

Received: April 6, 2016, reviewed, accepted: July 11, 2016

APPLICATIONS OF ANOVA IN MINERAL PROCESSING

Tomasz NIEDOBA^{*}, Paulina PIĘTA

AGH – University of Science and Technology, Department of Environmental Engineering and Mineral Processing Cracow, Poland

Abstract: The paper presents a scheme of conducting ANOVA, which can be a statistical method using for the analysis of various complex phenomena in different fields of science. It is mainly focused on presentation of different areas of application at statistical model, especially in mineral engineering. Selected experiments in which ANOVA was used successfully to study the process of flotation and properties of three types of coal are presented in this study. Mineral resources are characterized by different properties and the flotation process depends on many factors which must be carefully controlled to obtain an expected separation efficiency of the process under industrial conditions. These facts determine the complexity of mineral processing structure. This choice of ANOVA was dictated by the adjustment of methodology to the nature of the analyzed phenomena. Based on this analysis it can be stated that the analysis of variance is appropriate to consider problems associated with the mineral engineering. The examples described in the paper confirmed a high potential which ANOVA carries. However, it should be noted that the one-dimensional analysis of variance has some limitations.

Keywords: *mineral processing, copper-molybdenum sulphides flotation, hard coal properties, ANOVA, multidimensional analysis*

INTRODUCTION

Variability is a property which characterizes all recurring scientific measurements and observations of various phenomena. Both controlled and uncontrolled factors influence each stage of the experiment and its final result. A number of random factors, i.e. aberrations appearing during the experiment, could disturb the value of measurements. As a result, the more complex phenomenon is being analyzed, the bigger is the probability of measuring diversity of some features. The problem was considered by Fisher, who was the founder of the analysis of variance.

^{*} **Corresponding author:** tniedoba@agh.edu.pl (T. Niedoba)

The beginning of the ANOVA is connected with the development of a field known as experimental design in the 1920s and 1930s. During this period Ronald A. Fisher worked in Agricultural Experimental Station in Rothamsted (England), and he formulated the ANOVA as a basic statistical method used for an analysis of experimental data (1923). In retrospect, it can be said that the theory of experiment planning together with the analysis of variance both originate from the field of agriculture, where they were of crucial significance. Originally, analyses were related to the methods and conditions for crops, so the main process of planning and carrying out the experiment with obtaining results lasted long because of the vegetation period of plants (Stanisz, 2007; Wątroba, 2002).

It can be claimed that the ANOVA is a very useful method, as it allows the assessment of the influence of some controlled factors on experimental results. The analysis of variance can be carried out according to different schemes. The first consists in examining the variance in a single classification, and the other provides for a multivariate analysis. It is also possible to carry out the analysis of variance in a combined scheme (Masiya & Nheta, 2014).

THE ANALYSIS OF VARIANCE

The analysis of variance (ANOVA) is an examination technique of experiment or observation effects. What constitutes the basis to carry out the ANOVA is verifying the assumptions concerning the homogeneity of variances and the normal distribution of these variables.

This chapter shows a schematic analysis of variance for one-way classification. The null hypothesis one-dimensional analysis of variance assumed equal mean values of the studied phenomenon characteristics. For the one-dimensional analysis, the null-hypothesis assumed that $H_0: \bar{\mu}_{1j} = \bar{\mu}_{1j} = \bar{\mu}_{1j} = \dots = \bar{\mu}_{nj}$, and the alternative hypothesis H_1 was: $\bar{\mu}_{1j} \neq \bar{\mu}_{1j} \neq \bar{\mu}_{1j} \neq \dots \neq \bar{\mu}_{nj}$. A measure of differences between the actual and hypothetical is F statistics, which took its name after its originator – Fisher (Wątroba, 2002).

The impact of individual parameters that are more or less important in the examined phenomenon can be estimated with the use of the analysis of variance. A description of the test and the equations used in the ANOVA are presented below.

For the analysis, it is necessary to calculate the mean values \bar{x}_i of the observed characteristics in each group and the total size of the mean values \bar{x} , taking into account all the observations of individual groups. The next step is to determine the sums of squares effects SS_{Effect} and sums of squares errors – SS_{Error} . The values SS_{Error} reflect the amount of errors that occur during the observation resulting from the natural randomness. Taking into account the number of samples obtained following formulas (Stanisz, 2007):

$$SS_{\text{Effect}} = \sum_{i=1}^a n_i (\bar{x}_i - \bar{x})^2 \quad (1)$$

$$SS_{\text{Error}} = \sum_{i=1}^a \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \quad (2)$$

where:

n_i – the number of random variables of the studied properties,

\bar{x} – the general mean for all observations,

\bar{x}_i – the arithmetic mean of a characteristic in the studied group,

x_{ij} – the value of a single property measurement.

It is clear that in the conditions of the experiment different factors affect the total variability between the analyzed groups SS_{Total} , which is a sum of SS_{Effect} – measure of the differences between the groups and the SS_{Error} – variation within groups. Here, it can be concluded that if the mean value of the measurements in the groups is different, it is likely that the variation between groups is greater than within them. Another essential stage of the analysis, the ANOVA, is to determine the number of degrees of freedom according to the formulas (3) and (4):

$$df_{\text{Effect}} = a - 1 \quad (3)$$

$$df_{\text{Error}} = n - a \quad (4)$$

where:

n – the number of independent observations x_{ij} equal $n = \sum_{i=1}^a n_i$

a – the number of independent variables,

The basic formulas (1), (2), (3) and (4) enable to calculate mean sums of squares between and within groups, which are the estimators of variance for all population. Mean sums of squares between and within groups are determined according to the formulas (5) and (6).

$$MS_{\text{Effect}} = \frac{SS_{\text{Effect}}}{df_{\text{Effect}}} \quad (5)$$

$$MS_{\text{Error}} = \frac{SS_{\text{Error}}}{df_{\text{Error}}} \quad (6)$$

The verification the null hypothesis is based on the performing F-test, comparing variances:

$$F = \frac{MS_{\text{Effect}}}{MS_{\text{Error}}} \quad (7)$$

The value of F-test indicates an essential diversity or rather the lack of it between groups. When the quotient $MS_{\text{Effect}}/MS_{\text{Error}}$ reaches a value close to unity, the null hypothesis is correct, hence, the mean values of examined characteristics in the groups did not differ in a statistically significant way. Otherwise, the value of the F-test is deflected upward by one. If the deviation is large enough to conclude that there is a differentiation between groups, the Fisher-Snedecor distribution with degrees of freedom of $a-1$ and $n-a$ will be applied. If the F statistic is greater than the critical value F_{α} read from tables, the null hypothesis shall stand as rejected, and simultaneously the alternative hypothesis of the occurrence of variation between groups will be approved (Stanisz, 2007).

The computational scheme of the ANOVA that was discussed, explains successive stages of implementing this statistical method of data analysis. It was considered relevant to present reliable mechanism of variance analysis, because it is one of the most important methods of statistics, which is more and more widely used in various fields of science.

The ANOVA is used in different fields of empirical research, which are the subject of the phenomenon with a complex structure. Issues related to medicine, psychology, chemistry, and agriculture can be listed as an example; moreover, the analysis is used in various industries to solve complex technological problems and optimize the production processes, too (Aranowska & Rytel, 2010; Raykov & Marcoulides, 2008; Güler et al., 2012; Rao et al, 2012).

In the next part of the article selected experiments are presented, in which ANOVA was used successfully to study the complex processes of mineral processing and to study the characteristics of grained material.

STATISTICAL ANALYSIS OF PROCESSING PROCEDURES

All technical sciences aim to formalize the description of processes, their optimization and study of the properties of material and products, hence mathematical models are widely used and developed. There is no doubt that different statistical methods allow for a precise identification of the processes, as well as their development. This involves both: process engineering and mineral engineering (Tumidajski, 2010; Brożek & Surowiak, 2010).

THE APPLICATION OF ANOVA IN THE OPTIMIZATION OF FLOTATION

Froth flotation is based on the separation of particles, which are dispersed in a flotation pulp. Complex chemical and physical phenomena occur during the separation process. This is associated with the interaction between a solid phase, liquid and gas, which are in continuous contact with one another. The hydrophobic particles attach to

the air bubbles rising to the surface and form a froth phase. The flotation process depends on many factors which must be carefully controlled to obtain an expected separation efficiency of the process under industrial conditions. A complete description of the flotation is therefore a complex and multidimensional issue. Hence, the optimization entails the fact that a number of parameters describing the conditions of the process, as well as material properties must be taken into account (Drzymala, 2007). The conclusion is that the application of the analysis of variance to study the optimal parameters for the flotation process is fully justified. The studies conducted by scientists from Amirkabir University of Technology and Shahid Bahonar University concerned the selection of optimal conditions for the process of selective flotation of copper-molybdenum ore in the Sungun Copper Concentrator Plant. A thorough description of the technology, which is used in the processing plant, is presented in the article *Application of D-optimal design for optimizing copper-molybdenum sulphides flotation* by Nakhaei, Irannajad, Sam and Jamalzadeh.

Tab. 1. Experimental parameters and their relative levels *Source: Nakhaei et al., 2015*

Collector			Frother		Diesel dosage [g/Mg]	d_{80} [%]	Reagent dosage [g/Mg]				
A	B	C	D	E	F	G	H	J	K	L	M
Z ₁₁	R407	0	A65	A70	0	80	10	15	0	5	5
Z ₆	Nascol	C4132	DF 250	MIBC	10	100	15	20	7	10	10
-	A7240	X231	Flomin	Pine oil	20	120	20	25	15	15	15

The flotation tests were performed according to the D-optimal factorial design of experiments. The random variables in this study were: type of collectors (A, B, C), type of frothers (D, E), collector and frother dosages (H, J, K, L, M), dosage of diesel (F) and flotation feed size (G). The levels of variables were given in Table 1 (Nakhaei et al., 2015).

First of all, particular assumptions were verified, which constituted the basis for the analysis of variance. The tests confirmed that the investigated factors had a normal distribution and they characterized by a homogeneity of variance.

In the experiment, the analysis of variance was conducted separately to examine the impact of various factors on Cu and Mo recovery and grade. According to the analysis of variance, the Fisher test (Tab. 2) with a very low probability values $P < 0.05$ indicates that the models for Cu and Mo recovery and grade are highly significant (Nakhaei et al., 2015; Stanisiz, 2007).

The lower value of the ratio P, the greater the impact of a factor to the flotation effectiveness. Collectors from group B had the greatest importance for Cu recovery. In the case of Mo recovery a dosage of diesel was the most significant. However, Cu and Mo grade were dependent on collectors Z₁₁ (isopropyl xanthate sodium), Z₆ (amil

xanthate petasium) in the greatest extent. A more detailed analysis of the data was possible on the basis of the graphs shown in Figures 1 and 2. The diagrams for the most important factors influencing the recovery of Cu and Mo are listed as an example (Nakhaei et al., 2015).

Tab. 2. Analysis of variance for Mo, Cu recovery and grade *Source: Nakhaei et al., 2015*

Source		F-value	P-value	
Cu RECOVERY	Model	19,93	0,0018	
	Variables	A	16,80	0,0094
		B	20,29	0,0040
		D	10,15	0,0174
		H	9,81	0,0186
		J	8,44	0,0249
CV = 0.85%; $R^2 = 0.99$; Adj. $R^2 = 0.93$, C-E-F-G-K-L-M = not significant				
Cu GRADE	Model	31,19	0,0006	
	Variables	A	138,57	<0,0001
		B	16,69	0,0061
		D	13,82	0,0092
		E	13,18	0,0102
		H	8,90	0,0225
L	46,14	0,0006		
CV = 4.2%; $R^2 = 0.99$; Adj. $R^2 = 0.96$, C-F-G-J-K-L-M = not significant				
Mo RECOVERY	Model	32,97	0,0005	
	Variables	C	15,94	0,0068
		D	6,55	0,0401
		F	54,62	0,0004
		G	22,83	0,0031
		K	7,98	0,0278
M	5,32	0,0478		
CV = 1.92%; $R^2 = 0.99$; Adj. $R^2 = 0.96$, A-B-E-H-J-L = not significant				
Mo GRADE	Model	15,07	0,0034	
	Variables	A	46,60	0,0010
		C	5,48	0,0442
		E	12,85	0,0107
		G	16,35	0,0064
CV = 2.71%; $R^2 = 0.98$; Adj. $R^2 = 0.92$, B-D-F-H-J-K-L-M = not significant				

Figure 1 shows the effect of flotation process parameters on Cu recovery. In general, higher recovery of Cu was obtained for Z₁₁, which was used as the collector. Also, R407 is the best choice among reagents from B group. Furthermore, the highest Cu recovery could be achieved with the A65 as the frother. As shown in Fig. 1, the Cu

recovery increased by increasing the primary collector (H) dosage and reached its maximum value at 20 g/Mg.

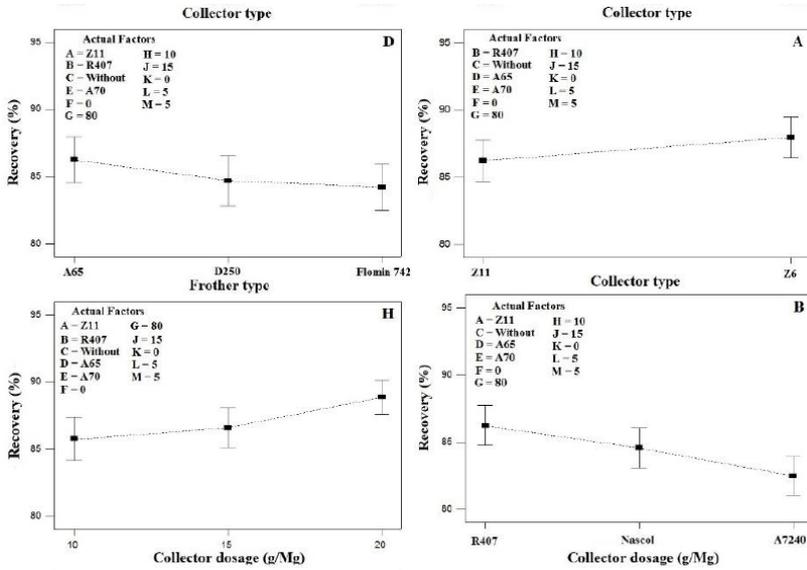


Fig. 1. Effect of process parameters on Cu recovery
Source: Nakhaei et al., 2015

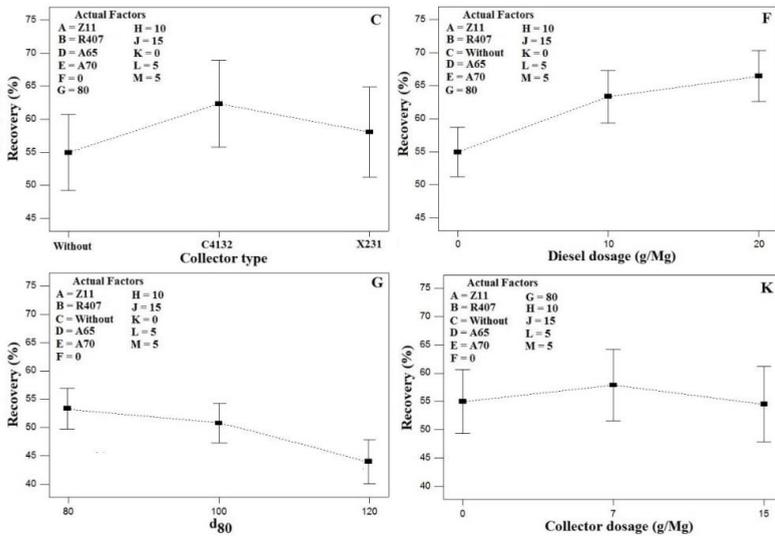


Fig. 2. The effect of process parameters on Mo recovery
Source: Nakhaei et al., 2015

The highest Mo recovery is obtained when C4132 (7 g/Mg) is used as a second promoter collector as well as the diesel dosage and d_{80} of feed are equal to 20 g/Mg and 80 μm , respectively. An increase of size has a negative impact on recovery and grade.

Generally, the analysis of variance showed that the primary promoter collector was the most significant parameter affecting the recovery of Cu, while diesel dosage and d_{80} were the most significant parameters influencing the Mo recovery. It can be said that the statistical method works when examining the impact of certain factors on the beneficiation process. ANOVA does not take into account the interactions between factors and their further effects on the results of the experiment. In order to investigate the interaction between the factors the factorial design was carried out. It enabled to optimize a froth flotation. The results of optimization showed that the highest possible Cu and Mo grade and recovery were obtained for (Nakhaei et al., 2015):

1. Z₁₁ (a primary collector) 20 g/Mg,
2. R407 (a first promoter collector) 20 g/Mg,
3. X231 (a second promoter collector) 7 g/Mg,
4. A65 (a frother) 15 g/Mg,
5. pine oil (a frother) 5 g/Mg,
6. d_{80} (a feed size) 80 μm ,
7. diesel dosage 20 g/Mg.

THE APPLICATION OF ANOVA IN THE STUDY OF GRAINED MATERIALS PROPERTIES

Statistical analyses in mineral processing are difficult due to the complexity of the operations unit, but the main difficulty is the fact that the processed material is a set of various grains of raw materials. A granular material has many properties that determine its susceptibility to a particular manner of beneficiation. Such features are: a content and a liberation efficiency index of valuable minerals, particle size, particle density, physicochemical properties and other physical and geometrical characteristics (Niedoba, 2013; Tumidajski, 2010; Tumidajski et al., 2012).

The other direction of the ANOVA application is presented in the research paper *Multidimensional analysis of different hard coal properties*. It presents an analysis of variance which allows examining the differences between three types of coals. Thanks to the ANOVA analyses, three main properties of coal were determined, which differentiated the aforementioned coal types, and a comprehensive study enabled a thorough characterization of variability properties of the tested grained material [Pięta, 2015]. Coals classified, according to Polish classification, as types 31, 34.2 and 35. Samples came from some mines, located within the Upper Silesian Industrial Region. The raw coal was determined by the following characteristics: particle size, density, ash con-

tent, volatile matter content, sulfur content, heat of combustion, and analytical moisture.

Data analysis began by verifying the hypothesis of normal distribution of the dependent variables and performance tests of homogeneity of variance. The results showed that not all parameters of coals are characterized by a normal distribution and homogeneity of variance. A Shapiro-Wilk test allowed to determine the characteristics of coal, which did not characterize by a normal distribution, however, making reference to the resistance of Brown-Forsythe test to the relinquishment of the assumption of normal distribution of variables, the homogeneity of variance was checked for all parameters. In this way it was possible to exclude two properties (sulfur content, moisture analysis) from further analysis, which breached the assumption of homogeneity of variance.

The results obtained by F-test identified the characteristics of raw materials, which are essential as far as diversity in coal types is concerned. The F statistic was determined for all parameters in all size fractions. Critical value F_{α} of Fisher-Snedecor distribution was set, as well, which allowed the classification of the characteristics influencing the differentiation of coal types.

While analyzing the results of F test it was concluded that all the properties of coal caused differentiation between types 31, 34.2 and 35 in size fractions: 3.15-6.3, 8-10, 10-12.5, 16-20 mm. However, the complex structure of the analyzed data does not allow the identification of hard coal types with an overall analysis of the observed characteristics in all size fractions. At the 0.5-1, 1-3.15, 6.3-8, 12.5-14, 14-16 mm size fractions, the parameters: density, heat of combustion, ash content did not affect coal differentiation significantly, from a statistical point of view (Fig. 4).

Tab. 3. The results of F test (significant results have been highlighted)

Source: Pięta, 2015

F test				
H_0 must be rejected $\leftrightarrow F_{emp} > F_{\alpha}$, where $F_{\alpha}(2,97) = 3,0903$				
Size fraction [mm]	Density	Heat of combustion	Ash content	Volatile matter content
0.5-1	1.466594	5.487025	2.467011	44.13694
1-3.15	1.134825	0.073532	1.815244	3.441177
3.15-6.3	4.583223	6.009332	6.946015	73.70378
6.3-8	2.399904	13.15684	4.973166	20.0778
8-10	6.414866	10.15947	7.465963	15.53309
10-12.5	6.549542	8.846621	7.088464	14.30923
12.5-14	0.335797	2.175174	2.395407	6.058163
14-16	7.623756	24.66782	0.098286	14.31082
16-20	12.08655	12.44888	10.54814	15.00584

Looking at the overall results for the parameters it turned out that only the volatile matter content is a feature differentiating the tested coals. It can be therefore concluded that the value of this parameter is a major factor, by which coal is classified according to its properties and use (Fig. 3). The F test results are presented in Table 3.

The collected data also served to draw up charts representing the relationship between the parameters of coal types, taking into account the granulation of material. The study introduced and elaborated on the charts for those qualities, which set the four most important characteristics that had the greatest impact on the diversification of raw materials (Fig. 3-4). Despite demonstrating significant differences between types of coal based on the analysis of variance performed on the properties of energy resources, in some cases there might be occur interactions between the coals. The relationships can be observed during graphical interpretation of charts representing the mean parameter values (Fig. 3-4).

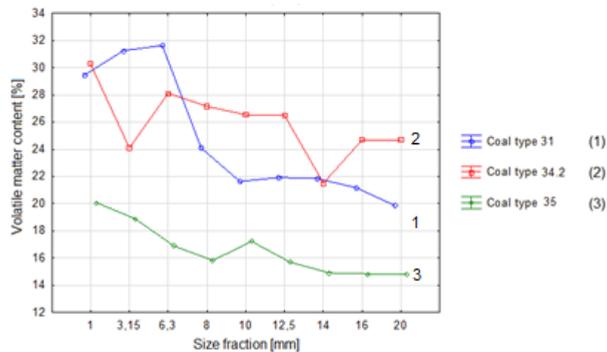


Fig. 3. Graphs of the mean values volatile matter content value of 31, 34.2, 35 coal types
Source: Pięta, 2015

CONCLUSIONS

Statistical methods have been used for ages into various fields of science, but more frequently an increased interest in multivariate analyses is observed, which allows a more complete examination of the object. There is a tendency for a comprehensive study of various phenomena to be carried out in order to explore thoroughly a number of factors affecting them. The ANOVA analysis brings such a possibility. This is a statistical method widely used in various fields. As it was shown in the article, the analysis of variance is appropriate to consider the problems associated with the mineral engineering. The examples described in this paper confirmed the high potential which the ANOVA carries. The consideration of optimization problems and differentiation of selected raw materials are only some of the issues with can be faced in min-

eral processing. A number of works serve as an evidence of the widespread use of the analysis of variance.

Analyzing multivariate phenomena that are associated with material processing and other areas, it should be noted that the one-dimensional analysis of variance has some limitations. It assumes the relationship between the parameters under consideration, although almost always the structure features is connected. Hence, the one-dimensional analysis of variance will show only independent influence of each variable on the differences between the groups. The method must not be ignored, however, using the multiple factor ANOVA to analyze such complex phenomena will allow to examine the strength of interactions between the studied properties.

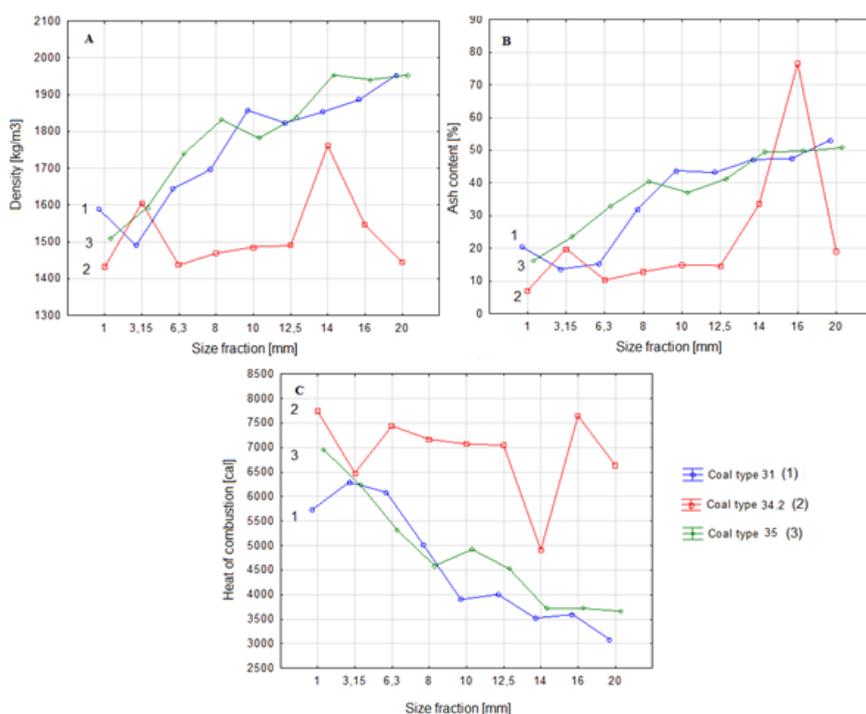


Fig. 4. Graphs of the mean parameter values (A – density, B – ash content, C – heat of combustion) of 31, 34.2, 35 coal types

Source: Pięta, 2015

ACKNOWLEDGEMENT

The paper is the effect of the scientific project no. 15.11.100.021.

BIBLIOGRAPHY

- BROŹEK M., SUROWIAK A., 2010. *Argument of separation at upgrading in the jig*. Arch. Min. Sci., Vol. 55, No. 1, 39-40.
- ARANOWSKA E., RYTEL J., 2010. *Multidimensional variance analysis – MANOVA*, Social Psychology, Vol. 5, No. 14, 117-141 [in Polish].
- DRZYMAŁA J., 2007. *Mineral Processing. Foundations of theory and practice of metallurgy*. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
- GÜLER T., AKDEMIR U., 2012. *ANOVA modeling on sintering parameters and frequencies, affecting microstructure and dielectric constant of Nb doped BaTiO₃*, Transactions of Nonferrous Metal Society of China, Vol. 22, 199–205.
- MASIYA T., NHETA W., 2014. *Flotation of nickel-copper sulphide ore: optimisation of process parameters using Taguchi method*, Proceedings of the International Conference on Mining, Material and Metallurgical Engineering, Prague (Czech Republic), No. 113, 1-11.
- NAKHAEI F., IRANNAJAD M., SAM A., JAMALZADEH A., 2015. *Application of d-optimal design for optimizing copper-molybdenum sulphides flotation*, Physicochemical Problems of Mineral Processing, Vol. 52, No. 1, 252–267.
- NIEDOBA T., 2013. *Methodological elements of applying two- and multi-dimensional distributions of grained materials properties to coal beneficiation*, Mineral Resources Management, Cracow, Vol. 29, No. 2, 155-172 [in Polish].
- PIĘTA P., 2015. *Multidimensional analysis of different hard coal properties*, X Krakowska Konferencja Młodych Uczonych, AGH - University of Science and Technology, Cracow (Poland), 98.
- RAO S., PADMANABHAN G., 2012. *Application of Taguchi methods and ANOVA in optimization of process parameters for metal removal rate in electrochemical machining of Al/5%SiC composites*, International Journal of Engineering Research and Applications (IJERA), Vol. 2, No. 3, 192-197.
- RAYKOV T., MARCOULIDES G.A., 2008. *An introduction to applied multivariate statistics*. New York, Routledge.
- STANISZ A., 2007. *The accessible statistics course using STATISTICA PL on examples from medicine*. Volume 2. Linear and nonlinear models. Cracow, StatSoft [in Polish].
- TUMIDAJSKI T., 2010. *Actual tendencies in description and mathematical modelling of mineral processing*, Mineral Resources Management, Cracow, Vol. 23, No.6, 111-123, [in Polish].
- TUMIDAJSKI T., NIEDOBA T., SARAMAK D., 2012. *Introduction to mathematical statistics of grained materials*, AGH Journal of Mining and Geoengineering, Cracow, Vol. 36, No. 4, 167-177.
- WĄTROBA J., 2002. *Introduction to analysis of variance* [on-line], StatSoft Polska, [access: 21.04.2015]. At the following website: <http://www.statsoft.pl/Czytelnia> [in Polish].