IOP Conf. Series: Materials Science and Engineering **427** (2018) 012008 doi:10.1088/1757-899X/427/1/012008

Analysis of distributions of various coal types properties by means of statistical methods

T Niedoba, P Pięta and A Surowiak

AGH University of Science and Technology, Faculty of Mining and Geoengineering, Department of Environmental Engineering and Mineral Processing, Mickiewicza 30, 30-059 Krakow, Poland

tniedoba@agh.edu.pl

Abstract. Coal can occur in various natural forms and its quality depends on its various properties. Polish coal mines extract and process various coal types, according to the Polish nomenclature. The paper presents an attempt of determining whether is possible to statistically differ the typical coal types occurring in Polish hard coal mines by means of traditional statistical tests. To this purpose three coal types, namely 31, 34.2 and 35 were selected from three Polish mines located in Upper Silesia. After appropriate preparation of the raw coals selected coal properties were examined, including combustion heat, ash contents, sulfur contents, volatile parts contents and analytical moisture. Then, by means of several statistical tests it was verified whether it is possible to prove statistically that these coal types differ significantly not knowing the type of distribution they represent. To this purpose such tests as Kruskal-Wallis test, test of series, Wilcoxon test and median test were applied. The results were statistically evaluated and the conclusions were made. The presented methodology may serve to compare not only coal types but also other raw materials.

1. Introduction

Coal as a fuel is still crucial one in Poland. Despite the environmental aspects and European Union policy its production is still very important for the country. However, its quality is also very important taking into consideration both environmental and economic aspects. Polish hard coal is located mainly in Upper Silesian Industrial Region, where several mines are situated. They produce various sorts of hard coal. There is a question – is it possible to easily recognize the type of coal basing only on its selected properties. The Authors already tried to answer this question but this time the analysis was conducted by means of traditional non-parametric statistical tests, without assumption of known form of distribution function. The whole calculating procedure is described in further part of the paper.

2. Preparation of data

Three types of coal are considered in the paper, namely 31, 34.2 and 35 according to Polish nomenclature. On the basis of the results of their basic properties investigations such parameters as combustion heat, ash contents, sulfur contents, volatile parts contents and analytical moisture were examined. These properties are treated as random variables of unknown distribution functions. By means of selected statistical methods the distributions of the same property for various coal types were compared in purpose of determining their similarity. To achieve this goal the separation of coal feeds of all considered coal types into particle size fractions 0.50-1.00; 1-00-3.15, 3.15-6.30, 6.30-8.00, 8.00-10.00, 10.00-12.50, 12.50-14.00, 14.00-16.00, 16.00-20.00 was performed. The measurements of the researched properties were done and the obtained results are presented in tables 1-3 for coal types 31, 34.2 and 35, respectively. Remark: data for combustion heat are given in J, but the calculations were made for coal.

Table 1. Juxtaposition of the results for coal, type 31.

Size *d*, mm	Combustion heat, J	Ash, %	Sulfur, %	Volatile parts, %	Analytical moisture W_a
0.50-1.00	21218.70	31.59	1.31	26.02	3.32
1.00-3.15	20506.95	30.92	1.45	26.35	2.78
3.15-6.30	20929.81	31.71	1.42	27.58	2.87
6.30-8.00	19770.07	33.53	1.23	25.43	2.78
8.00-10.00	18526.59	34.20	0.99	25.90	2.77
10.00-12.50	19430.94	32.88	1.23	26.02	2.94
12.50-14.00	18258.63	34.56	1.10	25.59	2.70
14.00-16.00	18802.92	34.74	0.98	25.78	2.79
16.00-20.00	19929.17	33.17	0.97	25.62	2.98

Table 2. Juxtaposition of the results for coal, type 34.2.

Size *d*, mm	Combustion heat, J	Ash, %	Sulfur, %	Volatile parts, %	Analytical moisture W_a
0.50-1.00	23169.75	28.75	0.89	25.19	0.66
1.00-3.15	23077.64	28.98	0.66	24.30	1.05
3.15-6.30	22792.94	27.49	0.64	24.74	0.96
6.30-8.00	23278.61	31.91	0.58	24.55	0.88
8.00-10.00	22591.97	31.05	0.50	23.92	1.12
10.00-12.50	23211.62	31.74	0.30	24.60	1.05
12.50-14.00	22324.02	29.44	0.46	25.65	0.97
14.00-16.00	26456.39	28.09	0.66	22.63	0.74
16.00-20.00	19313.71	28.78	0.34	22.25	1.01

Table 3. Juxtaposition of the results for coal, type 35.

Size *d*, mm	Combustion heat, J	Ash, %	Sulfur, %	Volatile parts, %	Analytical moisture W_a
0.50-1.00	22922.73	31.17	0.60	17.56	1.08
1.00-3.15	22051.88	33.29	0.57	17.55	1.02
3.15-6.30	22139.80	32.41	0.58	17.16	1.37
6.30-8.00	21490.84	34.09	0.58	14.40	1.23
8.00-10.00	22085.37	34.88	0.61	17.54	1.37
10.00-12.50	21788.11	34.44	0.50	17.31	1.36
12.50-14.00	21507.59	34.33	0.57	17.81	1.40
14.00-16.00	21930.46	33.89	0.52	17.39	1.32
16.00-20.00	21411.30	34.77	0.69	17.25	1.39

3. Methodology and results

Let mark the random variable related to j^{th} property in i^{th} coal type (i = 1, 2, 3; j = 1, 2, 3, 4, 5) as X_{ij} and distribution function of X_{ij} as $F_{ij}(x)$. First, the verification of H_0 hypothesis was performed that the distribution function of j^{th} coal property is identical for all considered coal types. That means that

IOP Conf. Series: Materials Science and Engineering **427** (2018) 012008 doi:10.1088/1757-899X/427/1/012008

$$H_0: F_{1j}(x) = F_{2j}(x) = F_{3j}(x) \tag{1}$$

To verify this hypothesis the Kruskal-Wallis χ^2 statistics was applied, which represents asymptotic χ^2 distribution function of two degrees of freedom [1-4]. The χ^2 statistics is in form:

$$\chi^2 = \frac{12}{n(n+1)} \sum_{i=1}^{3} \frac{R_i{}^2}{n_i} - 3(n+1) \tag{2}$$

where: n_i – quantity of the sample for i^{th} coal type (i = 1, 2, 3); $n = n_1+n_2+n_3$; R_i – sum of ranges of the results of i^{th} sample (i = 1, 2, 3).

The values of the χ^2 test for individual coal properties are presented in the table 4.

Table 4. Values of χ^2 test.

Coal property	χ^2
Combustion heat	20.174
Ash	17.620
Sulfur	20.056
Volatile parts	25.580
Analytic moisture	25.056

The critical area for Kruskal-Wallis test is range $\left(\chi^2_\alpha(2), +\infty\right)$ [2-4]. For the significance level α=0.05 the value of χ^2_α=5.991. On the basis of the results presented in table 4 the values of the test for all coal properties belong to the critical area, so the H_0 hypothesis should be rejected. This means that the distributions of the investigated coal properties depend on the coal type.

Next, the distributions of the researched coal properties were compared in pairs. So, the H_0 hypotheses were in form:

$$H_0: F_{i,j}(x) = F_{k,j}(x) \tag{3}$$

where: i, k = 1, 2 ,3; j = 1, 2, ..., 5; i≠k

To this purpose the test of series and the Wilcoxon's test were applied [2-4]. The results of both selected samples were put in growing order. In the test of series the test statistics is number of series *K*, where a series is each maximum sub-series including only elements of the same sample [2]. The critical area for test of series is a range $<2, k_\alpha(n_1, n_2)>$, where $k_\alpha(n_1, n_2)$ is the value read from the statistical tables for series distribution and significance level α.

In Wilcoxon test U statistics is the tested value which is equal to the sum of inversions of the elements of one of the samples according to the elements of the second sample (equation (4).

$$U = u_1 + u_2 + \cdots + u_k \tag{4}$$

where u_i is the number of elements of first sample being not bigger than the i^{th} element of second sample. The critical area of Wilcoxon test is sum of ranges $<0, u_\alpha(n_1, n_2)> \vee <n_1n_2 - u_\alpha(n_1, n_2), n_1n_2>$., where $u_\alpha(n_1, n_2)$ is a value read from statistical tables for significance level α.

The results of the test of series and Wilcoxon test are presented in tables 5 and 6.

Table 5. Test of series – values of *K*.

	(31, 34.2)	(31, 35)	(34.2, 35)
Combustion heat	4	2	5
Ash	6	10	4
Sulfur	2	2	9
Volatile parts	4	2	2
Analytic moisture	2	2	6

Table 6. Wilcoxon test – values of *U*.

	(31, 34.2)	(31, 35)	(34.2, 35)
Combustion heat	75	81	13
Ash	7	56	79
Sulfur	81	81	49
Volatile parts	0	0	77
Analytic moisture	2	81	81

The critical area for test of series (significance level α=0.05) is range <2, 6> while for Wilcoxon test is the sum of ranges <0, 11> and <70, 81>. On the basis of the results of both test it can be said that only in case of the random variable describing ash contents for coal types 31 and 35 as well sulfur contents for coal types 34.2 and 35 the H_0 hypotheses of identical distributions can be accepted. In case of the random variable describing combustion heat for coal types 34.2 and 35 the test of series rejected the hypothesis of distributions equality while Wilcoxon test did oppositely. In other cases, both tests rejected H_0 hypotheses. In these three cases the median test was applied additionally with the same form of H_0 hypothesis (distributions are identical) [3,4]. To this purpose the division of researched material into particle size and particle density fractions was done. In this way 72-element samples were obtained [1]. For each considered coal pair and their properties the median value was determined from all results considering certain property.

To verify the hypothesis of equality between distributions the χ^2 test was applied

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(n_{ij}-np_{ij}\right)^2}{np_{ij}} \tag{5}$$

where: n_{1i} – number of results being bigger than the median value for i^{th} coal type (i=1, 2); n_{2i} – number of results being smaller than the median value for i^{th} coal type (i=1, 2); $n=n_1+n_2$ – number of results being part of both considered samples; $p_{ij} = p_i \cdot p_j = \frac{n_{i1}+n_{i2}}{n} \cdot \frac{n_{1j}+n_{2j}}{n}$.

The results of median test are presented in table 7.

Table 7. Median test – results of χ^2 test.

	median	χ^2
Ash (31, 35)	29.44	0.0007
Sulfur (34.2, 35)	0.52	0.0279
Combustion heat (34.2, 35)	5448	0.2500

The χ^2 test is used in median test where the number of degrees of freedom is equal to 1. The critical area for this test is a range $(\chi^2_\alpha, +\infty)$, where χ^2_α is value read from χ^2 distribution statistical tables for significance level α. For α=0.05 the value of $\chi^2_\alpha = 3.841$. Because in all three cases the inequality $\chi^2 < \chi^2_\alpha$ occurs then the hypothesis about equality between distributions can be accepted.

4. Conclusions

As it was shown application of traditional non-parametric statistical tests to determine whether there are significant differences between the analyzed coal types is not always efficient. This means that either these differences are not so big, either the selection of the coal properties was not ideal. However, after applying several tests to this aspect it occurred that eventually the differences with combustion heat, ash contents, sulfur contents and analytic moisture taken into consideration can be treated as statistically important.

In works [1, 5, 6, 7, 8], in which the observational tunnels method was applied to identify the coal type on the basis of the measurements of their properties was proved that the sufficient coal features to identify the coal type are moisture, volatile parts contents and sulfur contents. The methods being applied in this particular work also proved that such properties as moisture, sulfur contents and volatile parts contents significantly differentiate individual coal types.

Acknowledgment

The paper is an effect of the statutory project no. *11.11.100.276*.

References

[1] Niedoba T 2013 *Multidimensional characteristics of random variables in description of grained materials and their separation processes* (Kraków: Wydawnictwo Instytutu Gospodarki Surowcami Mineralnymi i Energią PAN) [in Polish]
[2] Krysicki W, Bartos J, Dyczka W, Królikowska K and Wasilewski M 2012 *Probability theory and mathematical statistics in examples 2* (Warszawa: PWN) [in Polish]
[3] Dobosz M 2001 *Statistical analysis of investigation results* (Warszawa: Akademicka Oficyna Wydawnicza EXIT) [in Polish]
[4] Stanisz A 2007 *Easy course of statistics 3* (Kraków: Statsoft) [in Polish]
[5] Jamróz D and Niedoba T 2014 *Physicochemical Problems of Mineral Processing* **50** pp 185-203
[6] Jamróz D and Niedoba T 2013 *Archives of Mining Sciences* **58(4)** pp 1317-31
[7] Niedoba T, Pięta P and Surowiak A 2016 *E3S Web of Conferences* **8 01036** pp 1-10
[8] Jamróz D 2017 *Information Visualization* **16(4)** pp 346-60