios of moments of N. If we consider X_1, X_2 and X_3 to be independent random variables, we will end up overestimating the error on C_p, because X_1, X_2 and X_3 are actually correlated (which is a correlation unrelated to correlation we considered previously, which was due to using the same generated set of N values to calculate each moment). Clearly, we would expect greater values of $\langle N^2 \rangle$ correspond to greater values of $\langle N \rangle$. To take this correlation into account, we assume that the joint distribution of X_1, X_2 and X_3 can be described with a multivariate normal distribution:

$$f(\mathbf{X}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^3 |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2} (\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \right), \qquad (15)$$

where $\boldsymbol{\Sigma}$ is a 3x3 covariance matrix of $\mathbf{X}$, and $\boldsymbol{\mu}$ is mean of $\mathbf{X}$.

The covariance of $\langle N^k \rangle$ and $\langle N^l \rangle$ can be shown to be equal to $\frac{1}{m}(\langle N^{k+l} \rangle - \langle N^k \rangle \langle N^l \rangle)$, which gives us a way of calculating the covariance matrix $\boldsymbol{\Sigma}$.

We want to generate random values of X_1, X_2 and X_3 according to this multivariate distribution. It can be done by doing a variable transformation, for which matrix $\boldsymbol{\Sigma}$ becomes diagonal. The diagonal elements will be variances for random variables, which in these coordinates are independent, and therefore it is trivial to

generate them. We then get the values of $\mathbf{X}$ components by transforming back to original coordinates.

We get a final set of random values of N moments by replacing the values in the original random set with newly generated values of X_1, X_2 and X_3. We have now taken into account every source of uncertainty. The random set of C_p values is calculated trivially from the sets of moments of N using equation (2). We should, of course, remember to preserve the correlations between moments of N.

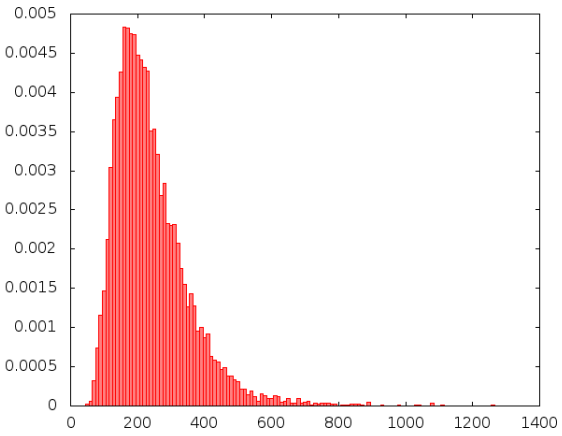

Figure 8: The distribution of $\langle N \rangle$ obtained from the numerical error analysis

Now that we have distributions of moments of N as well as C_p, we can estimate errors. An example of distribution we got for $\langle N \rangle$ can be seen in figure 8. Since distributions for several parameters can't be approximated by normal distribution (the distribution is skewed), we take a 68.26 % confidence interval in the following way : upper limit is defined by including 68.26 % of all the values that are greater than mean, and lower limit is defined by including 68.26 % of all the values that are less than mean.

Errors on the other parameters of interest ($\langle R_g \rangle$, $\langle R_g \rangle_z$, mean volume $\langle V \rangle$) are calculated numerically, using the same method as described above for moments of N and C_p.

6 Results

In this section we present the results obtained from the SEM images. The figures and graphs presented here are obtained from PBCS25 sample, but we generally get similar figures from other samples.

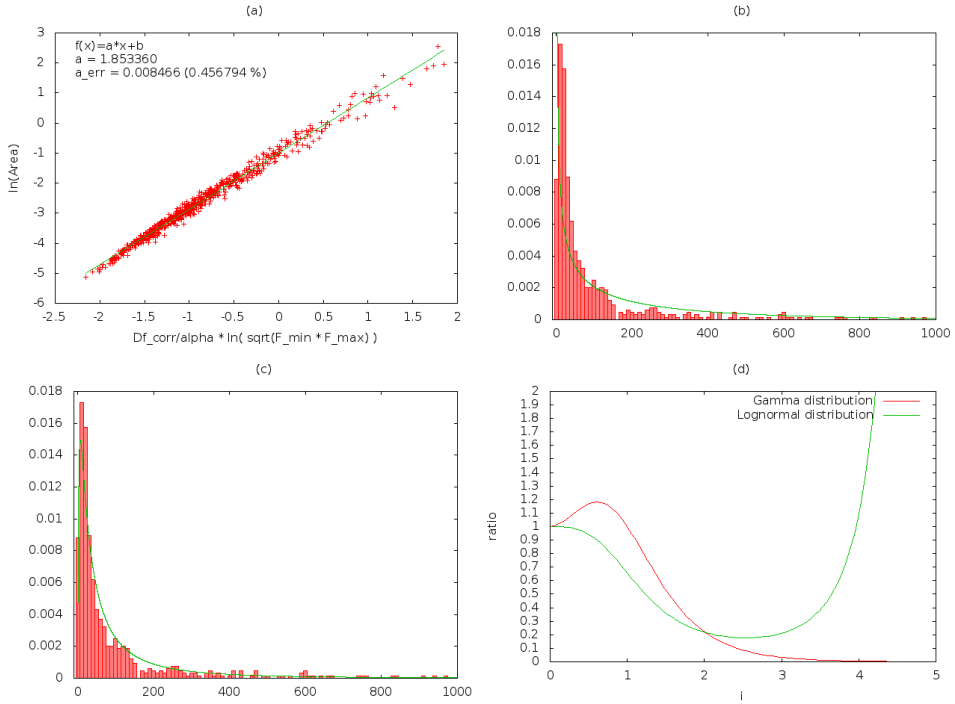

Figure 9: Results obtained from PBCS25 sample. (a) The linear relationship between $\ln(A_a)$ and $\ln(D_{gm})$, from which we derive D_f. (b) Distribution of aggregate sizes N fitted to a gamma distribution. (c) Distribution of aggregate sizes N fitted to log-normal distribution. (d) Ratio of moments calculated from log-normal and gamma distributions to moments calculated explicitly from the data as a function of power of moment i.

On figure 9(a), we have a relationship between $\ln(D_{gm})$ (geometric mean of minimum and maximum calliper diameter) and $\ln(A_a)$ (projected aggregate area) fitted to a linear function. This relationship is important for D_f calculation. We can conclude that it fits to a linear relationship reasonably well.

Figure 9(b) shows the comparison of observed distribution of $\langle N \rangle$ to a log-normal distribution that was fitted to data using the maximum likelihood method. It appears that log-normal distribution provides a reasonably decent fit to observed distribution.

Figure 9(c) shows the comparison of observed distribution of N to gamma distribution, whose parameters were also calculated using the maximum likelihood method. On visual inspection, the gamma distribution seems to be a worse fit than

log-normal distribution.

On figure 9(d) we can see a ratio of moments of distribution calculated using gamma and log-normal distributions to moments of distribution calculated explicitly from the data. It appears that neither log-normal nor gamma distribution can predict the values of the moments particularly well, except in vicinity of 1st order moment.

Finally, we present the results obtained from all measured samples in tables below.

Sample	Sample size	D_f	$\langle R_g \rangle$, µm	$\langle R_g \rangle_z$, µm	$\langle N \rangle$	C_p	$\langle V \rangle$, µm^3
PBCS13	1837	1.74	0.207	0.75	82.7	3.12	0.0064
PBCS20	96	1.82	0.636	1.59	599.8	1.63	0.0465
PBCS23	237	1.84	0.281	3.92	370.6	1.77	0.0288
PBCS25	648	1.85	0.267	3.56	209.8	2.55	0.0163
PBCS27	959	1.85	0.311	2.42	355.6	2.62	0.0276
PBCS29	1355	1.88	0.273	1.96	296.5	2.61	0.0230

Table 1: Values of fractal and size parameters for all analysed samples

	Gamma distribution		Log-normal distribution	
Sample	τ	λ	μ	σ
PBCS13	1.72	0.0208	4.10	0.74
PBCS20	0.93	0.0015	5.77	1.25
PBCS23	0.36	0.0010	4.03	1.34
PBCS25	0.46	0.0022	3.93	1.41
PBCS27	0.56	0.0016	4.75	1.20
PBCS29	0.60	0.0020	4.66	1.14

Table 2: Parameters of gamma and log-normal distributions that were fitted to observed distribution of N

Sample	$D_f \pm$	$\langle R_g \rangle +$	$\langle R_g \rangle -$	$\langle R_g \rangle_z +$	$\langle R_g \rangle_z -$	$\langle N \rangle +$	$\langle N \rangle -$	$C_p +$	$C_p -$	$\langle V \rangle +$	$\langle V \rangle -$
PBCS13	0.09	0.003	0.003	0.13	0.11	33.3	22.4	0.50	0.76	0.0026	0.0017
PBCS20	0.10	0.047	0.046	0.21	0.20	300.5	198.3	0.11	0.11	0.0231	0.0154
PBCS23	0.09	0.028	0.028	0.70	0.88	311.5	157.8	0.59	0.35	0.0241	0.0157
PBCS25	0.09	0.018	0.018	0.44	0.42	109.1	69.5	0.37	0.38	0.0084	0.0055
PBCS27	0.09	0.012	0.012	0.42	0.41	185.8	116.6	0.32	0.36	0.0146	0.0090
PBCS29	0.09	0.008	0.008	0.24	0.21	146.9	95.6	0.28	0.30	0.0113	0.0075

Table 3: Errors (68.26 % confidence interval) of fractal and size parameters for all analysed samples

7 Appendix: Descriptions of SEM images

7.1 Not analysed

PBCS10: No aggregates visible in an image, except a lot of very small ones; not feasible to analyse.
PBCS11: No aggregates visible in an image, except for very small ones (much fewer than in PBCS10); not feasible to analyse
PBCS12: No aggregates visible in an image, except for very small ones, and those are quite blurry; not feasible to analyse
PBCS14: No aggregates visible, except for few very small ones; not feasible to analyse.
PBCS31: It appears that there are no fractals in this image at all, only some dust particles.
PBCS35: An extremely large number of what appears to be a lot of completely fractured particles; not feasible to analyse.
PBCS37: This sample has a comparably small number of fractured aggregates; however all (or almost all) the aggregates were fractured, so this image was not analysed

7.2 Analysed

PBCS13: Good image, unusually very large number of aggregates. Particle selection had to be done manually, so a decent amount of smaller aggregates could have been missed.
PBCS20: It looke like several aggregates were fractured in smaller pieces. The smashed particles were mostly ingnored in analysis, and only ”good”, whole particles were analyzed.
PBCS23: Great image, no issues with analysis.
PBCS25: Great image, no issues with analysis.
PBCS27: Some particles were fractured, although much fewer than in sample PBCS20. Smashed particles were for the most part ignored.
PBCS29: Great image, no issues with analysis.

References

SC Choi and R Wette. Maximum likelihood estimation of the parameters of the gamma distribution and their bias. *Technometrics*, 11(4):683–690, 1969.

AV Filippov, My Zurita, and DE Rosner. Fractal-like aggregates: relation between morphology and physical properties. *Journal of Colloid and Interface Science*, 229(1):261–273, 2000.

ÜÖ Köylü and Gerard M Faeth. Structure of overfire soot in buoyant turbulent diffusion flames at long residence times. *Combustion and Flame*, 89(2):140–156, 1992.

ÜÖ Köylü, GM Faeth, TL Farias, and MG Carvalho. Fractal and projected structure properties of soot aggregates. *Combustion and Flame*, 100(4):621–633, 1995.

CM Sorensen. Light scattering by fractal aggregates: a review. *Aerosol Science & Technology*, 35(2):648–687, 2001.