Volume 2 No.7, JULY 2011 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

Algorithms in "UniModBase" Information System for determine Rosin-Rammler and Gaudin-Shumann equations of particle size distribution using regression analysis

[1]SAŠA PUŠICA, [2]VELIMIR ŠĆEKIĆ, [3]MILUNKA PUŠICA
[1]M.Sc.., Agencija "Smaj Business", Crnovrških brigada 6/1, 19210 Bor, Serbia
[2]prof. D.Sc., FIM, Majke Jugovića 4, 37000 Kruševac, Serbia
[3]B.Sc., Regionalni centar za talente Bor, 3.oktobar 71, 19210 Bor, Serbia
smajsoft@yahoo.com

ABSTRACT

"UniModBase" is freeware information systems for modeling of grain size analysis and determine Rosin-Rammler and Gaudin-Shumann equations of particle size distribution based on algorithms for regression analysis. The algorithm is precisely defined procedures and instructions how to solve a task or problem, and this paper show the algorithms for determine the Rosin-Rammler and Gaudin-Schuhmann equation using regresion analysis. Author of this paper these algorithms used as starting point in the construction of "UniModBase" information system.

Keywords— *Information system, algorithm, regression analysis, Rosin-Rammler, Gaudin-Schuhmann*

1. INTRODUCTION

Grain size of sample is a mixture of grains with different shape and size in narrow size class. Narrow size class contains the grains whose particle diameter is in the range of class between d1 to d2. This range depends on the methods and devices for determine the particle size distribution [1].

Sieve analysis is one of method for determine the composition of the raw material, where this material is classified according to geometric size of grains. Sieving through the sieve, the sample is classified into two class size. Large classes with pieces of grain that is greater than the holes of sieve, is retained on the sieve. Small class with grains smaller than the holes of sieve, passing through the sieve. Sieving of sample on multiple sieve from a series (Fig.1) obtained narrow size class of sample.

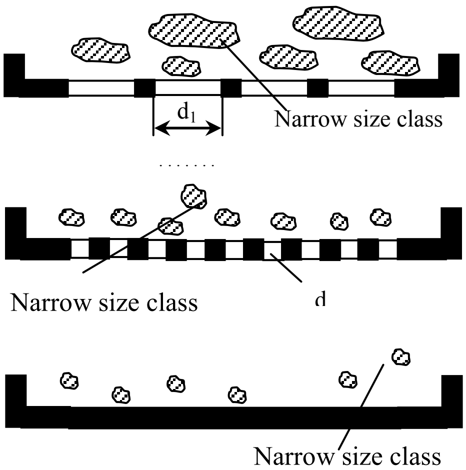


Table 1: Tabular presentation of grain size analyse

Narrow size class [mm]	Partial content m [g]	Partial content W [%]	Cumulative content R [%]	Cumulative content D [%]
1	2	3	4	5
+ 3.327	140.3	7.11	7.11	100.00
- 3.327 + 2.362	105.8	5.36	12.47	92.89
- 2.362 + 1.651	179.1	9.08	21.55	87.53
- 1.651 + 1.168	256.7	13.01	34.56	78.45
- 1.168 + 0.833	223.6	11.33	45.89	65.44
- 0.833 + 0.589	160.0	8.11	54.00	54.11
- 0.589 + 0.417	230.1	11.66	65.66	46.00
- 0.417 + 0.295	122.5	6.21	71.87	34.34
- 0.295 + 0.208	154.8	7.84	79.71	28.13
- 0.208 + 0.147	109.0	5.52	85.23	20.29
- 0.147 + 0.104	69.4	3.52	88.75	14.77
- 0.104 + 0.074	71.3	3.61	92.36	11.25
- 0.074 + 0	150.8	7.64	100.00	7.64
TOTAL	1973.4	100.00		

2. GRAIN SIZE ANALYSE PRESENTATION

There are three type for grain size presentation: tabular, graphical, analytical.

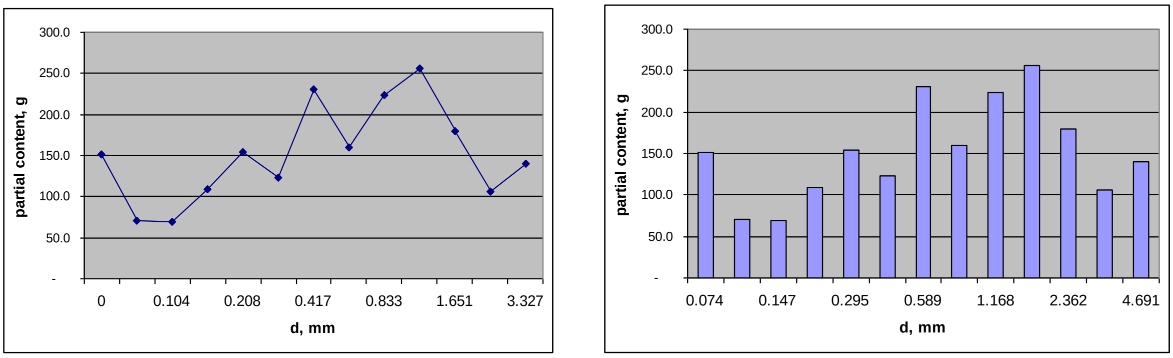
Principle for tabular presentation is based on a combination of direct results of grain size analysis (input data) and calculated data [1]. Input data are the diameter of holes of sieve d [mm] and mass of sample m [g] which is retained on the sieve. Other data are obtained by computation with the mass m of the sample. Example for

Volume 2 No.7, JULY 2011 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

tabular presentation is showed in Table 1. Calculated data in this table refer to the mass of size class and they are: W - mass of size class [%], R - cumulative content of mass retained on sieve [%], D - cumulative content of mass that passed through the sieve [%]

All type of graphical presentation for grain size analyse is based on data obtained from the tabular presentation (Table 1). Graphical presentation can be constructed based on partial (Fig.2) or cumulative characteristics (Fig.3) of grain size from Table 1.

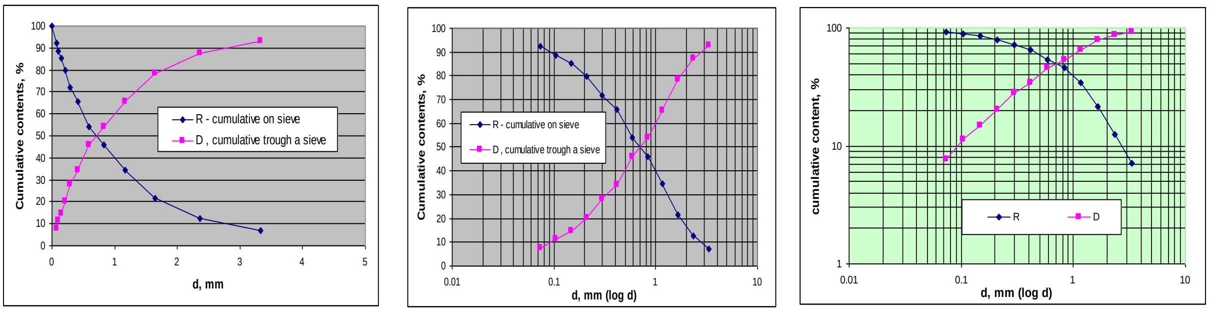


a) Partial characteristics

b) Histogram of partial characteristics

Figure 2 - Grain size analyse - parial graph

Graph of cumulative characteristics is plotted with narrow size class on the abscissa and the cumulative participation of individual class fractions (R or D) on the ordinate. If both characteristics (R and D) are shown in a graphic (Fig.3a), ordinates of their intersection is always 50%. For the sample with a wide range of grain size, drawing graphics is on semi logarithmic (Fig.3b) or logarithmic graph (Fig.3c) of cumulative characteristics R and D [2,3].



a) Cumulative characteristics

b) Semi logarithmic graph

c) Logarithmic graph

Figure 3 - Grain size analyse - cumulative graph

2.1 ANALYTICAL PRESENTATION FOR GRAIN SIZE ANALYSE

Many authors have proposed several different mathematical models and most famouse are Rosin-Rammler (Rosin and Rammler, 1933) and Gaudin-Schuhmann equation (Gates- Gaudin-Schuhmann, Gates 1915, Schuhmann 1940) [1,2].

2.1.1 Rosin-Rammler equation

Rosin-Rammler equation is the analytical presentation of the cumulative size distribution R in the form of exponential function:

$$R = 100 \ . \ e^{-\left(\frac{d}{d'}\right)^n} \qquad (1)$$

R - cumulative content of mass retained on sieve [%]
e - base of natural logarithms, e = 2.718
d - diameter of holes of sieve [mm]
d' , n - parameter
Rosin-Rammler equation can be reduced

$$\frac{100}{R} = e^{\left(\frac{d}{d'}\right)^n} \qquad (2)$$

$$\log \ \log \frac{100}{R} = n \ . \ \log d - n \ . \ \log d' + \log \ \log e$$

(3)

In equation (3) by replacing the following expression

$$B = \log \ \log e - n \ . \ \log d' \qquad (4)$$

Volume 2 No.7, JULY 2011 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

We gets the Rosin-Rammler equation of straight line in the coordinate system of log d, log log shown in Figure 4.

With graphical method it is possible to determine the value of n , d' parameters (Fig.5) where the choice of representative points for line p is subjective choice. This is the main reason for define algorithm to determine Rosin-Rammler equation using regression analysis - precision to avoid subjective errors with graphical method.

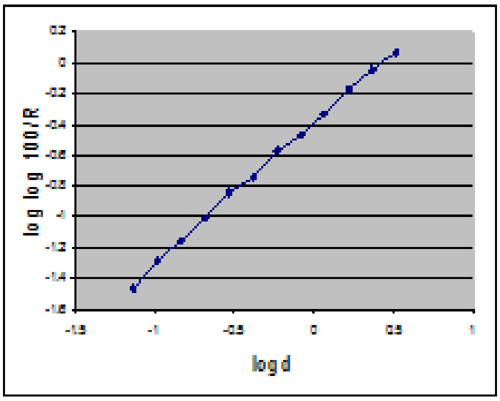


Figure 4 - Example of graph in a coordinate

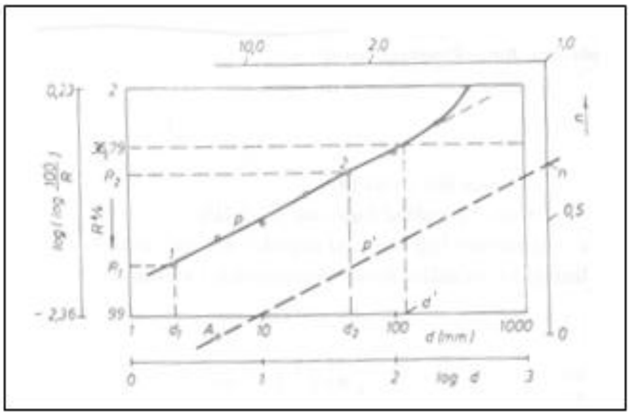


Figure 5 - Graphic principle for determining n, d' parameters of Rosin-Rammler equation [Error! Reference source not found.]

2.1.2 Gaudin-Schuhmann equation

Gaudin-Schuhmann equation is the analytical presentation of cumulative size distribution:

$$D = 100.\left(\frac{d}{d_{\max}}\right)^m \qquad (5)$$

D - cumulative content of mass that passed through the sieve [%]
d - diameter of holes of sieve [mm]
dmax - diameter of holes of sieve that passed 100% of sample [mm]
Gaudin-Schuhmann equations can be reduced

$$\log\ D = 2 + m\ .\ \log d - m\ .\ \log d_{\max} \qquad (6)$$

In equation (6) by replacing the following expression

$$A = 2 - m\ .\ \log d_{\max} \qquad (7)$$

We get the Gaudin-Schuhmann equation of straight line in the coordinate system of log d, log D shown in Figure 6. The value of m parameters (Fig.7) is possible to determine with graphical method where the choice of representative points and angle for line is subjective choice. To avoid subjective errors with graphical method, we can define algorithm to determine Gaudin-Schuhmann equation using regression analysis.

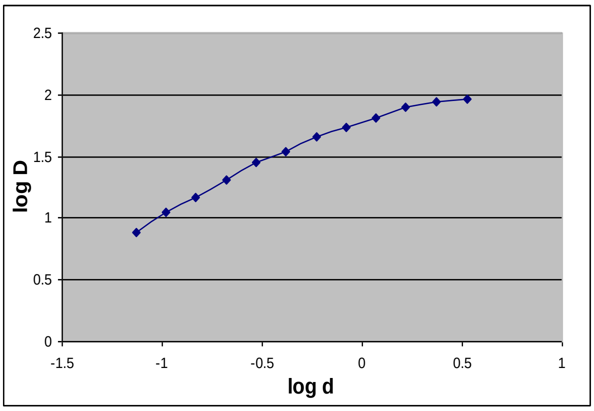


Figure 6 - Example of graph in a coordinate system log d , log D

3. REGRESSION ANALYSIS

3.1 Linear regression

A linear relationship y = ax + b is the simplest mathematical model of the functional dependence of related variables. The parameter a is called the regression coefficient, while the parameter b represents the value of a regression function when the independent variable equals zero [4,6,9]. To determine this parameters, it is assumed that the values of variable yi associated with the values of the independent variable xi as follows:

$$y_i = a.x_i + b, \quad i = 1,\dots, n \qquad (8)$$

- x1, x2, ... xn - measured values of independent variables
- y1, y2, ... yn - measured values of dependent variables

Numerical values of the dependent variable is obtained when the equation (8) incorporate the measured values of independent variables:

$$\hat{y}_i = a.x_i + b, \quad i = 1,\dots, n \qquad \textbf{(9)}$$

Volume 2 No.7, JULY 2011 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

Differences between measured and theoretical values of the dependent variable represent the deviation of this mathematical model:

$$\varepsilon_i = y_i - \hat{y}_i = y_i - (a.x_i + b), \quad i = 1,\ldots, n \qquad (10)$$

For a minor deviation, the calculated values of the dependent variables are very close to measured values and the proposed mathematical model adequately demonstrates the measured data.

According to this principle, model y = ax + b is the most accurate when the sum of squared deviations is minimal:

$$f(a,b) = \sum_{i=1}^{n}[y_i - (a.x_i + b)]^2 = \sum_{i=1}^{n}\varepsilon_i^2 \qquad (11)$$

The sum of squared deviations is minimal if the conditions are met:

$$\frac{\partial f(a,b)}{\partial a} = 0 \qquad \frac{\partial f(a,b)}{\partial b} = 0 \qquad (12)$$

We get the system of two equations with two unknowns a, b parameters for mathematical model:

$$\sum_{i=1}^{n} y_i x_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i = 0 \qquad (13)$$

$$\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i - nb = 0 \qquad (14)$$

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad (15)$$

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad (16)$$

These equations can be simplified by introducing the following expressions:

$$S_x = \sum_{i=1}^{n} x_i \qquad S_y = \sum_{i=1}^{n} y_i \qquad S_{xx} = \sum_{i=1}^{n} x_i^2 \qquad (17)$$

$$S_{xy} = \sum_{i=1}^{n} x_i . y_i \qquad DE = n \;.\; S_{xx} - S_x^2 \qquad (18)$$

The parameters a, b of mathematical model is now calculated:

$$a = \frac{n.S_{xy} - S_x.S_y}{DE} \qquad (19)$$

$$b = \frac{S_{xx}.S_y - S_x.S_{xy}}{DE} \qquad (20)$$

In accordance with these principles of linear regression, is defined and on Figure 8 is presented the algorithm for determining the coefficients a, b of linear regression (KOEF_A, KOEF_B) to be applied in determining the Rosin-Rammler and Gaudin-Schuhmann equation.

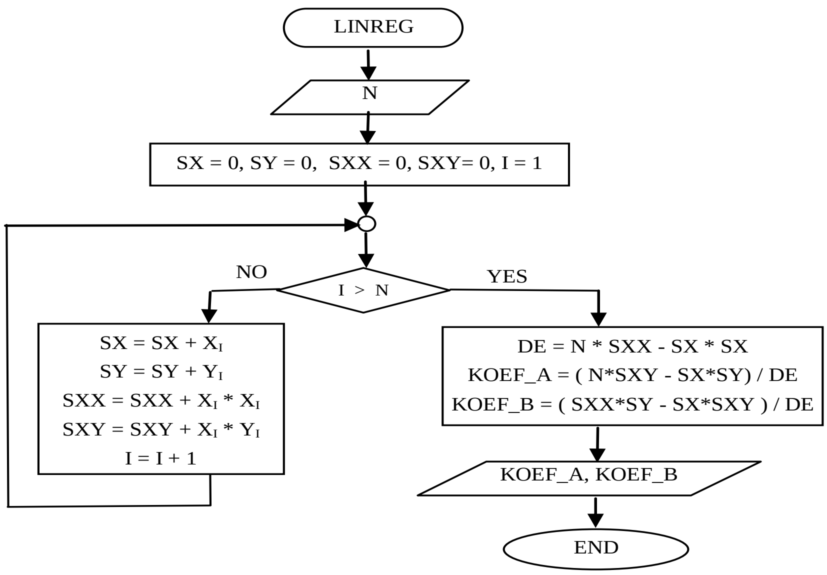


Figure 8 - The algorithm for determining the coefficients of linear regression

3.2 The coefficient of determination and correlation coefficient

The coefficient of determination is the ratio of the total and interpreted deviation [7,10] and value of this coefficient indicates the representativeness of the model: model is more representative if the coefficient of determination close to 1.

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}, 0 \le R^2 \le 1 \qquad (21)$$

The correlation coefficient r measures the linear correlation of paired measurement data (x1, y1), ..., (xn, yn) and this values are always in the range from -1 to 1. The value 1 or -1 indicates a perfect correlation and value 0 show that these are independent variables. On Figure 9 is presented the algorithm for determining the coefficient of determination and correlation coefficient (DETER, KOREL) to be applied for the Rosin-Rammler and Gaudin-Schuhmann equation obtained by linear regression.

$$r = \sqrt{R^2} \qquad (22)$$

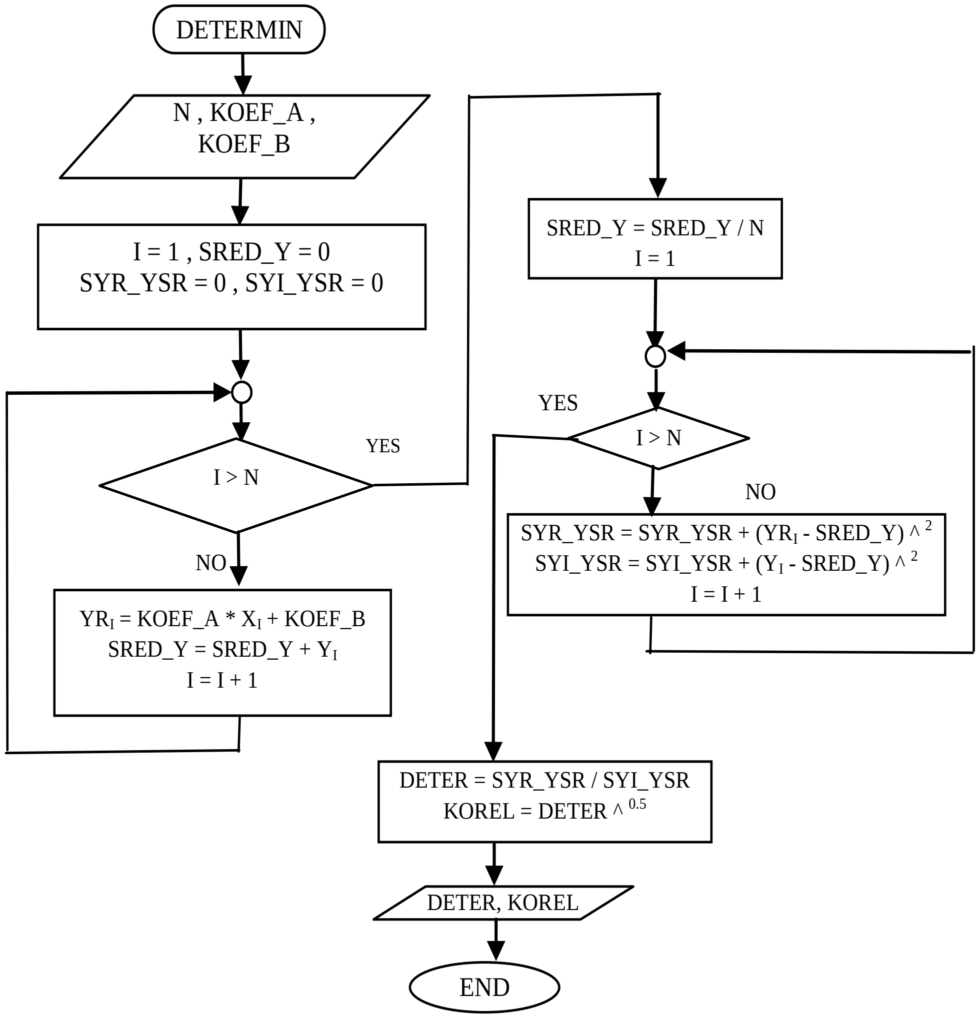


Figure 9 - The algorithm for determining the coefficient of determination and correlation coefficient

3.3 Standard deviation, variance, coefficient of variation for regression

Standard deviation of regression, shows the average deviation of measured values of the dependent variable from its regression value, expressed in the same units of measurement as the dependent variable [11,12] and calculated as the square root of variance for regression:

$$\sigma_{\hat{y}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}\right)^2 \qquad (23)$$

Standard deviation of regression:

$$\sigma_{\hat{y}} = \sqrt{\sigma_{\hat{y}}^2} \qquad (24)$$

Coefficient of variation for regression is unnamed number that is expressed in percentages, based on the standard deviations of regression and the mean value of the dependent variable.

$$V_{\hat{y}} = \frac{\sigma_{\hat{y}}}{\bar{y}}.100, \quad [\%] \qquad (25)$$

In accordance with this, is defined and on Figure 10 is showed the algorithm for determining the standard deviation, variance and coefficient of variation for regression (RDEVIA, RVARIA, KVARIA) to be applied for the Rosin-Rammler equation obtained by linear regression.

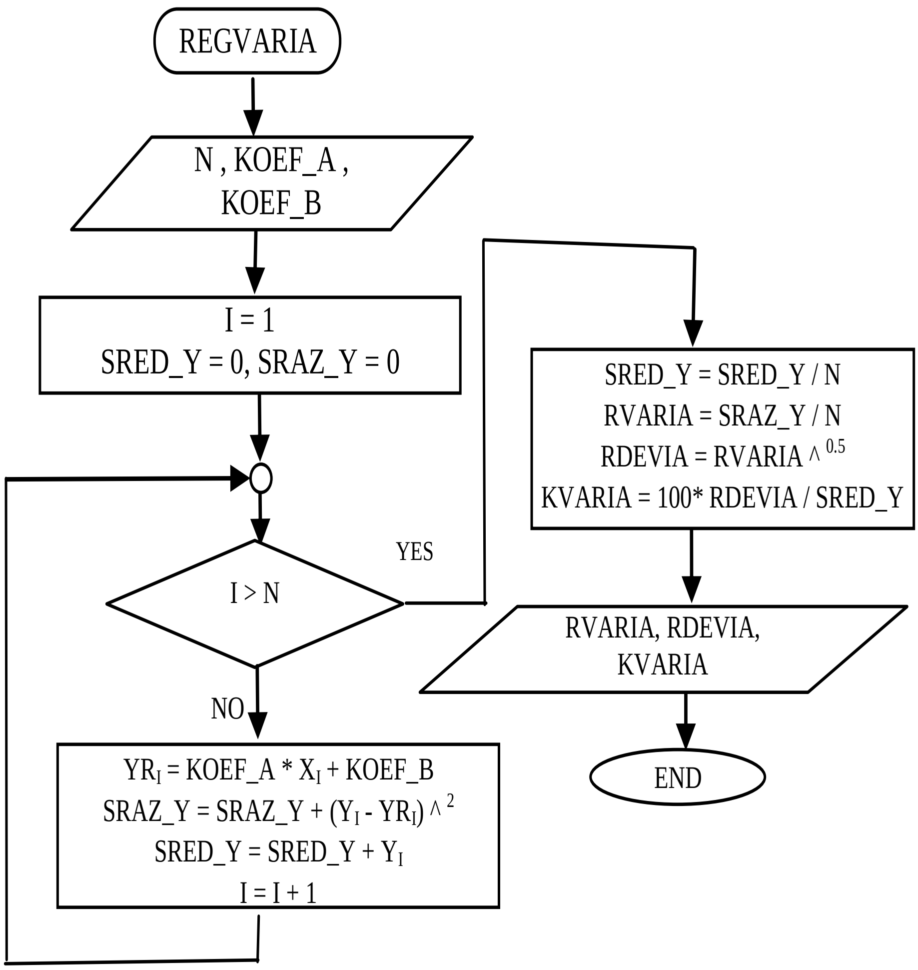


Figure 10 - The algorithm for standard deviation, variance and coefficient of variation for regression

4. ALGORITHM FOR ROSIN-RAMMLER LINEAR REGRESSION

On Figures 11, 12 and 13 is presents the algorithm for determining the Rosin-Rammler equation using linear regression. Firs step is to determine masses of samples Wi in percentages which is calculated on masses mi of size class. For this, is must calculated the total weight of sample mSUMA as the sum of all partial class fractions mi and then is performs the recalculation of the masses of samples Wi in percentages. With this data, Ri is calculated as the sum of the mass of samples Wi that would be retained on the sieve diR.

Using converted original data, is calculated the coefficients A, B for linear regression and values D_PRIM, N_KOEF who representing the parameters d', n of Rosin-Rammler equation obtained by linear regression method. After that is calculated the coefficient of determination and correlation coefficient for this model (equation), as well as standard deviation, variance and coefficient of variation for regression.

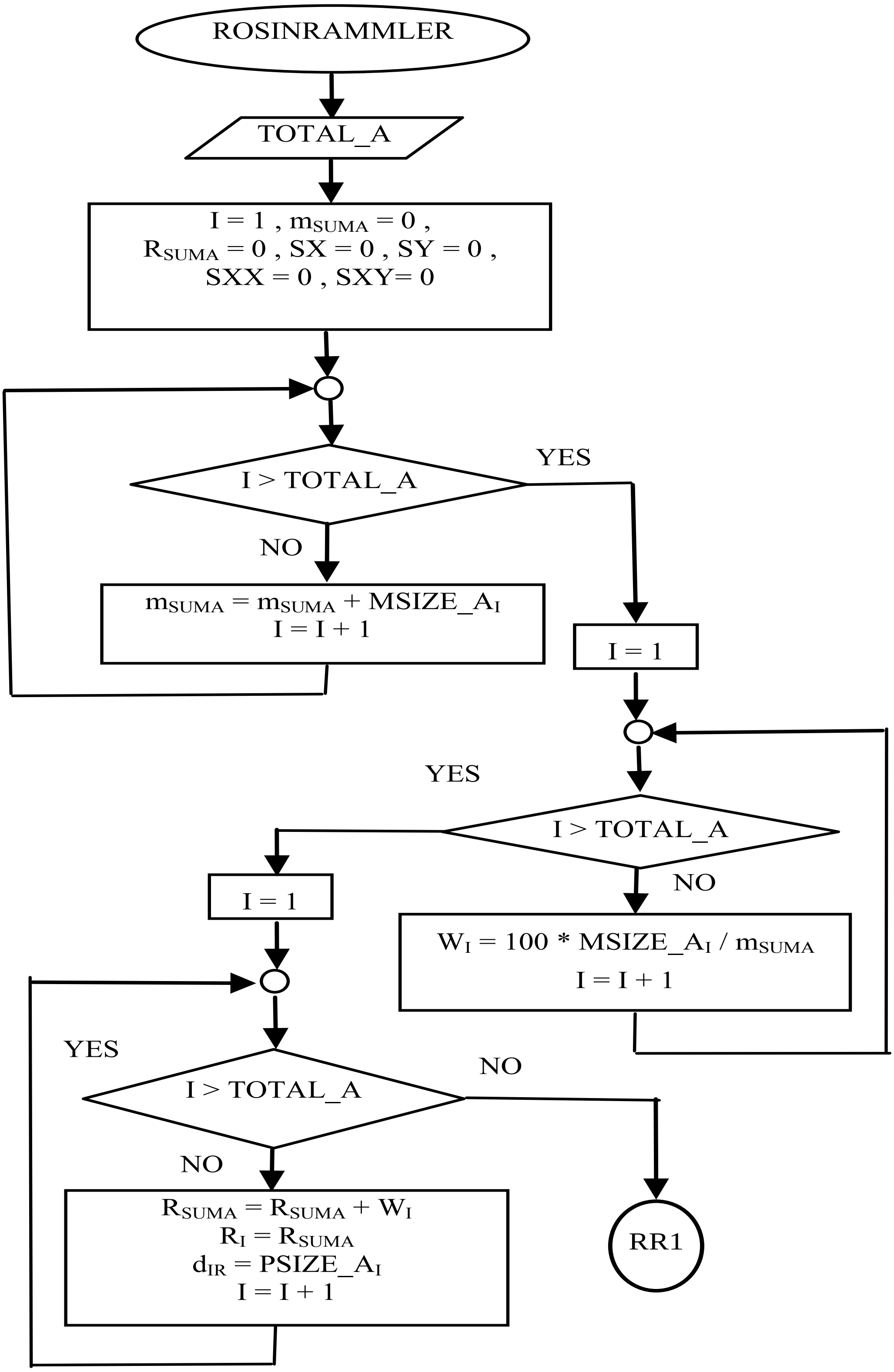


Figure 14 - The algorithm for Gaudin-Schuhmann equation and statistical analysis of this model

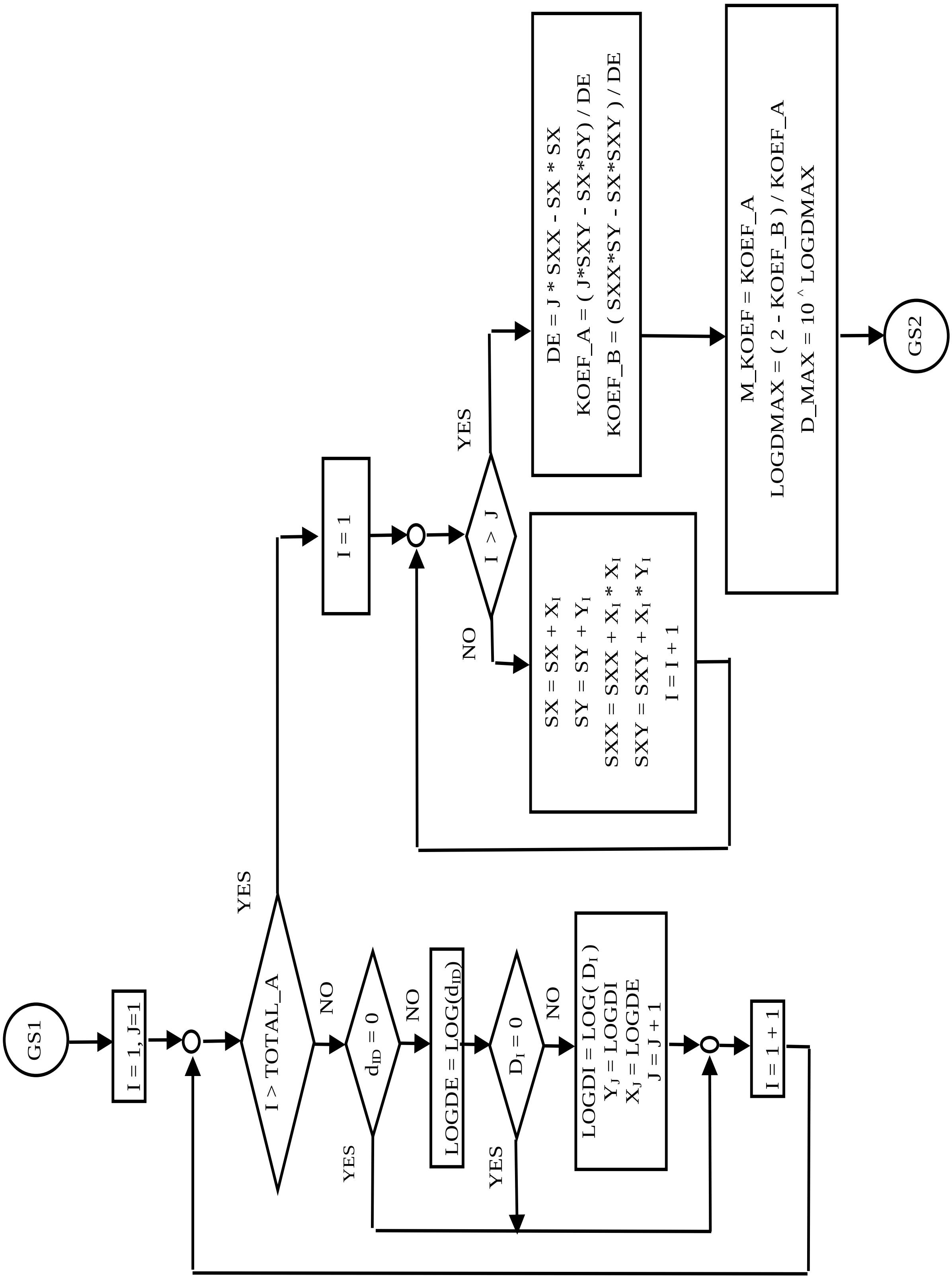


Figure 15 - The algorithm for Gaudin-Schuhmann equation and statistical analysis of this model

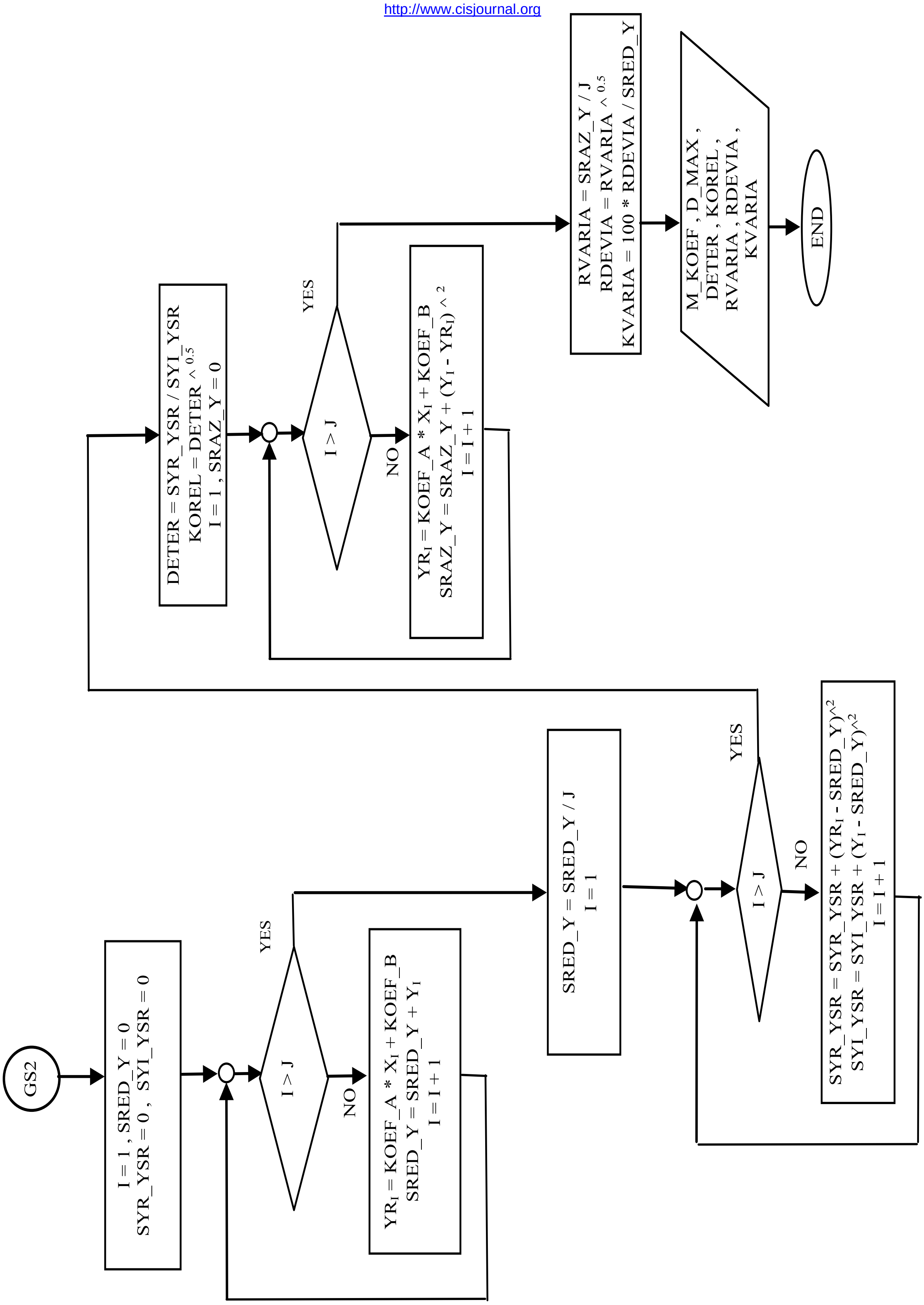


Figure 16 - The algorithm for Gaudin-Schuhmann equation and statistical analysis of this model

Volume 2 No.7, JULY 2011 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

CONCLUSION

This paper show the algorithms for determine the Rosin-Rammler and Gaudin-Schuhmann equation of particle size distribution using regresion analysis that author of this paper used as starting point in the construction of "UniModBase" information system [13, 14, 15].

In this information system (Fig.17 - 21) the statistical analysis and modeling for results have been simplified for using and "hiden" from users. With entered data, preparation of all required reports is automated and always available for future analysis and theoretical considerations.

On this way is solving the practical problems of analytical presentation of grain size analysis and remove the subjective error in determining the parameters of this model using graphical method.

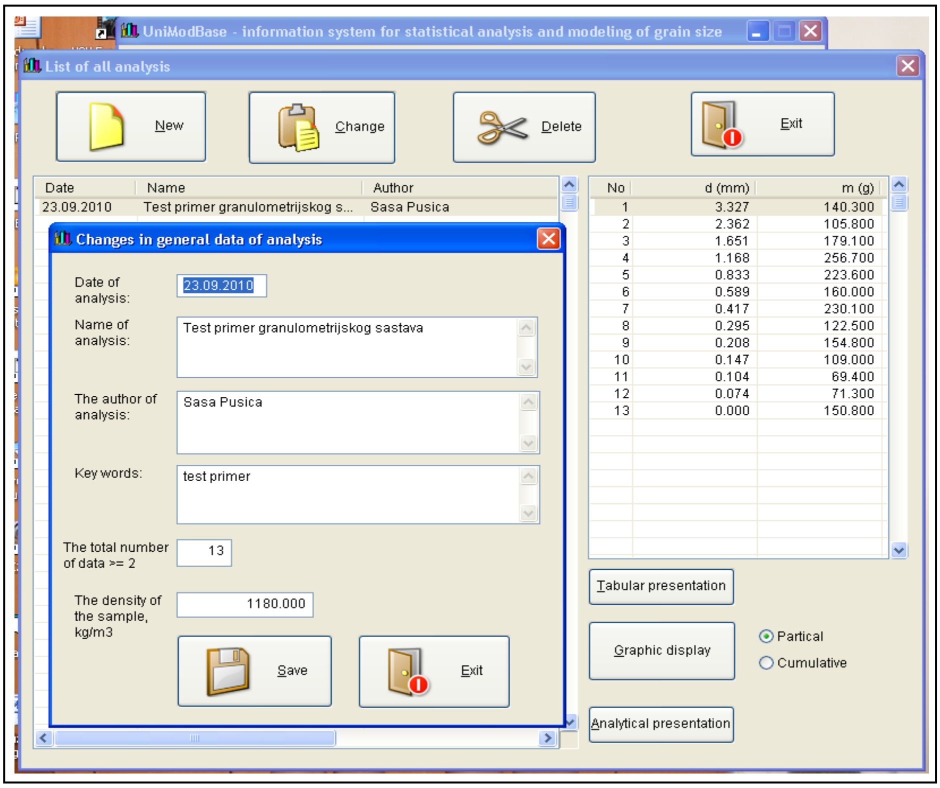


Figure 17 - Insert new analysis in "UniMod Base" information system

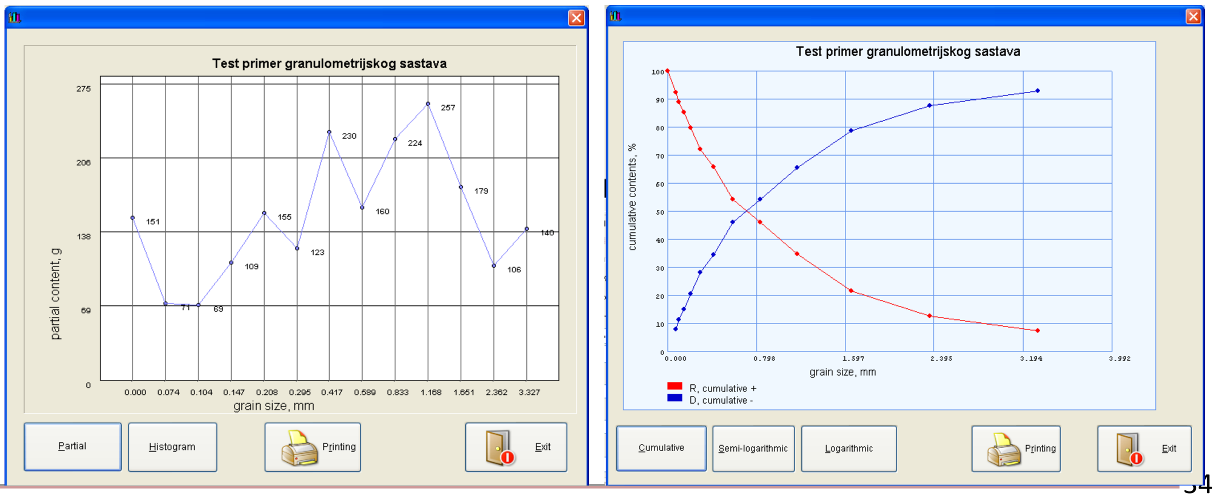


Figure 18 - Graphical presentation of grain size distribution in "UniMod Base" information system

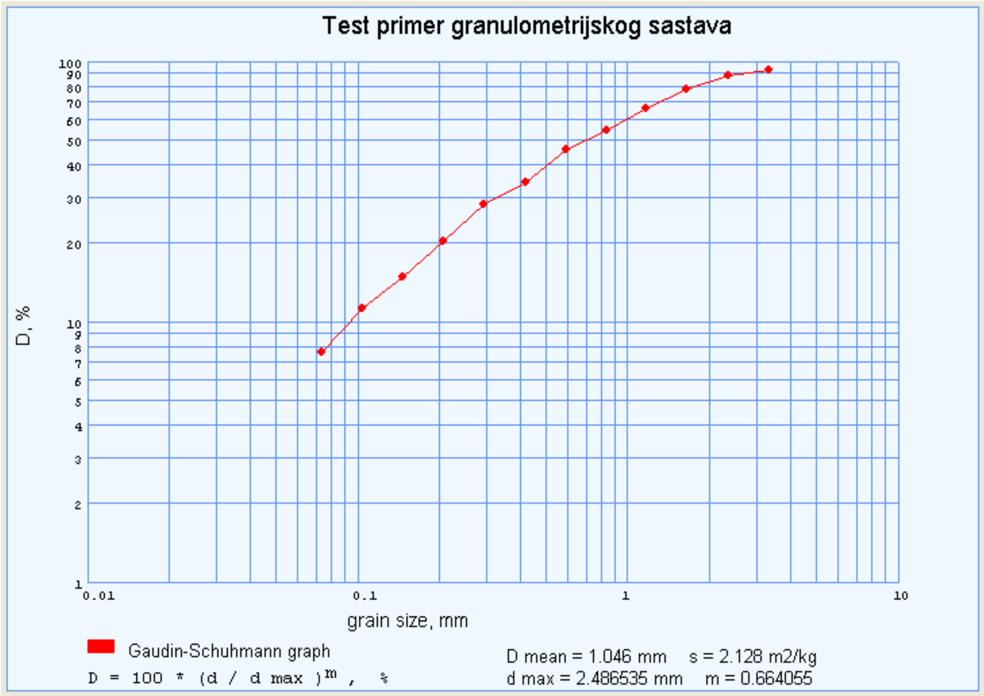


Figure 19 - Rosin-Rammler graphical presentation maked with "UniMod Base" information system

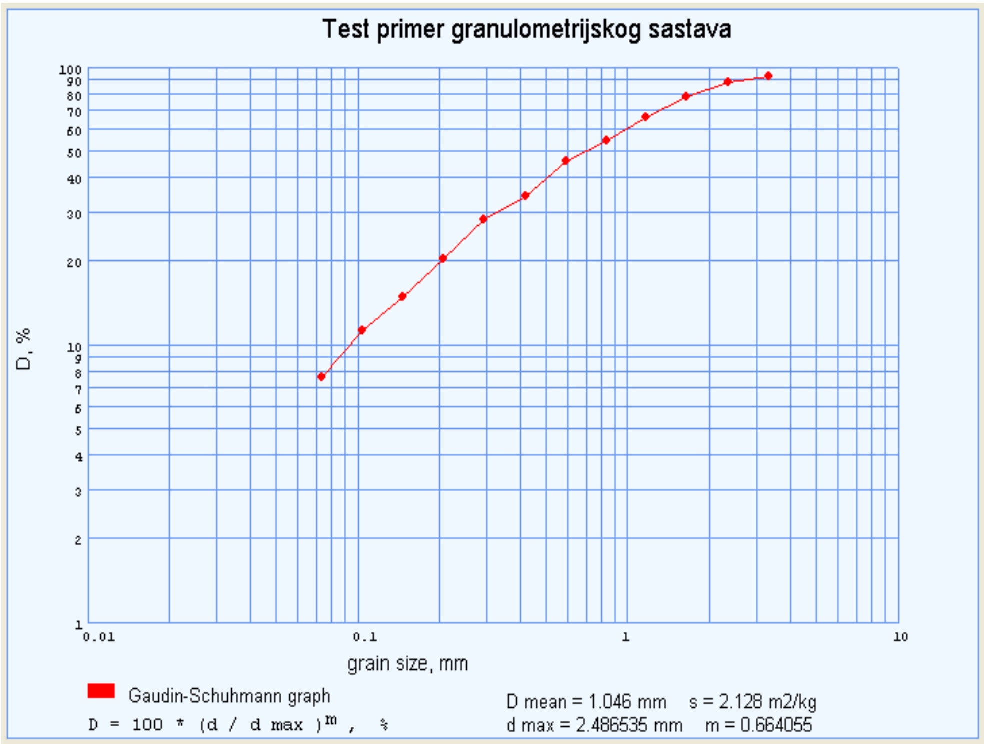


Figure 20 - Gaudin-Schuhmann graphical presentation maked with "UniMod Base" information system

Volume 2 No.7, JULY 2011 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

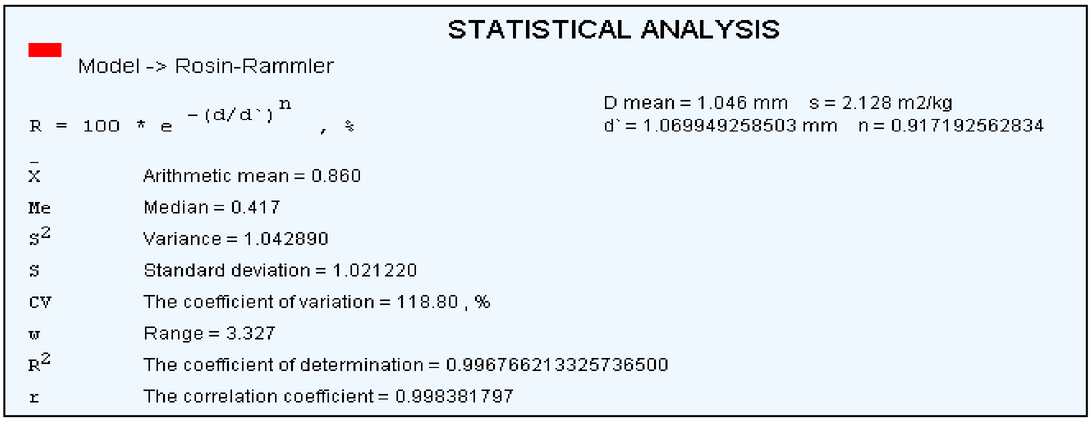


Figure 21 - Statistical analysis and Rosin-Rammler equation with "UniMod Base" information system

REFRENCES

[1] N. Magdalinović, "Usitnjavanje i klasiranje mineralnih sirovina", Naučna knjiga, Beograd, 1991,

[2] N. Ćalić, "Teorijski osnovi pripreme mineralnih sirovina", Univerzitet u Beogradu, 1990.

[3] M.M. Gifing, "Priprema mineralnih sirovina", Univerzitet u Beogradu, Rudarsko-Geološki fakultet, Beograd, 1986.

[4] L.D. Schroeder, "Understanding Regression Analysis, Sage Pub, 1986.

[5] K. Surla, "Aproksimacija funkcija i regresiona analiza", Univerzutet u Novom Sadu, 1998.

[6] A.S. Hadi and S. Chatterjee, "Regression Analysis by Example", (fourth edition), Wiley, 2006.

[7] M. Crowder, D.J. Hand, and N. Adams, "Methods and Models in Statistics", Imperial College Press, 2004.

[8] N.R. Draper and H. Smith, "Applied Regression Analysis", 3rd Edition, Wiley, 1998.

[9] D.C. Montgomery, E.A. Peck, and G.G. Vining, "Introduction to Linear Regression Analysis" (3rd ed.), New York: Wiley, 2001.

[10] K. Kero and B. Bojanić-Glavica, "Statistika u primjerima", Fakultet organizacije i informatike, Varaždin, 2003.

[11] I. Šošić, "Metode statističke analize", Ekonomski fakultet, Zagreb, 1991.

[12] N. Vuković, "Statističko zaključivanje", Fakultet Organizacionih Nauka, Beograd, 2007.

[13] S. Pušica and M. Pušica, “DATABASE INFORMATION SYSTEM FOR THE RESULTS IN MINERAL PROCESSING”, International Journal of Reviews in Computing, Volume 1, p.61-66, December, 2009.

[14] S. Pušica, Z. Nikolić, and V. Šćekić, “INFORMATION SYSTEM FOR MODELING THE GRAIN SIZE DISTRIBUTION”, Journal of Theoretical and Applied Information Tecnology, Volume 23 No.1, p.7-15, January, 2011.

[15] S. Pušica, Z. Nikolić, and V. Šćekić, “DEVELOPMENT OF INFORMATION SYSTEM FOR MODELLING OF GRAIN SIZE DISTRIBUTION USING IDEF METHOD”, Journal of Theoretical and Applied Information Tecnology, Volume 27 No.2, p.105-118, May, 2011.

AUTHOR PROFILES

M.Sc. SAŠA PUŠICA graduated at the Technical Faculty in Bor 1993 on the Department of Mineral Processing. Since 1995 until today, he is professionally engaged in developing business information systems. He completed master's thesis in 1999 at the Technical Faculty in Bor. On the FIM in Krusevac, 2010 he started doctoral dissertation in the field of information systems, which is in the final phase.