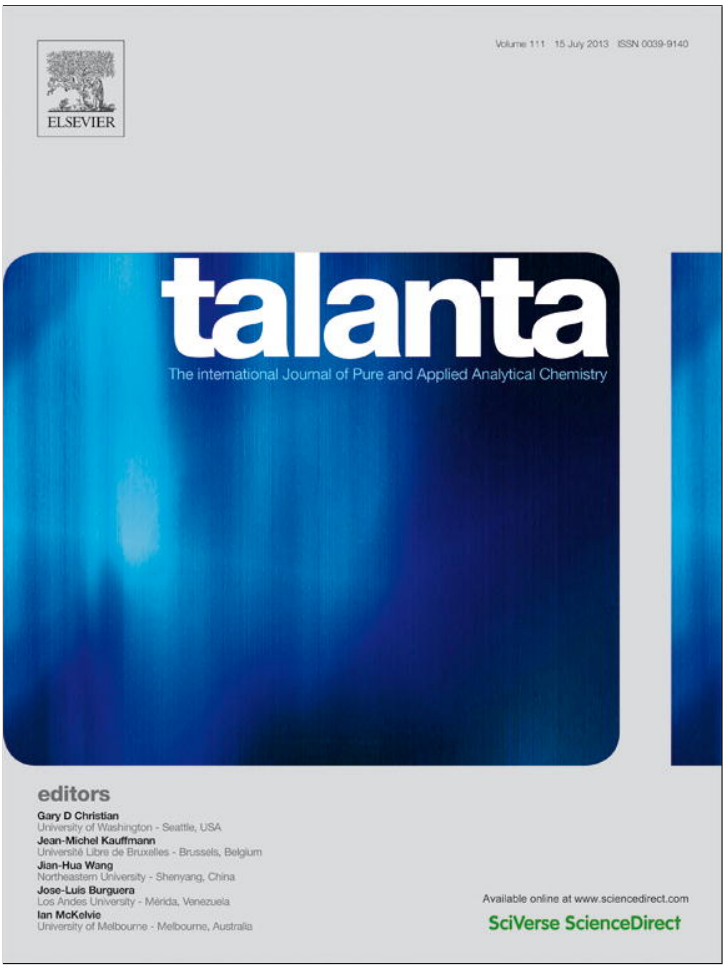

Talanta 111 (2013) 69–75

Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Quantification of uncertainty in mercury wastewater analysis at different concentration levels and using information from proficiency test with a limited number of participants

Marcelo T. Pereyra *, Adriana G. Lista, Beatriz S. Fernández Band

Analytical Chemistry Section, INQUISUR (UNS-CONICET), Department of Chemistry, Universidad Nacional del Sur, Av. Alem 1253, B8000CPB, Bahía Blanca, Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 3 October 2012
Received in revised form
15 February 2013
Accepted 17 February 2013
Available online 28 February 2013

Keywords:
Proficiency testing
Wastewater analysis
Mercury wastes
Measurement uncertainty
Metrological compatibility
Limited number of participants

ABSTRACT

Measurement uncertainty is one of the most required parameters of analytical quality in environmental decision-making. Several approaches have been reported for estimating uncertainty in analytical measurements. The use of proficiency test (PT) is an alternative, not only for assessing the performance of individual analytical laboratories, but also for estimating the concordance of an analytical method and their measurement uncertainties. Here we develop a PT scheme for a limited number of analytical laboratories that took part in the analysis of total mercury in samples with complex matrix (chlor-alkali process wastewater effluent) in absence of references of the highest metrological hierarchy (e.g. primary method, primary CRM). Two in-house reference materials (IHRMs) were prepared at different levels of analyte concentration and the homogeneity required was verified for the intended use. A set of parametric and robust statistical tests were applied to evaluate the assigned values of each IHRM. Metrological compatibility assessment of PT results, evaluation of individual and global performance of the laboratories, and determination of uncertainty of the analytical measurements were evaluated in this restricted study scenario. Between-laboratory differences were found not significant ($\alpha=0.05$). For both concentration levels we computed a relative standard uncertainty of 30% for the total Hg content.

1. Introduction

Most environmental management decisions are the result of analysis of information obtained from measurements made in both internal and external monitoring of the control routines. The correct interpretation of the analytical results of such measurements requires knowledge about their quality, especially with regard to measurement uncertainty [1]. Therefore, in order to achieve the required quality on these results, it is necessary to apply principles of quality assurance (QA) and quality control (QC) in the monitoring process to eliminate causes of unsatisfactory performance [2,3]. These quality control systems should include, among other requirements, verification of the results of traceability, or at least a quantitative assessment of bias and sustained in order to ensure comparability of results over time. In practice it is difficult to demonstrate the traceability of the environmental measurements in complex matrices, like soils, sediments, wastes water, sludges. A common practice in the analytical laboratories is the use of certified reference materials (CRM), pure substances or by using other documented standards to verify the traceability or at least the assessment of trueness [4]. However, when CRM are not available (absence on the market, difficulty in acquiring) and the use of pure calibrating substance is not appropriate because the matrix sample is complex, it is possible to evaluate the performance of analytical laboratories and calculate measurement uncertainties by using test schemes data [5]. There are not available in the market CRM of mercury in a complex matrix sample, such as the wastewater effluent of the chlor-alkali process. It should be noted that mercury is one of the heavy metals most currently found in the international, national and regional list of regulated pollutants [6–8], and is one of the pollutants discharged in the estuary of Bahia Blanca, more subject to official controls. Therefore, it is necessary to demonstrate the quality of the results of the determination of total mercury in the wastewater industrial chlor-alkali process and is useful to know the value of the measurement uncertainty of these analytical results to improve the scientific basis for management decisions. In this study, we carried out a PT scheme in order to investigate the technical competence of participating laboratories and the metrological comparability of their results. For this purpose, two in-house reference materials (IHRM) of industrial wastewater containing mercury were prepared by the Accredited Laboratory of Executive Technical Committee (CTE), government office for environmental protection in the area of the Bahía Blanca Petrochemical

* Corresponding author. Tel.: +54 291 4595100; fax: +54 291 4595160.
E-mail address: mpereyra@criba.edu.ar (M.T. Pereyra).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.02.045

complex [9]. The homogeneity of these IHRM was checked following the procedure described by Fearn and Thompson [10] and the International Harmonized Protocol [11]. Stability of these IHRM was also tested in spite of the fact that the analyses performed by the participants were started immediately after the sample preparation. The small number of test participants was other limiting condition [12,13]. Only six laboratories located in the region of Bahía Blanca, Argentina, are in conditions to carry out this determination as a routine task. There are not reported studies in the literature for such a restrictive scenario and evaluating more than one level of the analyte concentration in the samples. Based on these limited conditions, both materials were used as test samples in the PT scheme with five of the six existing laboratories, applying the same analytical method in experimental conditions of repeatability.

The objectives of this study included: evaluation of metrological compatibility of the assay results, evaluation of individual and global performance of the participating laboratories, and the determination of the measurement uncertainty using the information obtained from a PT scheme with the following scenario: (a) appropriate reference materials are not available, (b) the limited number of participants, (c) the analytical determination is not simple and, (d) materials with different level of concentration (split-level).

2. Experimental

2.1. Organisation of PT

The scheme was organised by the CTE of the government of Bahia Blanca, Argentina, in cooperation with private laboratories, research laboratories and researchers from the Universidad Nacional del Sur. The PT coordinator was responsible for: (a) designing the scheme, (b) sampling the test materials, (c) preparation and validation of the IHRM, (d) development and simultaneous distribution of instructions and IHRM to participating laboratories, and (e) collection and statistical analysis of the data obtained from the PT. The materials investigated were identical to those used in compliance with external and internal tests in order to be sure that the representativeness of each IHRM was enough.

This PT scheme was performed using the combination of criteria and guidelines proposed by the IUPAC/CITAC Guide [6], ISO Standard 5725 [14], AOAC Guidelines [15], ASTM Standard E 691 [16], taking into account lack of CRM (primary and secondary standard), the limited number of participants and the different levels of Hg in the test items.

Participating laboratories had to use the analytical official method consistent with their normal routine practice, i.e., methods were not adjusted for participate in this PT.

2.2. Description and collection of IHRMs

All participating laboratories in the PT had to analyse each of the IHRMs. Candidates IHRM were collected directly from two different chlor-alkali process effluents, containing Hg residues and they were divided into two groups; samples with low level concentration of Hg, and samples with high level concentration of Hg. All procedures for sampling and bottling were established according to international standards [17]. The PT coordinator established that all analyzes had to be carried out within 24 h because the stability in these types of samples is very important. Physicochemical characteristics of this complex material matrix are shown in Table 1.

Table 1
Wastewater characteristics.

Parameter	Units	Range	Number of analysis
pH		6–12	162
Conductivity	μScm^{-1}	5000–60,000	139
Turbidity	NTU	13–500	109
Temperature	°C	20–43	137
Total solids	$mg\,L^{-1}$	1600–30,000	109
Settable solids (10 min.)	$ml\,L^{-1}$	0.1–200	157
Settable solids (2 h)	$ml\,L^{-1}$	0.1–100	153
Chemical oxygen demand	$mg\,L^{-1}$	20–900	115
Total petroleum hydrocarbons	$mg\,L^{-1}$	0.1–30	36
Chloride	$mg\,L^{-1}$	1300	53
Total lead	$mg\,L^{-1}$	0.003–0.007	45
Total cadmium	$mg\,L^{-1}$	0.0006–0.007	47
Total mercury	$mg\,L^{-1}$	0.0001–0.0374	118

2.2.1. Sample 1 (LL)–IHRM

This sample is an industrial effluent discharge into the Bahía Blanca estuary. This effluent is subjected to continuous monitoring and audits to verify compliance with the maximum allowable limit of total mercury in $0.0050\,mg\,L^{-1}$; typical value is around $0.0015\,mg\,L^{-1}$. About 10 L of this material was collected in a suitable preconditioned polyethylene tank and immediately transported to the laboratory. The precondition step included an acid wash [filling the tank with 2% (v/v) nitric acid solution] and subsequent washings with the sample itself.

2.2.2. Sample 2 (HL)–IHRM

This material corresponds to an intermediate effluent of the same industrial process, with concentrations of total mercury in the level of $0.005\,mg\,L^{-1}$. The coordinator of PT collected a similar volume of this sample (10 L) applying the same Sample 1 preparation protocol, but now applied to Sample 2.

2.3. Preparation of the IHRMs

Preparation of samples for the participating laboratories is a fundamental and critical step of the PT scheme [18]. Thus, the homogenization and sub-sampling of Sample 1 and Sample 2 were treated strictly in order to ensure that all participating laboratories analyzed identical subsamples of these test samples. Five portions of 200 mL of Sample 1, with thorough mixing, were sequentially added to the bottles preconditioned, thus obtaining each final subsamples which were immediately sealed. Subsequently, Sample 2 was subsampled in the same way of Sample 1. Finally, these bottles of both IHRM were immediately labeled and randomly distributed among the participating laboratories, emphasizing the importance of the complete homogenization of each IHRM prior to the execution of the determination of total mercury. Each laboratory analyzed one sample of each level of analyte concentration.

2.4. Homogeneity of samples

To test for sufficient homogeneity of the IHRM, respect to total mercury content, a high analytical precision of $0.0001\,mg\,L^{-1}$ (LL) and $0.0002\,mg\,L^{-1}$ (HL) were necessary. For this purpose, three random subsamples of each IHRM were sent to an external laboratory after sealing the samples. The analyses were performed following the approach proposed by Fearn and Thompson [10] by triplicate instead of duplicate, applying the SM 3112B method [17].

Following the approach of these authors, three rapid tests were applied to calculate statistical homogeneity experimental parameters, which were compared with corresponding critical values, i.e.,: (a) Cochran's criterion procedure to check data homoscedasticity, outlier detection; (b) precision of the used analytical method, analytical standard deviation; (c) homogeneity test, between-sample standard deviation. Results obtained in the analysis of homogeneity data were summarized in Table 2. As the critical values were higher than the experimental values, both IHRM were considered sufficiently homogeneous.

2.5. Stability of samples

The stability of the total mercury content was tested to determine the suitability of Sample 1 and Sample 2 as reference materials in this study. Bottles were kept at 6 °C ± 2 °C over a period of 48 h and total mercury was determined at the beginning of the storage period and after 24 h and 48 h. Samples were analyzed by duplicate using the same procedures as for the homogeneity study. The recommended statistical test used was a *t*-test, conducted at the 95% confidence level on the measurement results [19].

The results obtained for both IHRM showed no significant effects ($\alpha \ll 0.05$) at 24 h and 48 h of storage period. On the basis of these results, it was concluded that no instability could be demonstrated, and the uncertainty contribution due to possible instability during the proficiency testing period was considered negligible.

2.6. Total Hg determination

Both samples were analyzed by triplicate by all participating laboratories by using standard method, SM 3112B [17], established in international and regional environmental regulations for monitoring wastewater quality [6,8]. Additionally, each participant had shown satisfactory results in proficiency tests already performed for the determination of mercury in natural waters using the same standardized method.

2.7. Data treatment

First, it conducted a series of statistical tests with the results to assess the outliers (outlying laboratories or outlying results for individual laboratories at each level). To carry out these assessments, we used the internationally recommended methodology: (a) the methodology of ISO 5725-5 was performed to assess consistency within-laboratory and between-laboratory and to establish the assigned value (best estimate of the true value of the measurand in the PT scheme) for each IHRM with their corresponding uncertainties [20,21]; (b) Cochran test was applied to check homogeneity of variances; (c) Grubbs' tests (single and pair) were done in order to investigate outlying laboratory averages.

Second, due to the limited population of laboratories, we applied robust statistics in order to contrast the results obtained from parametric statistics (modified *z*-scores, modified Huber test and MAD test) [22]. Finally, the proficiency assessment was carried out using the *z*-scores and generalized Youden-plots methodology designed for two non-uniform levels [23,24].

Table 2
Homogeneity of IHRMs. Summary of statistical results.

IHRM-Sample	Test	Experimental value	Critical value	Result
Sample 1-LL	Cochran	0.176	0.871	Pass
	Analytical standard deviation	0.137	0.167	Pass
	Homogeneity	0.018	0.110	Pass
Sample 2-HL	Cochran	0.382	0.871	Pass
	Analytical standard deviation	0.199	0.221	Pass
	Homogeneity	0.041	0.223	Pass

2.7.1. Establishing the assigned value (Xa) of IHRMs

There are several possibilities for the choice of assigned value Xa. Taking into account the lack of reference materials, absence of accredited test and a high complex matrix, which prevents the formulation of a quality reference material, the traceability of the total Hg determination in wastewaters is not commonly easy to check. The coordinator of PT should decide which way is most appropriate to resolve this problem. The value often used is the consensus value of the results reported by participating laboratories [25]. Then, assigned value Xa for both IHRM was established by evaluating the results of the participating laboratories using robust algorithm descript in ISO 5725-5 and arithmetic mean.

2.7.2. Estimating the uncertainty of IHRMs

The standard uncertainty of each IHRM was calculated as the standard deviation of each assigned value, taking into account the effect of the limited population of the PT participants. (N_p=6):

$$\mu_a = s_a \left[\sqrt{\frac{N_p - n}{n N_p}} \right] \quad (1)$$

where s_a is the standard deviation of the assigned value *Xa*, *n* is the number of participating laboratories.

2.7.3. Assessment of performance of an individual laboratory

For this purpose *z*-scores were calculated for each IHRM according to the following equation, after checking $\mu_a^2 < 0.1\rho_p^2$:

$$z = \frac{(x_i - Xa)}{\rho_p} \quad (2)$$

where x_i is the laboratory experimental reported value; *Xa* is the assigned value by the PT coordinator and ρ_p is the fitness for purpose-based "target value for the acceptable deviation from the assigned value" [25]. Acceptable limits for *z*-scores are:

- *z*-score = 0. Perfect result.
- *z*-score ≤ ±2. Satisfactory or acceptable result. No action is required.
- ±2 < *z*-score ≤ ±3. Questionable or doubtful result. Preventive action is required.
- *z*-score > ±3. Unsatisfactory or unacceptable result. Corrective action is required.

To select ρ_p the PT coordinator used three recognized sources of information about interlaboratory precision for this specific analytical determination (Table 3).

As can be seen in the table, Horwitz equation gives similar values to EPA Method 245.1, then for Sample 1 – IHRM (LL, 1.5 μg L^{-1}) was selected RSD = 0.45 (ρ_{LL} = 0.7 μg L^{-1}), and for Sample 2 – IHRM (HL, 5.5 μg L^{-1}) was selected RSD = 0.44 (ρ_{HL} = 1.5 μg L^{-1}) for *z*-score calculations.

2.7.4. Assessment of metrological compatibility of results

Metrological compatibility was considered satisfactory if the bias exceeds μ_a only by a value which is insignificant in comparison with random interlaboratory errors, i.e., the chosen null

Table 3
Consulted sources and data for selection ρ_p values.

Source	Concentration level ($\mu g\ L^{-1}$)	Relative standard deviation (%)	References
Standard method 3112B	0.34 4.2	22.6 13.3	[17]
EPA method 245.1	0.60 3.4	55 44	[26]
Horwitz SD	0.6 1.5 3.5	49 45 38	[27,28]

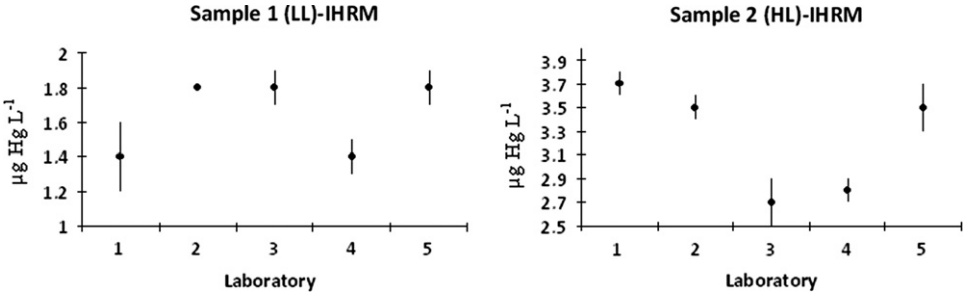

Fig. 1. Participants' results of Sample 1 (LL) and Sample 2 (HL) IHRMs.

hypothesis was stated as:

$$H_0 = |x_i - Xa| \le [\mu_a^2 + (0.3\rho_p)^2]^{0.5} = \text{critical bias} \quad (3)$$

All individual results, x_i, were checked against the critical values for bias, calculated according the IUPAC/CITAC Guide [6].

2.7.5. Youden plot

In the present PT, IHRMs were not equal in their total Hg concentrations, so assumptions underlying in the original Youden's article [29] are not obeyed. In these cases it is more appropriate to apply the procedure proposed by Mandel and Lashof [23]. Confidence ellipse region ($\alpha=0.05$) for the two-sample plot is expressed in terms of Hotelling's T^2 distribution [24]:

$$[x_1 - X_{a(\text{Sample1}-\text{LL})}]^2 - 2r_s[x_1 - X_{a(\text{Sample1}-\text{LL})}][x_2 - X_{a(\text{Sample2}-\text{HL})}] + [x_2 - X_{a(\text{Sample2}-\text{HL})}]^2 = (1-r_s^2)T^2 \quad (4)$$

where r_s is the Spearman rank correlation coefficient, which was selected to be less sensitive to lack of data normality, and T^2 is distributed as:

$$\frac{2(n-1)}{(n-2)} F_{(0.05;2,n-2)} = 25.47$$

2.7.6. Assessment of expanded uncertainty for an individual laboratory

Calculations of expanded uncertainties U_i (coverage factor, $k=2$) for any participating laboratory at each level, was based on the recommended methodology proposed in the Nordtest Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories [30]. To carry out these estimations, were used the repeatability information reported by each participant in terms of standard uncertainties μ_i and contributions for the bias and the corresponding IHRM, μ_a, i.e.:

$$U_i = k \times \sqrt{\mu_i^2 + \mu_{bias}^2} = 2 \times \sqrt{\mu_i^2 + bias_i^2 + \mu_a^2} \quad (5)$$

where: $bias_i = (x_i - X_a)$.

Relative expanded uncertainties, U_i/x_i, were also calculated in order to study between-material variability.

3. Results and discussion

3.1. Individual results

The experimental values reported by all participating laboratories (83% of population) are summarized in graph form (Fig. 1), expressed as $x_i \pm s_i$ ($\mu g\ Hg\ L^{-1}$). Within-laboratory variability expressed by repeatability relative standard deviation (RSDr) was acceptable; only one laboratory reported a value of RSDr higher than 10% in Sample 1 (LL)–IHRM.

3.2. Evaluation of results obtained for each IHRM

3.2.1. Consistency study of data.

The ISO 5725-2 and ASTM E 691 standards describe useful graphical techniques to visually assess the consistency of results and both organizations recommend Mandel's k and h statistics. These statistics can also be used to evaluate the quality of laboratories in laboratory-performance studies [21]. Mandel's k and h statistics indicate the within-laboratory and the between-laboratory consistency respectively, and are calculated using the following equations:

$$k_i = \frac{S_i}{\sqrt{\sum_{i=1}^{n} S_i^2/n}} \quad \text{and} \quad h_i = \frac{x_i - \overline{x}}{\sqrt{1/(n-1)}\left(\sum_{i=1}^{n}(x_i-\overline{x})^2\right)}$$

where: $\overline{x}$ is the global arithmetic mean of the individual laboratory averages, x_i.

Fig. 2 shows the k and h results with their corresponding critical values at the significance levels of 1% and 5%. These critical values are determined according ISO or ASTM standards listed above.

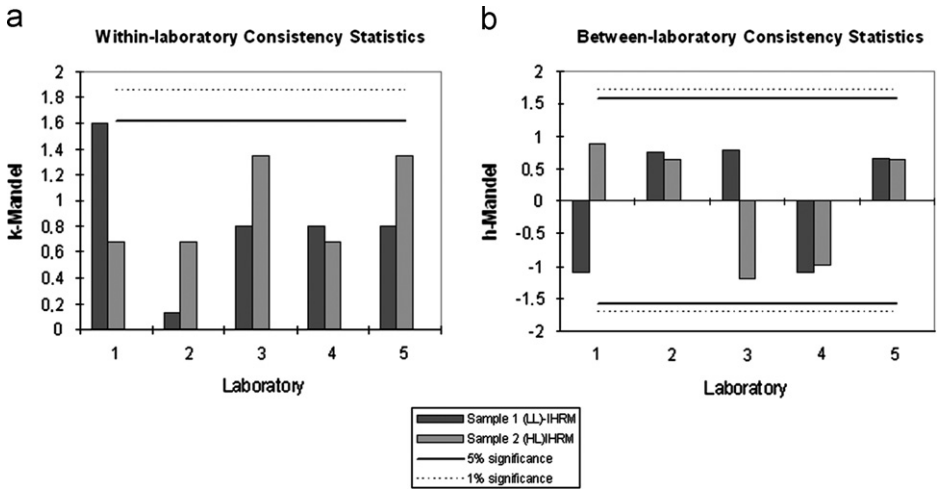

Fig. 2. Mandel's k (a) and h (b) statistics of each participant for each IHRM.

Table 4
Summary of obtained results with parametric and robust statistics.

IHRM	ID Lab (i=)	Parametric statistics			Robust statistics		
		Grubbs single	Grubbs doble	Cochran	Modified z-scores	Modified Huber test	Modified MAD test
Sample 1 (LL)	1	1.09	0.0022	0.57	0.69	1.03	1.20
	2	0.75	—	0.03	0.54	0.80	0.90
	3	0.79	—	0.14	0.57	0.84	1.00
	4	1.09	0.0022	0.14	0.69	1.03	1.20
	5	0.66	—	0.14	0.48	0.71	0.80
	Critical value:	**1.76**	**0.0018**	**0.79**	**3.50**	**1.50**	**5.00**
	Rejected values:	0	0	0	0	0	0
Sample 2 (HL)	1	0.88	—	0.09	0.63	0.94	1.1
	2	0.65	—	0.09	0.48	0.71	0.8
	3	1.20	0.0089	0.36	0.78	1.15	1.3
	4	0.97	0.0089	0.09	0.62	0.92	1.1
	5	0.65	—	0.09	0.48	0.71	0.8
	Critical value:	**1.76**	**0.0018**	**0.79**	**3.50**	**1.50**	**5.00**
	Rejected values:	0	0	0	0	0	0

From Mandel's plots it can be observed that all data were consistent. In k graph can be look that Laboratory No. 1 has large k value, but only in Sample 1 (LL)-IHRM, indicating potential within-laboratory imprecision at this low level of Hg concentration. Moreover, both Laboratory No. 2 k values were the lowest, representing the best within-laboratory precision. On the other hand, there is a general pattern in h graph indicating that most laboratories (80%) tend to report higher results, however between-laboratory variability was more consistent than within-laboratory variability ($\alpha=0.20$ versus $\alpha \cong 0.05$).

3.2.2. Homogeneity of variances and outliers assessments

Results of the parametric test as well as the robust statistics are summarized in Table 4.

In this context, outliers should be eliminated at the 1% level unless there is a strong reason to keep it in the study [14,20]. Then, critical values were calculated at the significance levels of 1% and no rejected values were found, as can be seen in Table 4.

3.2.3. Assigned value and associated uncertainty

Null hypothesis of Shapiro–Wilk test [31] was not rejected ($\alpha=0.05$) so data normality can be assumed. Homogeneity of variances was also verified. Then IHRM assigned value X_a, and s_a^2 could be calculated as the average of results for Sample 1 (LL) and Sample 2 (HL) and as the reproducibility variance (sum of the repeatability variance and the between-laboratory rowsep="1" variance). However, mean and standard deviation from robust algorithms A and S included in ISO 5725-5, were also calculated in order to corroborate these approaches. Standard uncertainties were calculated according Eq. (1) using both standard deviation and robust standard deviation. These results are shown in Table 5.

Table 5
Results of parametric and robust estimators and calculated confidence intervals for each sample (μg Hg L^{-1}).

IHRM	Average	Robust mean	Standard deviation	Robust std. dev	μ	μ robust	$Xa \pm Ua$
Sample 1 (LL)	1.6	1.8	0.3	0.2	0.04	0.04	1.6 ± 0.1
Sample 2 (HL)	3.2	3.4	0.4	0.3	0.08	0.08	3.2 ± 0.2

As can be seen, standard uncertainties values obtained through parametric statistic were identical to those obtained by robust statistic. Moreover, means and standard deviation were very similar from both methodologies. For these reasons, average was selected as assigned value X_a for each IHRM and their expanded uncertainties U_a ($k=2$) were 0.1 and 0.2 μg Hg L^{-1}, respectively, for Sample1 (LL) and Sample 2 (HL).

3.2.4. Metrological compatibility assessment

According to Eq. (3) critical biases were: 0.21 μg Hg L^{-1} and 0.47 μg Hg L^{-1} for each IHRM level, respectively. Then, as shown in Fig. 1, no x_i results exceeded the maximum allowable biases and H_0 was not rejected ($\alpha=0.05$). Therefore, the metrological compatibility of the results for any IHRM was demonstrated.

3.3. Laboratory performance studies

3.3.1. Comparison of z-scores

Statistical *z*-scores were calculated by using Eq. (2) and they are reported in Fig. 3. As shown in the figure all values reported by participating laboratories were satisfactory ($z < \pm 0.5$) and the performance was similar for each IHRM. Moreover, it seems that 60% of participant would tend to overestimate the results. Additionally, the sum of squared scores (SSZ) [24] for each laboratory was lower than the critical value ($\chi^2=0$), then overall performance of the laboratories was also very good.

3.3.2. Youden plots

Confidence elliptical region (Fig. 4) calculated with Eq. (4) shows that the results of the five participating laboratories are included within this graph. Therefore, as all laboratories applied the same method, it was decided that between-laboratory differences were probably not significant in the PT ($\alpha=0.05$).

3.3.3. Assessment of expanded uncertainty for an individual laboratory

According to Eq. (5), both expanded uncertainty and relative expanded uncertainty were calculated for each participating laboratory. Table 6 summarizes these results and the individual 95% confidence interval for each laboratory and each IHRM, $x_i \pm U_i$.

Additionally, there are no evidences of statistically significant differences between all values of U_i/x_i when a two-factor ANOVA test was carried out, the $H_{0(laboratory)}$ and $H_{0(IHRM)}$ were not rejected ($\alpha=0.05$). This is a significant finding, since all laboratories applied the same analytical method, and relative expanded uncertainties were uniform regardless of the concentration level of the IHRM. Then it can consider a relative expanded uncertainty, as:

$$\frac{\overline{U}(x)}{x} = \sqrt{\frac{\sum_{i=1}^{10}(U_i/x_i)}{10}} = 0.3$$

This value is equal to that obtained using the information of reproducibility standard deviation calculated for each IHRM, i.e.,: from Table 5, root of the mean square of relative standard deviation were calculated as:

$$\frac{\overline{U}(x)}{x} = 2 \times \sqrt{\frac{\left[(s_{R(LL)}/x_{LL})^2 + (s_{R(HL)}/x_{HL})^2\right]}{2}} = 0.3$$

These two approaches provided identical values of relative expanded uncertainty ($\alpha < 0.05$).

All these affirmation could indicate rectangular distribution for $U-(x)/x$ versus x.

4. Conclusion

- A PT was developed, organized and conducted, in the field of wastewater routine complex matrix for a limited number of participants applying a non simple analytical method. The interest in this PT was very high taking into consideration the lack of availability of CRM (primary or secondary measurement standard) and the very low number of analytical laboratories able to participate in a PT scheme.
- Two test items were prepared from real wastewaters with different concentration levels and they were studied as certified in-house reference materials. Both IHRM demonstrated a proper homogeneity, stability and suitability for use as quality control material in proficiency tests.
- 83% of laboratory population took part in the performance study and reported analytical measurement results for the total mercury content in each item, all using the international official method, SM 3112B. Individual proficiency assessment was satisfactory for every laboratory, and a successful consistency rate was also achieved (100%).
- Performed statistical evaluations of the PT also provided useful information for calculating uncertainty associated to assigned values for each IHRM, and uncertainty associated to individual analytical results of participants. Relative expanded uncertainty was therefore statistically equal for both source of variation: laboratories and IHRM, this is useful information for decision making that should be investigated further.
- Taking into account a rectangular distribution for the $U-(x)/x$, the quantification of the expanded uncertainty at any

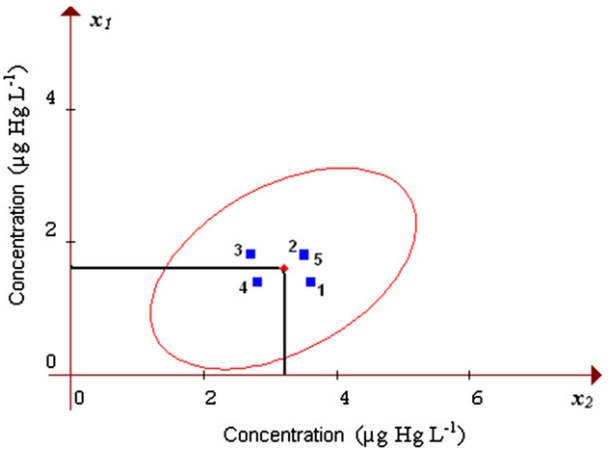

Fig. 4. Confidence ellipse region (95%). x_1 and x_2 are the concentrations of Sample 1 (LL) IHRM and Sample 2 (HL) IHRM, respectively.

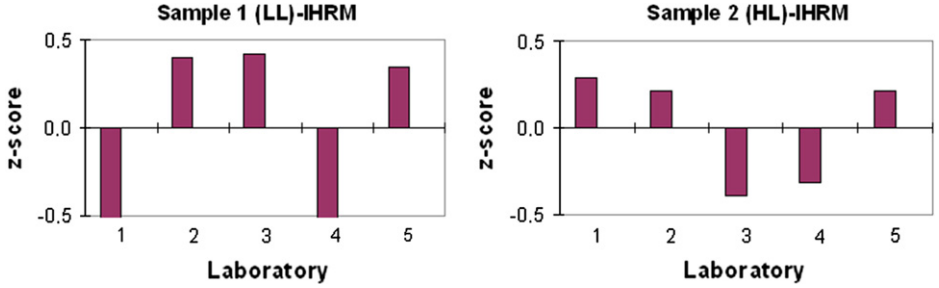

Fig. 3. *z*-score values of each participant for each IHRM.

Table 6
Expanded uncertainties for each laboratory.

IHRM	$(\text{Bias})^2$ $(\mu g\ L^{-1})^2$	$\mu_a^2(\mu g\ L^{-1})^2$	$\mu_i^2(\mu g\ L^{-1})^2$	$U_i(\mu g\ L^{-1})$	$x_i \pm U_i(\mu g\ L^{-1})$	U_i/x_i
Sample 1 (LL)	0.040	0.002	0.040	0.6	1.4 ± 0.6	0.4
	0.048	0.002	0.000	0.4	1.8 ± 0.4	0.2
	0.053	0.002	0.010	0.5	1.8 ± 0.5	0.3
	0.040	0.002	0.010	0.5	1.4 ± 0.5	0.3
	0.040	0.002	0.010	0.5	1.8 ± 0.5	0.3
Sample 2 (HL)	0.160	0.006	0.010	0.8	3.6 ± 0.8	0.2
	0.090	0.006	0.010	0.7	3.5 ± 0.7	0.2
	0.250	0.006	0.040	1.1	2.7 ± 1.1	0.4
	0.160	0.006	0.010	0.8	2.8 ± 0.8	0.3
	0.090	0.006	0.040	0.7	3.5 ± 0.7	0.2

concentration level, within the study range, is very simple. It can be calculated, only by multiplying the obtained concentration value by the calculated relative expanded uncertainty value.

Acknowledgement

The authors gratefully acknowledge participating laboratories for their contribution, and also thank to Universidad Nacional del Sur and Comité Tecnico Ejecutivo, Bahía Blanca, Argentina for the Financial support. Dra. Fernández Band acknowledges to Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

References

[1] P. Dehouck, Y. Vander Heyden, J. Smeyers-Verbeke, D.L. Massart, J. Crommen, Ph. Hubert, R.D. Marini, O.S.N.M. Smeets, G. Decristoforo, W. Van de Wauw, J. De Beer, M.G. Quaglia, C. Stella, J.-L. Veuthey, O. Estevenon, A. Van Schepdael, E. Roets, J. Hoogmartens, Anal. Chim. Acta 481 (2003) 261–272.
[2] Ph. Quevauviller, Quality Assurance in Environmental Monitoring: Sampling and Sample Pretreatment, John Wiley & Sons Inc., Weinheim, Germany, 2008, pp. 1–24.
[3] Ph. Quevauviller, O.F.X. Donard, Trends Anal. Chem 20 (2001) 600–613.
[4] Ph. Quevauviller, Trends Anal. Chem 23 (2004) 171–177.
[5] R. Companyó, R. Rubio, A. Sahuquillo, R. Boqué, A. Maroto, J. Riu, Anal. Bioanal. Chem. 392 (2008) 1497–1505.
[6] U. S. Code of Federal Regulations, Title 40, Part 415, Subpart F, Section 415.66, U.S. Government Printing Office, Washington, DC, 2012.
[7] European Commission, Directive 84/156/EEC of 8 March 1984, Off. J. European Communities L74 (1984), 49–54.
[8] Resolución No 336/03 de la Autoridad del Agua, Boletín Oficial de la Provincia de Buenos Aires No 24835, La Plata, Argentina, 2003.
[9] Ley No 12530, Programa Especial para la Preservación y Optimización de la Calidad Ambiental, Congreso de la Provincia de Buenos Aires, Boletín Oficial de la Provincia de Buenos Aires No 24176 del 16 de noviembre de 2000, La Plata, Argentina, 2000.
[10] T. Fearn, M. Thompson, Analyst 126 (2001) 1414–1417.
[11] M. Thompson, S.L.R. Ellison, R. Wood, Pure Appl. Chem. 78 (2006) 145–196.
[12] I. Kuselman, A. Fajgelj, Pure Appl. Chem. 82 (2010) 1099–1135.
[13] I. Pankratov, S. Elhanany, S. Henig, S. Zaritsky, I. Ostapenko, I. Kuselman, Accredit. Qual. Assur. 15 (2010) 459–466.
[14] ISO, Accuracy (trueness and precision) of measurement methods and results. Part 5: Alternative methods for the determination of the precision of a standard measurement method. ISO 5725-5:1998 (E), International Organization for Standardization, Geneva, Switzerland, 1998, pp. 1–56.
[15] Association of Official Analytical Chemist (AOAC), (1995), J. AOAC Int. 78 143A–160A.
[16] ASTM, Standard Practice for Conducting an Interlaboratory Study to Determinate the Precision of a Test Method. E691-99, Annual Book of ASTM Standards, vol. 14.02. American Society for Testing and Materials, West Conshohocken, PA, USA 2004, pp. 1–22.
[17] APHA-AWWA-WPCF, Standard Methods for the Examination of Water and Wastewater, 20th ed., American Public Health Association (APHA), American Water Works Association (AWWA) and Water Pollution Control Federation (WPCF) Washington, DC, 1999.
[18] F. Raposo, M.A. de la Rubia, R. Borja, M. Alaiz, J. Beltrán, C. Cavinato, M. Clincksporr, G. Demirer, E. Diamadopoulos, B. Helmreich, P. Jenicek, N. Artí, R. Mendéz, J. Noguerol, F. Pereira, S. Picard, M. Torrijos, Talanta 80 (2009) 329–337.
[19] B. Brookman, R. Walker, Guidelines for the In-House Production of Reference Materials, Version 2, LGCVAM1998040, LGC Publisher, Middlesex, UK, 1998, pp. 22–25.
[20] E. Hund, D.L. Massart, J. Smeyers-Verbeke, Anal. Chim. Acta 423 (2000) 145–165.
[21] Y. Vander Heyden, J. Smeyers-Verbeke, J. Chromatogr. A 1158 (2007) 158–167.
[22] J.N. Miller, J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, sixth ed., Pearson Education Limited, Gosport, UK, 2010, pp. 175–179.
[23] J. Mandel, T.W. Lashof, J. Qual. Technol. 6 (1974) 22–36.
[24] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part A, Elsevier, Amsterdam, NL, 1997, pp. 452–454.
[25] M. Thompson, Ph.J. Lowthian, Notes on Statistics and Data Quality for Analytical Chemists, Imperial College Press, London, UK, 2011, pp. 215–218.
[26] USEPA, Determination of Mercury in Water by Cold Vapor Atomic Absorption Spectrometry, EPA Method 245.1, 3rd rev., United States Environmental Protection Agency, Ohio, PA 45268, 1997.
[27] K.W. Boyer, W. Horwitz, R. Albert, Anal. Chem. 57 (1985) 454–459.
[28] R. Albert, W. Horwitz, Anal. Chem. 79 (1997) 789–790.
[29] W.J. Youden, Int. Qual. Control 15 (1959) 24–28.
[30] B. Magnusson, T. Näykki, H. Hovind, M. Krysell, Handbook for calculation of measurement uncertainty in environmental laboratories, second ed., Nordtest Technical Report 537, Espoo, Finland, 2004.
[31] S.S. Shapiro, M.B. Wilk, Biometrika 52 (1965) 591–611.