Theory of the Stereological Analysis of Spheroid Size Distribution – Validation of the Equations

D. Gurgul, A. Burbelko *, T. Wiktor
AGH University of Science and Technology,
23 Reymonta Str., 30-059 Krakow, Poland
* Corresponding author. E-mail address: abur@agh.edu.pl

Received 14.07.2017; accepted in revised form 24.08.2017

Abstract

The paper presents validation tests for method which is used for the evaluation of the statistical distribution parameters for 3D particles’ diameters. The tested method, as source data, uses chord sets which are registered from a random cutting plane placed inside a sample space. In the sample space, there were individually generated three sets containing 3D virtual spheres. Each set had different Cumulative Distribution Function (CDF3) of the sphere diameters, namely: constant radius, normal distribution and bimodal distribution as a superposition of two normal distributions. It has been shown that having only a chord set it is possible, by using the tested method, to calculate the mean value of the outer sphere areas. For the sets of data, a chord method generates quite large errors for around 10% of the smallest nodules in the analysed population. With the increase of the nodule radii, the estimation errors decrease. The tested method may be applied to foundry issues e.g. for the estimation of gas pore sizes in castings or for the estimation of nodule graphite sizes in ductile cast iron.

Keywords: Stereological analysis, Linear analysis, Nodule size distribution

1. Introduction

The problem of the mapping of an unknown probability distribution of the spherical particle sizes in a non-transparent substance is one of the classical stereological tasks. When the density of the analysed particles is very much different from the density of the matrix phase, it is possible, for this kind of the analysis, to use the X-ray micro-tomography method [1]. There is also a possibility to evaluate the statistical distribution of the 3D nodule size by using stereological methods which base on a statistical distribution of the diameters on 2D or base on a random chord length distribution obtained by linear analysis from random secants.

2. Model assumption

Let us assume that the fraction of the visible circular sections with radius \( r \leq t \) on a random plane section is given by the function \( F_{i}(t) \) named Cumulative Distribution Function (CDF). Let us assume also that the fraction of the chords with the length \( 2r \leq 2t \) placed inside the particles is given by the function \( F_{j}(t) \) (designated as CDF).

The largest section radius, as well as the largest half-length of chord for a random round cross-section of particles are limited by the maximal radius \( R_{\text{max}} \) of the largest sphere.

Each \( F_{i}(t) \) presented above has its Probability Density Function \( f_{i}(t) \) (designated as PDF):
For this scale of the chord length proposed by Cahn and Fullmann [13], Lord and Willis [14], and distribution of spheres randomly distributed in 3D space was examination of matching the empirical function and one of the perturbations of the solution.

Quantitative metallography methods, equation (2) can be solved usually used not for the designation of the PDF3 form, but for the with respect to the integrand

\[ f_2(t) = \frac{t}{E[r_3]} \int_0^{R_{\text{max}}} \frac{f_3(x)}{\sqrt{x^2 - t^2}} \, dx \]

were \( R_{\text{max}} \) is the radius of the biggest particle in the probe, \( E[r_3] \) is the expected (mean) values of the nodule radius, \( x \) is the variable of integration.

When the empirical function \( f_2(t) \) was estimated by quantitative metallography methods, equation (2) can be solved with respect to the integrand \( f_3(x) \) by the implicit solution. In this task such a solution (named sometimes as inverse) according to [3] gives the unsatisfactory results. This is the reason why the numerical solutions of this task are used most often. The analysis of the planar section which bases on the direct Wicksell equation (2) has been used for the volume size distribution of the spheroidal particles by Scheil [4], Schwartz [5], Saltykov [6,7], Li at al. [8]. A similar solution to the mineralogy task has been presented in [9]. Unfortunately, small numerous errors of the empirically estimated function \( f_2(t) \) result in the "arbitrarily large perturbations of the solution" [10,11]. In this context, Eq. (2) is usually used not for the designation of the PDF1 form, but for the examination of matching the empirical function and one of the selected statistical distribution law, e.g. normal, log-normal, Weibull or uniform-sized [3,8,12].

Other stereological method for the calculation of the size distribution of spheres randomly distributed in 3D space was proposed by Cahn and Fullmann [13], Lord and Willis [14], and Spektor [15]. This method bases on the measurements of the PDF1 of the corresponding chord-length distribution obtained through lineal analysis:

\[ f_3(t) = \frac{N_i}{2\pi N_3} \left[ \frac{f_i(t)}{t^2} \cdot \frac{1}{t} \frac{df_i(t)}{dt} \right] \]

where \( N_3 \) is the number of the nodules and \( N_i \) is the number of the chord with the diameter (length) that does not exceed \( l_i \). The limitation of this solution is the low resolution for the large diameter of the particles which is a result of the logarithmic nature of the scale.

According to [17] the empirical CDF3 for spherical particle radii should be estimated on the basis of the empirical estimation of PDF1 of half-length chords \( t \) as follows:

\[ F_3(t) = 1 - \frac{S}{8\pi t} f_3(t) \]

where \( S \) is the mean external surface of nodules in the probe.

### 3. Data preparation

In order to test the precision of the proposed method, for the evaluation of the CDF3 form on the basis of the PDF1 given by eq. (5), three virtual sets of nodules were generated. The nodules are randomly distributed in the volume of size equal to \( 8a-b-R_{\text{max}} \) (Fig. 1). These particles are cut by a base plane on which a set of parallel secants are placed. The 2D section diameters and length of the chords, made by the secants, are measured and collected. The virtual probe space, in which the particles were generated, fulfils the following conditions:

- The mean grain density in the specimen volume is equal to \( N_1 \) (number grains per unit volume).
- The spatial distribution of the grain centres is defined by the Poisson statistical distribution.
- The probability of intersection between the particles is negligible (or for the intersected particles it is possible to indicate and measure the diameter of every nodule).

Each linear dimension which will be mentioned in the later part of the paper, is expressed by convention in a unit length (ul). Similarly, all measured and calculated areas are also expressed by convention in a surface unit (1 us = 1 ul²).

The virtual sets of nodules were generated as follows: In the first stage the set of radii \( r_{ij} \) (for \( i = 1 \ldots N \)) is generated by using a pseudo-random number generator for three types of CDF1 which will be described in the next section. In the next step among all generated particles, it is necessary to find the radius \( R_{\text{max}} \) of the greatest nodule. The last step is to generate, by using the uniform pseudo-random number generator, the centre coordinates \( (x_i, y_i, z_i) \) for all nodules. The coordinates are distributed in the following range:

\[ x \in (-a + R_{\text{max}}; a - R_{\text{max}}) \]
\[ y \in (-b + R_{\text{max}}; b - R_{\text{max}}) \]
\[ z \in (-R_{\text{max}}; R_{\text{max}}) \]

Having fulfilled the above conditions, the results of the virtual stereological measurements are registered from the rectangular test plane \( T \) shown in Fig. 1:
The measuring plane \( T \) is parallel to \( XY \) axes and cut the \( Z \) axis in zero. All secants are parallel to the \( X \) axis and have the same distance between each other equal to \( k \). The secants cut through the 2D section and form the chords (the fragments of the secant inside the nodules sections). Only these particles and sections are taken into account which the central points (\( x \) and \( y \) coordinates) lie on the \( T \) area (excluding the \( R_{\text{max}} \) margins). The aim of such an approach is to eliminate the bias error created by the overestimation of the nodules with small sizes.

Between the nodules and their sections, chords and central points, there are the following geometrical relations:

- for the nodule \( i \) with the radius \( r_{3i} \), its radius \( r_{2i} \) of the section is equal to:
  \[
  r_{2i} = \sqrt{r_{3i}^2 - z_i^2}
  \]  
  (7)

- the relation between the radius of the nodule \( i \) and its length of the chord \( (r_{i,j})_k \), the coordinates of the particle centre \( (y_j) \) and the coordinates of the measuring secant \( j \) \((y_{a,j})\) is as follows:
  \[
  r_{3i}^2 = z_i^2 + r_{i,j}^2 + (y_j - y_{a,j})^2
  \]  
  (8)

Hence the half-length of the random chord \( j \) of the nodule \( i \) will be equal to:

\[
  r_{i,j} = \sqrt{r_{3i}^2 - z_i^2 - (y_j - y_{a,j})^2}
  \]  
  (9)

4. Data analysis

In order to check the application of the equation (5) for the evaluation of the \( F_3(t) \), three sets of data have been prepared. Each set has different PDF3 of the nodules radii:

A. The set of nodules with a constant radius equal to \( r_3 = 200 \). The total Number of particles is \( N = 1024 \).

B. The normal distribution with the mean value \( \bar{r}_3 = 120 \), and the standard deviation \( \sigma = 15 \), \( N = 2048 \).

C. The bimodal distribution as the superposition of two normal distributions with the parameters \( \bar{r}_{3,1} = 90 \), \( \sigma_1 = 10 \), \( N_1 = 2048 \) and \( \bar{r}_{3,2} = 150 \), \( \sigma_2 = 15 \), \( N_1 = 1024 \).

In order to obtain the sets with the normal distribution, the inverse cumulative distribution function method was used. The data was generated by using the \texttt{vdRngGaussian} function from the MKL library [18]. Whereas for generating the nodules’ central points, the \texttt{vRngUniform} function from this library was used.

The cumulative distribution functions and probability density functions for the analysed sets of the half-length chords are calculated by empirical equations:

\[
  F_1(t_i) = \frac{N_i}{N}
  \]  
  (10)

\[
  f_1(t_i) = \frac{N_i - N_{i-1}}{N(t_i - t_{i-1})}
  \]  
  (11)

where \( N_i \) is the chord numbers with a length no longer than \( 2t_i \). For the calculation the sets of \( r_1 \) were sorted in an ascending order. In the calculations the length of every five hundred chord or every thousandth was taking into account. The results for 1 out of 500 chords are presented in the fig. 2.

If in the sample there are no spherical particles with a radius less than \( R_{\text{min}} \) then it is possible to estimate the value of the mean external surface of the nodules \( \bar{S} \) on the basis of measurements obtained for interval \((0 < t < R_{\text{min}})\) by the following equation:

\[
  S_{\text{est}} = 8\pi \frac{t_i}{f_1(t)}
  \]  
  (12)

By substituting the equation (11) into (12) we obtain:

\[
  S_{\text{est}} = 8\pi Nt_i \frac{N_i - N_{i-1}}{N_i - N_{i-1}}
  \]  
  (13)

This means that for \( t_i < R_{\text{min}} \) the estimation of the \( S_{\text{est}} \) may be obtained by equation (13) after the substitution of the data for any interval \( t \) of the chord length (provided that the measured data is free of noise). The results received by this equation are presented in fig 3. In the case of the A distribution for the whole interval of the measured chords, the calculated results are characterised by the mean value \( S_{\text{est,A}} = 502\cdot10^3 \) us, with the standard deviation \( \sigma_A = 34\cdot10^3 \) us. The results for two first intervals are clearly bigger than the rest and were eliminated from the calculation. The precise value of the outer area for the sphere with the radius \( r_3 = 200 \) ul is equal to 502 655 us.
Fig. 2. The shapes of the empirical cumulative distribution functions and probability density functions for the prepared sets of nodules: a) constant $r_3$ – A; b) normal – B; c) bimodal – C

In the case of B and C distributions, there are distinguishing horizontal intervals that indicate the absence (or negligible low number) of the nodules with a radius less than around 70-80 ul. The mean values of the $S_{est}$ and the standard deviation for the radii which are less than 70 ul are following:

- for the B distribution: the mean value $S_{est}\_B = 181.9 \cdot 10^3$ us, the standard deviation $\sigma_B = 8.6 \cdot 10^3$ us,
- for the C distribution: the mean value $S_{est}\_C = 161.7 \cdot 10^3$ us, the standard deviation $\sigma_C = 7.3 \cdot 10^3$ us.

Similarly to the case of the A distribution, the results from the three first intervals (that give results clearly bigger than the rest) were eliminated from the calculation.

Fig. 3. Cumulative distribution functions and probability density functions of the half-length of the chords for the nodular particle sets for the distributions: a) constant radius $r_3$ – A; b) normal B; c) bimodal C
For comparison, the precise values of the mean outer areas, which were calculated directly on the base of the source data, are equal to 182 406 us for the set B and 162 774 us for the set C respectively. As it follows from the above data, the results of the estimation obtained by the proposed method are similar to the real parameters of the analysed set of nodular particles.

The cumulative distribution functions of the nodular particles returned by the eq. (5) on the base of the source data for the sets B and C are presented in Fig. 4. The results of the direct calculation on the basis of the bench-mark data are presented by solid lines. The dependence obtained by the means of eq. (5) is indicated by the dots. As it follows from the pictures, the proposed method has a lower accuracy at the interval close to 10% of the smallest nodules of the analysed population. With the increase of the nodule radii, the accuracy of the CDF3 mapping by the chords method increases and for the particles with big sizes, the accuracy of mapping is reasonable.

5. Conclusions

- The empirical CDF3 of the nodular size particles can be estimated on the basis of the empirical measurement of PDF1. For the estimation of CDF3, it is necessary to use Eq (5) and the statistical distribution of random chord lengths received from experimental data.
- In the calculations, the results must be used from the interval, where the estimation of the mean value of the outer particle surface by means of the eq. (12) is constant.
- For the used sets of the source data, the proposed chord method has a lower accuracy at the interval close to 10% of the smallest nodules of the analysed population. With the increase of the nodule radii, the accuracy of the CDF3 mapping by the chord method increases and for the particles with big sizes, the accuracy of mapping is reasonable.

Acknowledgment

The authors would like to express their thanks to The National Centre for Research and Development for supporting this study through Project No. PBS3/B5/38/2015.

References


[18] Intel Math Kernel Library, Developer Reference, Revision 0.11, MKL 2017.