Proficiency assessment of laboratories undertaking hard coal analysis

Agnieszka Plis · Mariusz Mastalerz

Received: 30 January 2013 / Accepted: 19 July 2013 / Published online: 4 August 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract The Institute for Chemical Processing of Coal in Zabrze, Poland, organized a proficiency testing round to determine the mass fraction of ash, total sulfur and total carbon and the value of gross calorific value. Forty-four participants from Polish heat and power plants, research entities and independent laboratories reported results. The performance of the participants was evaluated using z and E_n scores. Over 85 % of the participants reported satisfactory results indicating the high level of proficiency of the Polish laboratories in the field.

Keywords Proficiency testing · Interlaboratory tests · Hard coal · Performance statistics

Introduction

The first proficiency tests organized by the Institute for Chemical Processing of Coal in Zabrze, Poland (IChPW), were conducted in 1999. The structure of the Institute has included a proficiency testing provider (PTP) since 2008. Since December 2010, the IChPW PTP possessed an implemented management system conforming to the requirements that are included in standard PN-EN ISO/IEC 17043:2010 [1], as confirmed by a certificate issued by the Polish Centre for Accreditation. The scope of accreditation comprises proficiency testing programs in the field of hard coal, brown coal, biomass for power purposes, coke from hard coal and solid combustion by-product analyses.

The organization of the proficiency test (PT) proceeded according to ISO 17043:2010 [1], and the statistical treatment of data was performed according to ISO 13528:2005 [2] and ISO 5725-2 [3] standards.

The PT was designed to assess the analytical performance of laboratories in charge of confirmatory analysis as well as screening analysis, since most of the laboratories belonging to the Polish heat and power plants.

This paper presents the organization of a PT performed in 2012 for hard coal analysis: the determination of the mass fraction of ash, $w(A)$, total sulfur, $w(S)$, and total carbon, $w(C)$, and the gross calorific value, Q .

Experimental

Preparation of hard coal samples

All of the hard coal samples were prepared in accordance with standard PN-90/G-04502 [4]. First the material was collected and adequately conditioned. Then, it was air-dried and grounded to particle size <2 mm before homogenization and sample dividing. The samples weighing 50 g were separated from the material and placed in boxes.

Determination of homogeneity and stability of PT

The homogeneity of the samples of hard coal was tested by analyzing the ash mass fraction, and parallel determinations from 15 subsamples were performed. There were performed the investigations of coal properties and developed method of selection and conditioning of this material, which was distributed to participants as a part of the PT.

Materials distributed in the PTs must be sufficiently stable over the period in which the assigned values are

A. Plis (✉) · M. Mastalerz
Institute for Chemical Processing of Coal, Zamkowa 1,
41-803 Zabrze, Poland
e-mail: aplis@ichpw.zabrze.pl; mmastalerz@ichpw.zabrze.pl
URL: www.ichpw.zabrze.pl

valid. The stability of the samples of hard coal was tested by analyzing the ash mass fraction (similarly as in the case of homogeneity), and parallel determinations from the 10 subsamples that remained after sending all of the samples to the participants were performed.

The sample homogeneity and stability were carried out by analyzing different measurands and in all cases; the results exhibited the homogeneity and stability. It is therefore concluded that the analysis of the ash mass fraction will be sufficient for evaluation of sample homogeneity and stability.

The computational methods used to assess the homogeneity and stability of the test objects complied with the ISO 13528 standard [2].

Proficiency testing

The organization of the PT in hard coal analysis proceeded according to ISO 17043:2010 [1]. At the beginning, all instructions and relevant information for participants were made available at the IChPW Web site. After registration,

in accordance with confidentiality rules, each participant that took part in the PT was provided a code. All of the results, tables and performance assessments presented in this paper are presented according to this code. In total, forty-four participants from Polish heat and power plants, research entities and independent laboratories were included in the testing round.

The samples were sent to the participants, who measured the mass fraction of ash, $w(A)$, total sulfur, $w(S)$, and total carbon, $w(C)$, and the gross calorific value, Q in triplicate. The values were reported to the PTP, along with the mean value of the results and the expanded uncertainty of the measurements (for $k = 2$). All of the participants calculated the uncertainties according to European cooperation guidelines for Accreditation EA-04/16 (2003) [5]. The results were calculated to a dry-weight basis and presented in this state. Additional information regarding the techniques and instruments was also asked. Participants carried out the determination of investigated measurands according to specified standards and their own procedures. The methods and description of their main principle are shown in Table 1.

After return of results by participants, the statistical treatment of data and the evaluation of the laboratory performances were performed according to the recommendations of the ISO 17043 [1] and ISO 13528 [2] standards.

Statistical analysis of results

The assigned values X of the measurands, the corresponding uncertainties u_X and standard deviations σ were determined based on the results obtained by the participants, using a robust method in accordance with the ISO 13528 standard [2]. The expanded uncertainty of assigned value U_X is calculated as $U_X = k \cdot u_X$, where k is coverage factor, and it is determined from the Student's t distribution corresponding to the appropriate associated degrees of freedom and % confidence for each measurands ($k = 2$).

To use performance statistics in the assessment of participant performance, the arithmetic means of the studied measurands must feature a normal distribution, as described in the ISO 13528 standard [2]. To analyze the normality of the distribution of the results, outliers and warning signals must be identified using the Grubbs test, in accordance with ISO 5725-2 standard [3]. These values cannot be considered in the subsequent calculations. The normality of the mean values of the data obtained from the participants was verified using the Shapiro–Wilk test. The value of the test probability p was used as the test results. A value of 0.05 was the upper limit of the test probability.

To select an appropriate performance statistic, one must determine whether the uncertainty of the assigned value is

Table 1 Description of methods applied to measure the investigated measurands

Measurand	Methods
Mass fraction of ash $w(A)$, g/(100 g)	ISO 1171:2010 [6]
	Thermogravimetric method, heating at a determined rate to a temperature of 815 °C in air
	PN-G-04560:1998 [7]
Mass fraction of sulfur $w(S)$, g/(100 g)	Thermogravimetric method, oxygen atmosphere, temperature of 600 or 815 °C
	PN-80/G-04512/Az1:2002 [8]
	Thermogravimetric method, complete combustion in air and calcination of residual ash at 815 °C
Gross calorific value Q , J/g	PN-G-04584:2001 [9]
	High-temperature method, oxygen atmosphere, temperature of 1350 °C, SO_2 detection by IR
	ISO 351:1996 [10]
Mass fraction of carbon $w(C)$, g/(100 g)	High-temperature method, oxygen atmosphere, temperature of 1350 °C, SO_x detection via titration
	PN-81/G-04513 [11]
	Calorimetric method
	ISO 1928:2009 [12]
	Calorimetric method: a different calculation method was used compared to standard PN-81/G-04513
	PN-G-04571:1998 [13]
	High-temperature method, oxygen atmosphere, temperature ≥ 950 °C, CO_2 detection by IR

lower than the critical value (0.3σ). The proficiency standard deviation σ was assumed to be equal to the robust standard deviation s^* , which was determined from the results obtained by the participants in the current round. It is important to determine whether the uncertainty of the assigned value is lower than the critical value: if these conditions are satisfied, the uncertainty value may be not considered in the performance statistic calculations due to its low value. In this case, to specify the participant's performance, the z score should be determined. If the condition is not satisfied, the assigned uncertainty must be considered in the calculation, and the z' score should be used.

In the current round of proficiency testing, the above conditions were satisfied; thus, the z score was used, according to the following equation:

$$z = \frac{x - X}{\sigma} \quad (1)$$

where x is the participant's result, X is the assigned value and σ is the robust standard deviation for proficiency assessment [2].

The test results of the participants was interpreted as follows:

$$\begin{aligned} |z| < 2 & \quad \text{satisfactory result} \\ 2 < |z| < 3 & \quad \text{questionable result} \\ |z| \geq 3 & \quad \text{unsatisfactory result} \end{aligned}$$

The performance statistics which accounts for the uncertainty of the participants' measurements are E_n , ζ (zeta) and E_z scores.

In the PT report, which was sent to the participants, the E_z score was used to evaluate the measurement uncertainty according to the applicable standard ISO 13528:2005 [2]. Other statistics, such as E_n and ζ (zeta) scores, may be used when the assigned value is not calculated using the results reported by the participants. In the current round, the assigned value was calculated using a robust method and was correlated with the results reported by the participants. But E_n and ζ (zeta) scores are usually used by National Metrology Institutes and widely accepted, so for the purposes of this paper, the E_n score was used to evaluate the measurement uncertainty [2], using the equation:

$$E_n = \frac{x - X}{\sqrt{U_x^2 + U_X^2}} \quad (2)$$

where x is the participant's result, X is the assigned value, U_x is expanded measurement uncertainty ($k = 2$) and U_X is expanded uncertainty of assigned value ($k = 2$). When no uncertainty was reported, it was set to zero ($U_x = 0$).

Using E_n score, it is possible to evaluate the reported "range" by the laboratory $x \pm U_x$ ($k = 2$) to the assigned

"range" $X \pm U_X$ ($k = 2$). A value of $|E_n| < 1$ provides objective evidence that the estimate of uncertainty is consistent with the definition of expanded uncertainty given in the GUM [14].

Results and discussion

Evaluation of homogeneity and stability of PT

In homogeneity test, the standard deviation between the unit samples was lower than the critical value (0.3σ), indicating that the sample was homogeneous.

In stability test, the difference between the results obtained at the beginning of the round and those obtained after 4 weeks was lower than the critical value (0.3σ), indicating that the samples were stable during the time frame of the analyses.

The proficiency standard deviation σ was assumed to be equal to the robust standard deviation s^* , which was determined from the results obtained by the participants in the previous round in the case of homogeneity test and in the current round in the case of stability test, in accordance with the ISO 13528 standard [2].

Statistical analysis of results

From the 44 laboratories having registered to the PT, all laboratories reported measurement results for $w(A)$ and Q , while two and four laboratories did not report results for $w(S)$ and $w(C)$, respectively. However, three laboratories (code number 25, 28, 37) having reported results did not submitted uncertainty budgets for any of the measurands investigated.

The specifications of the calculated assigned values X and the corresponding uncertainties U_X and relative standard deviations σ are presented in Table 2.

Distribution of the arithmetic means of the studied measurands was proven to be normal according to the Shapiro–Wilk test (probability p values ranged from 0.05 to 0.88). In the current round of proficiency testing, the z and E_n scores were used. The values of the z and E_n

Table 2 Assigned value (X), associated expanded uncertainty (U_X , $k = 2$) and relative standard deviation for PT assessment (σ) for investigated measurands

Measurand	X	U_X , $k = 2$	σ (%)
$w(A)$	3.69 g/(100 g)	0.02 g/(100 g)	1.68
$w(S)$	0.756 g/(100 g)	0.008 g/(100 g)	2.65
Q	31000 J/g	24 J/g	0.2
$w(C)$	78.16 g/(100 g)	0.30 g/(100 g)	0.98

Table 3 Specification of obtained z and E_n scores values

Participant code	Studied measurand							
	$w(A)$, g/(100 g)		$w(S)$, g/(100 g)		Q , J/g		$w(C)$, g/(100 g)	
	z	E_n	z	E_n	z	E_n	z	E_n
1	2.46 ^(q)	0.36	0.45	0.24	−0.32	−0.10	−0.59	−0.39
2	−0.01	0.00	−0.40	−0.19	1.42	0.54	−0.06	−0.04
3	0.58	1.16 ^(us)	0.32	0.15	−0.97	−0.56	0.18	0.12
4	−0.82	−0.32	4.03 ^(us)	0.50	−0.74	−0.15	1.52	0.45
5	−0.01	−0.01	−0.78	−0.26	−0.46	−0.14	−0.51	−0.32
6	−1.68	−0.79	−0.38	−0.19	2.00	1.12 ^(us)	0.44	0.48
7	−0.82	−0.12	−1.30	−0.49	1.23	0.28	−0.91	−0.64
8	−0.17	−0.05	0.30	0.02	0.49	0.35	−0.88	−0.51
9	0.69	0.16	–	–	−1.18	−0.41	–	–
10	0.63	0.38	1.82	1.17 ^(us)	−0.41	−0.26	1.08	1.65 ^(us)
11	0.80	0.48	–	–	1.29	0.79	–	–
12	−0.71	−0.19	−0.08	−0.04	0.12	0.03	0.70	0.31
13	0.42	0.25	−0.83	−0.54	−0.68	−0.26	−0.45	−0.43
14	−0.55	−0.33	−0.02	−0.01	0.00	0.00	0.13	0.06
15	−1.30	−0.87	−0.82	−0.27	−0.62	−0.46	−0.63	−0.22
16	−0.55	−0.53	−0.25	−0.14	−0.60	−0.24	0.14	0.07
17	−0.12	−0.04	−0.90	−0.40	−0.80	−0.23	−1.10	−0.54
18	0.69	0.11	−1.17	−0.26	−0.41	−0.05	−1.33	−0.61
19	1.49	1.25 ^(us)	1.47	0.58	4.48 ^(us)	1.56 ^(us)	0.60	0.16
20	−0.49	−0.08	−0.83	−0.25	−0.11	−0.02	0.21	0.07
21	−1.25	−0.29	1.48	0.38	−0.67	−0.07	0.26	0.05
22	−0.39	−0.12	−0.40	−0.16	−0.83	−0.37	−0.45	−0.34
23	2.78 ^(q)	0.90	0.68	0.27	−3.61 ^(us)	−0.15	2.34 ^(q)	0.22
24	−0.49	−0.17	−0.68	−0.34	−1.23	−0.36	0.57	0.38
25	−2.59 ^(q)	−6.92 ^(us)	−0.55	−1.45 ^(us)	1.05	2.77 ^(us)	2.66 ^(q)	6.74 ^(us)
26	−0.71	−0.18	−1.48	−0.73	−0.32	−0.13	−1.20	−0.89
27	−0.23	−0.13	−0.78	−0.34	0.87	0.18	0.34	0.18
28	−0.33	−0.90	0.67	1.76 ^(us)	1.06	2.81 ^(us)	−1.31	−3.32 ^(us)
29	2.41 ^(q)	0.55	−0.43	−0.20	−0.15	−0.05	0.46	0.36
30	−0.28	−0.08	0.45	0.24	0.59	0.20	0.77	0.46
31	−0.44	−0.12	−0.47	−0.30	−0.96	−0.65	−0.77	−1.00
32	0.42	0.31	−0.80	−0.75	−0.09	−0.02	0.59	0.43
33	0.20	0.04	−0.30	−0.10	−0.56	−0.11	−0.12	−0.05
34	2.89 ^(q)	1.74 ^(us)	0.78	0.51	0.96	0.32	2.22 ^(q)	1.22 ^(us)
35	−0.23	−0.07	0.47	0.17	−0.86	−0.31	−0.43	−0.21
36	0.47	0.12	0.23	0.07	0.74	0.31	−0.20	−0.04
37	2.09 ^(q)	5.55 ^(us)	−5.28 ^(us)	−13.94 ^(us)	0.49	1.29 ^(us)	–	–
38	−0.23	−0.22	0.38	0.19	−0.39	−0.10	−0.75	−0.40
39	1.01	0.61	1.18	0.58	0.17	0.07	−1.83	−1.65 ^(us)
40	−0.44	−0.33	0.37	0.10	−1.63	−0.74	−1.01	−0.63
41	−0.17	−0.11	0.96	0.78	−0.22	−0.04	0.90	1.00
42	−0.55	−0.17	1.47	1.37 ^(us)	2.20 ^(q)	0.37	–	–
43	−2.59 ^(q)	−1.57 ^(us)	−0.10	−0.05	0.63	0.17	−0.70	−0.55
44	1.06	0.89	0.63	0.31	0.98	0.49	2.71 ^(q)	1.11 ^(us)

^(q) Questionable result, ^(us) unsatisfactory result

scores, obtained by the participants for the determined measurands and evaluation of measurement uncertainty, are presented in the Table 3.

The laboratories' performances appear to be good for all measurands with a percentage of satisfactory results of z and E_n scores of above 85 % (Table 3). The percentage of

Fig. 1 Measurement results and uncertainties (vertical bars) of determined measurands: **a** $w(A)$, **b** $w(S)$, **c** Q , **d** $w(C)$. The solid line represents the assigned value X , the dashed lines delimit the assigned interval $X \pm U_X$ ($k = 2$) and the dotted lines delimit the target interval ($X \pm 2\sigma$)

Fig. 2 Uncertainty evaluation: u_x , measurement uncertainty, u_X , uncertainty of assigned value, calculated from expanded uncertainties for $k = 2$, σ , standard deviation for PT assessment

the satisfactory results of z score is 84, 95, 90 and 90 % for $w(A)$, $w(S)$, Q and $w(C)$, respectively. The percentage of the satisfactory results of E_n score is 86, 88, 89 and 85 % for $w(A)$, $w(S)$, Q and $w(C)$, respectively.

In seven cases, participants obtained questionable result of z score and unsatisfactory result of E_n score [for $w(A)$ and $w(C)$] and two participants obtained unsatisfactory results of z and E_n scores [for $w(S)$ and Q].

Figure 1 presents the mean values of the investigated measurands of hard coal x , the expanded measurement uncertainties U_X and the assigned values X , their expanded uncertainties U_X and standard deviation for PT σ . The solid line represents the assigned value X , the dashed lines delimit the assigned interval $X \pm U_X$ ($k = 2$), which is relevant for the E_n score evaluation and the dotted lines

Fig. 3 Methods applied to measure the investigated measurands, number of participants, which obtained satisfactory, questionable and unsatisfactory results of z score (proprietary procedure—single-laboratory validated method)

delimit the target interval ($X \pm 2\sigma$), which is relevant for the z score evaluation.

Based on Fig. 1, the qualitative measurands and estimated measurement uncertainties could be evaluated. If the measurands were determined correctly, their values belong to region defined as $X \pm 2\sigma$. In this case, the participants obtained satisfactory z score.

Additionally, in accordance with the ISO 13528 standard [2], the qualitative parameters and estimated measurement uncertainties were properly determined, if their value covered the area defined as $X \pm U_X$. It can be seen that most participants (above 85 %) carried out valid calculation of the expanded uncertainties of their results. There are also some cases, in which uncertainties are correct, but the values are too high (for example participant with code 8 in the case of sulfur mass fraction or 21 in the case of gross calorific heat value). Some participants estimated the measurement uncertainties correctly, but the mean values of investigated parameter are too high or too low (for example participant with code 34 in the case of ash mass fraction, 6, 23 in the case of gross calorific heat value or 26, 39, 44 in the case of carbon mass fraction).

The E_n score states if the laboratory result agrees with the assigned value within the respective uncertainty. It includes all the most important parameters such as

measurement result x , assigned value X , expanded uncertainty of the assigned value U_X ($k = 2$) and the measurement uncertainty U_x ($k = 2$).

An additional assessment was provided to each laboratory, based on the evaluation of reported measurement uncertainties, similar to what is implemented by the IMEP program [15]. Figure 2 presents for each measurand the number of participants having underestimated, overestimated or estimated correctly their measurement uncertainties.

The minimum and maximum acceptable uncertainties were set equal to the uncertainty of assigned value (u_X) and to the standard deviation for PT assessment (σ), respectively. The measurement uncertainty reported by a laboratory (u_x) was considered as:

- acceptable when ranging between u_X and σ ,
- underestimated when smaller than u_X or
- overestimated when larger than σ .

In the last two cases, laboratories were advised to review their uncertainty budget. More than half of the participants seem to have overestimated measurement uncertainties in the case of $w(A)$, $w(S)$ and Q . However, few laboratories having overestimated their measurement uncertainties got unsatisfactory E_n scores, probably due to a large analytical bias that needs to be investigated.

Specification of methods

Participants used the methods described in Polish and international standards, as well as proprietary testing procedures (proprietary procedure—single-laboratory validated method). The main methods and their description are shown in the Table 1.

Figure 3 presents methods applied to measure the investigated measurands, which was assessed using the z score and the number of participants, which obtained satisfactory, questionable and unsatisfactory results.

Most of the participants having reported satisfactory results applied the following methods for the determination of the various measurands:

- thermogravimetry according to ISO 1171:2010 for $w(A)$,
- SO_2 detection by infrared (IR) according to PN-G-04584:2001 for $w(S)$,
- calorimetry according to PN-81/G-04513 for Q , and
- CO_2 detection by IR according to PN-G-04571:1998 for $w(C)$.

Many other single-laboratory validated methods (proprietary procedure) were used for $w(C)$. The performance of laboratories having reported questionable or unsatisfactory results will be monitored in our next proficiency testing rounds, in order to follow the improvement of their measurement capability.

Summary

The Institute for Chemical Processing of Coal in Zabrze (Poland) organized in 2012 a PT exercise to determine the mass fraction of ash, total sulfur and total carbon and the value of gross calorific value. Over 85 % of the 44 participants from Polish heat and power plants, research institutes and independent laboratories reported satisfactory results. Several Polish and/or international standard methods were successfully applied. However, most of the participants reported overestimated measurement uncertainties that should be carefully reevaluated. Only few

laboratories were identified having a significant bias. Their analytical performance will be monitored in the next PT rounds, to ensure the improvement of the measurement capabilities of the Polish laboratories in the field.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. ISO/IEC 17043:2010 Conformity assessment—general requirements for proficiency testing
2. ISO 13528:2005 Statistical methods for use in proficiency testing by interlaboratory comparisons
3. ISO 5725-2:1994 Accuracy (trueness and precision) of measurement methods and results—Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method
4. PN-90/G-04502 Coal and lignite. Methods of sampling and preparation for laboratory testing (in Polish)
5. EA-04/16 rev00 December 2003 EA guidelines on the expression of uncertainty in quantitative testing
6. ISO 1171:2010 Solid mineral fuels—determination of ash
7. PN-G-04560:1998 Solid fuels. Determination of moisture, volatile matter and ash mass fraction using automatic analyzer (in Polish)
8. PN-80/G-04512/Az1:2002 Solid fuels. Determination of ash mass fraction using weight method (in Polish)
9. PN-G-04584:2001 Solid fuels. Determination of total sulfur and ash mass fraction using automatic analyzers (in Polish)
10. ISO 351:1996 Solid mineral fuels—determination of total sulfur—high temperature combustion method
11. PN-81/G-04513 Solid fuels. Determination of gross calorific value and calculation of net calorific value (in Polish)
12. ISO 1928:2009 Solid mineral fuels—determination of gross calorific value by the bomb calorimetric method, and calculation of net calorific value
13. PN-G-04571:1998 Solid fuels. Determination of carbon, hydrogen and nitrogen mass fraction using automatic analyzers. Macro method (in Polish)
14. JCGM 100:2008 Evaluation of measurement data—guide to the expression of uncertainty in measurement (GUM (1995) with minor corrections)
15. Fiammegkos I, Cordeiro F, Robouch P, Emteborg H, Charoud-Got J, Leys H, Kortsens B, de la Calle B IMEP-114: Determination of total Cd, Pb, As, Hg and Sn in feed premixes, Interlaboratory Comparison Report, December 2012