Microchemical Journal 119 (2015) 145–151

Contents lists available at ScienceDirect

Microchemical Journal

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

Measurement uncertainty estimated from accuracy assessment by using an in-house reference material: two case studies for the extraction of total PAH in industrial wastewater sludges

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

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

ARTICLE INFO

Article history:
Received 11 September 2014
Accepted 2 December 2014
Available online 9 December 2014

Keywords:
Measurement uncertainty
In-house reference material
Total PAH
Extraction
Industrial wastewater sludge

ABSTRACT

Measurement uncertainty is a required parameter to assure the reliability of analytical results because many management decisions (scientific, technical, regulatory, and economical) are based on the results of analytical processes. Moreover, since ISO 17025 accreditation requirements specify protocols for estimating uncertainty in measurements, the community of the analytical laboratories had to start reporting their results with the corresponding associated uncertainty. Several approaches have been reported for estimating uncertainty, although the evaluation of uncertainty from the method accuracy assessment has the advantage to not require extra efforts for the analysts. In-house reference materials are useful both to perform the validation studies and to estimate the measurement uncertainty when certified reference materials are not available. In this study, we presented a practical methodology to calculate the uncertainty of the analytical results when the accuracy assessment is performed by using an in-house reference material developed to fit this purpose. For evaluating the suitability of this approach, we applied it to the assessment of two extraction methods for total PAH determination in industrial sludges. The uncertainty budgets revealed that the uncertainty arising from subsampling gives decisive contributions to the overall uncertainty (from 18% to 69%), depending on the homogeneity in the sample preparation process.

1. Introduction

Hazardous waste management is a predominant issue in environmental science research and policy. The physicochemical properties of sludge wastes must be investigated in origin before their safe disposal in order to determine their compliance with international, national or regional regulations [1–3]. These current regulations require analytical controls and periodic monitoring for some hazardous pollutants. Besides polychlorinated biphenyls and chlorobenzenes, PAH are well known class of priority hazardous organic pollutants [4,5]. In this regard, the total PAH concentration in sewage sludge was recently limited in many Member States of the European Community [6]. Therefore, the compliance of these regulations requires the availability of reliable data for decision-making and thus adequate analytical methods with well-established uncertainty estimates have to be applied for testing laboratories. The evaluation of the uncertainty associated with the overall measurement process should be estimated and reported in order to verify the "fit for purpose" of an analytical method and demonstrate the quality of their results. This not only supports the reliability of the analytical performance of the laboratories, but also prevents the data end-users from drawing misleading conclusions concerning the environmental risk and potential need for remediation or waste disposal.

Over the years, several spectroscopic and non-spectroscopic methods have been developed and standardized for the analysis of total PAH in soil, sediments and waste materials. Most of these methods need two pretreatment steps prior to the analytical determination: subsampling/splitting of the laboratory sample with a consecutive processing step. Sub-sampling is associated with the homogenization and splitting of the analytical portion from the laboratory sample, and sample processing involves extraction techniques followed by a clean-up procedure. These pretreatment steps have associated their contribution of uncertainty to the overall measurement uncertainty of the analytical process. Different approaches to perform the evaluation of the measurement uncertainty are described in the literature [7–10]. The approach proposed by Maroto et al. takes advantage of the information generated in assessing the accuracy of a given analytical procedure [10]. A common practice in the analytical laboratories is the use of certified reference material (CRM), pure substances or by using other documented standards to evaluate the accuracy of their results. 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 accuracy and measurement

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

http://dx.doi.org/10.1016/j.microc.2014.12.001
0026-265X/

uncertainty by using a suitable reference material produced in the laboratory, usually denominated "in house reference materials", IHRMs. Scarce works have been reported in the literature concerning the production of IHRMs. The National Measurement Office of the UK Government published an official document that provides a guidance to prepare IHRMs. However, the author remarks that this document does not provide a full set of instructions [11].

The objectives of this study included: (a) development of an IHRM for total PAH analysis in wastes, using the authentic matrix of routine samples, and assessing the assigned value and its associated uncertainty taking into account the preparation procedure, and the contributions of the homogeneity and the stability to the overall uncertainty of the IHRM, (b) evaluation of the suitability of the IHRM for assessing both the accuracy and the measurement uncertainty of the results, obtained both with a standard reference method as well as with a new developed extraction method, (c) evaluation of the uncertainty budget of the total PAH analysis in waste samples taking into account the analytical procedure, the bias component and the pretreatment steps, and (d) comparison of the uncertainty budget for both studied extraction methods.

2. Experimental

2.1. Chemicals

All chemicals used were of analytical grade or better. The standard solution of chrysene (192.8 μg ml^{-1} in methylene chloride) was supplied by Supelco (Bellefonte, PA, USA). Polyoxyethylene 10 lauryl ether (POLE, purity $p_{POLE} \geq 99\%$) was supplied by Sigma (St. Louis, MO). Chrysene (purity $p_c > 95\%$ AR grade) was obtained from BDH Ltd. (Poole, UK). Triton® X-100 (purity $p_T \geq 99\%$) and SDS (purity $p_{SDS} \geq 95\%$) were purchased from Merck (Darmstadt, Germany) and Anedra (Tigre, Argentina), respectively. For spectrofluorimetric analysis, hexane, acetone and methanol of spectrometric grade and water purified with a Milli-Q system (Millipore, Bedford, MA) were used throughout.

2.2. Samples

In order to ensure the representativeness of the matrix routine samples, authentic wastewater sludge was collected from the effluent channel in the Petrochemical Complex, Bahía Blanca, Argentina, with a pre-conditioned stainless steel sampler, screened to remove foreign objects and placed into a dark polypropylene container previously washed with methanol. Wet sample was immediately transferred to the research laboratory to sieve and mix for preliminary homogenization, and then allowed to settle to remove most of the occluded water. Wet settled sample was preserved at 4 °C in the dark according to the international protocols (ASTM and EPA).

2.3. In house reference material production

2.3.1. Preparation procedure

The authentic sludge sample was dried at room temperature during 2 hours, homogenized again, and then dried during 12 hours at 80 °C [12,13]. This preconditioned authentic material (PAM) was finally homogenized at room temperature and then stored in the dark at constant-humidity. Later, several sub-samples of the PAM were used to prepare the in-house reference material (IHRM) following this process: a) Two sub-samples of the PAM were analyzed in an ISO 17025-accredited laboratory for the determination of the native concentration of PAH by applying the EPA Method 3540 C (Soxhlet extraction) [14] and the EPA Method 8310 (HPLC-UV) [15]. Certificates of such analyses reported that 16 priority PAH were not detected (detection limit, 0.1 mg kg^{-1}) in both sub-samples of the PAM; b) another sub-sample of the PAM ($m_{PAM} = 121.0011$ g) was fortified with chrysene (analytical-grade) and aged in a stationary mode for 1 month in order to allow the analyte–matrix interactions [16]. Chrysene was selected as the target analyte because of its known environmental and analytical properties of concern and because is the target analyte used for calibration in the standard reference method and research works [12,17–20]. This fortification procedure was carried out according to the technique described by Llompart *et al* [21], i.e.: this sub-sample of the PAM was slurried with 200 mL of a hexane-acetone homogeneous solution (1:1) of chrysene (0.96×10^{-4} M) and then allowed to air-dry with occasional stirring at ambient temperature, protected from draught for 5 days, bottled and stored in a dry dark place at 4 °C. The mass of chrysene ($m_{chrysene} = 0.0044$ g) used for the preparation of the hexane-acetone solution and the mass of the sub-sample of the PAM m_{PAM} were measured on an analytical balance (Ohaus Pioneer PA214, Ohaus Corporation, Parsippany, NJ, USA) calibrated with a weight mass standard of 0.002 g nominal value (Class E2-OIML, manufactured by Dolz Hnos SRL Argentina) and also with a weight mass standard of 100 g nominal value (Class 1-ANSI/ASTM E617, manufactured by Troemner Inc., Philadelphia, USA).

2.3.2. Homogeneity and stability testing

The sufficient homogeneity of the IHRM, respect to the total content of PAH, was tested by using the approach proposed by Fearn and Thompson [22]. The analyses of the IHRM were performed by triplicate instead of duplicate, applying the official reference method [12]. 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; and c) homogeneity test, between-sample standard deviation. Results obtained in the analysis of homogeneity data were summarized in Table 1. As the critical values were higher than the experimental values, the IHRM was considered sufficiently homogeneous. On the other hand, it was concluded that no instability could be demonstrated because of the assay results obtained for the IHRM with the official reference method showed no significant effects ($p > 0.11$) during the 30 days of the storage period of the IHRM.

2.3.3. Calculation of assigned value and associated uncertainty

On the assumption there were no chrysene losses during stirring or storage, the assigned value of the total PAH concentration in the IHRM was assessed according to the preparation procedure and taking into account the purity p_c of chrysene used, the mass $m_{chrysene}$ of the chrysene measured, the mass m_{PAM} of the sub-sample of the PAM measured, the buoyancy correction factor $Bu_{chrysene} = 1.0008$ for $m_{chrysene}$ and the buoyancy correction factor $Bu_{PAM} = 1.0006$ for m_{PAM}. The assigned value c_{IHRM} of the mass concentration of total PAH in the IHRM was calculated by the following formula (1):

$$c_{IHRM} = \left[\frac{m_{chrysene} \cdot p_c . Bu_{chrysene}}{m_{PAM} \cdot Bu_{PAM}}\right] \cdot 10^6 = 34.6\ mg\ kg^{-1} \quad (1)$$

Table 1
Homogeneity of the IHRM.

Test	Experimental value	Critical value	Result
Cochran	0.382	0.967	Pass
Analytical standard deviation	1.1	1.2	Pass
Homogeneity variance	2.92	7.72	Pass

Summary of the statistical results.

Formula (1) enables also calculation of the standard uncertainty of the assigned value c_{IHRM} by applying the law of propagation of uncertainty to this formula (1), i.e.:

$$\frac{u_{C_{IHRM}}}{c_{IHRM}} = \sqrt{\left(\frac{u^2_{m_{chrysene}}}{m^2_{chrysene}}\right) + \left(\frac{u^2_{m_{PAM}}}{m^2_{PAM}}\right) + \left(\frac{u^2_{p_c}}{p^2_c}\right) + \left(\frac{u^2_{Bu_{chrysene}}}{Bu^2_{chrysene}}\right) + \left(\frac{u^2_{Bu_{PAM}}}{Bu^2_{PAM}}\right)} \quad (2)$$

The standard uncertainties of the masses, $u_{m_{chrysene}} = 0.0002$ g and $u_{m_{PAM}} = 0.0004$ g, were derived from: (i) precision experiments, (ii) calibration process of the balance at the nominal values of 0.002 g and 100 g of weight size, and (iii) balance drift due to both temperature changes and between calibration events. Both uncertainties, $u_{m_{chrysene}}$ and $u_{m_{PAM}}$, included the respectively contributions of the tare and gross measurements. The standard uncertainty of the purity $u_{p_c} = 1\%$ was derived from the manufacturer's information. The uncertainties of the buoyancy correction factors were neglected because the uncertainties of the densities of the air, chrysene, PAM and weight reference standards were neglected. With this assumption and taking into account the standard uncertainties of $u_{m_{chrysene}}$, $u_{m_{PAM}}$ and u_{p_c}, it is possible to calculate the standard uncertainty $u_{C_{IHRM}}$ by applying the formula 2. Then, the obtained result of the standard uncertainty of the assigned value was $u_{C_{IHRM}} = 2.03$ mg kg^{-1}.

The standard uncertainty of the homogeneity of the IHRM, $u_{hom} = 1.71$ mg kg^{-1}, was taken as the square root of the homogeneity variance presented in Table 1.

The standard uncertainty of the homogeneity and the standard uncertainty of the assigned value were combined by taking the square root of the sum of the squares; an additional component equal to the homogeneity uncertainty was included to take account of any possible instability of the IHRM. Thus, the combined uncertainty of the IHRM was calculated according to this formula:

$$u_{IHRM\ [combined]} = \sqrt{u^2_{IHRM} + 2u^2_{hom}} \quad (3)$$

The combined uncertainty was $u_{IHRM\ [combined]} = 3.2$ mg kg^{-1}, with $\nu_{IHRM} = 25$ freedom degrees, estimated with the Welch–Satterthwaite approach [23]. Then the assigned value of the total PAH content in the IHRM and its expanded uncertainty were:

$(c_{IHRM} \pm U_{IHRM}) = (c_{IHRM} \pm t_{(25,\frac{\alpha}{2})}.u_{IHRM\ [combined]}) = (34.6 \pm 6.6)$ mg kg^{-1} at the level of confidence 0.95. This concentration level would be useful to evaluate the international intervention level of 40.00 mg kg^{-1} for total PAH [24].

2.4. Evaluation of the suitability of the IHRM

2.4.1. Extraction procedures

Two extractive procedures were applied to study the suitability of the IHRM: a) the standard procedure A (Soxhlet extraction) was performed following the EPA Method 3540 C [14]. Samples of the IHRM were weighed into pre-extracted and air-dried cellulose extraction thimbles Whatman grade 603, and the loaded thimbles were inserted into a preconditioned Soxhlet extractor with 150 mL of *n*-hexane-acetone (1:1, v/v). Extractions were carried out during 24 h, resulting in more than 100 extraction cycles. Finally, the raw extract was transferred into volumetric flasks, previously calibrated according to the methodology described by Rius et al. [25]. An aliquot of 10 mL of each one of these extracts was cleaned-up according to EPA Method 3650 B [26] in order to check sample interferences in total PAH determination. Finally these extracts were used for measuring the analytical signal of fluorescence in order to estimate the total PAH content, according to the international official method [12]; b) a non-standardized procedure B (UAME) was carried out in an open rectangular ultrasonic bath (Cole Parmer®, USA) with an operating power of 300 W at frequency of 40 Khz. The optimization of this UAME is described in the subsequent sections. Portions of 0.2 g of the IHRM were accurately weighed and sonicated during 25 min with 10 mL of POLE 0.050 M. After the extraction, the IHRM extracts were centrifuged at 2,500 rpm during 3 min and filtered. Then, these micellar extracts were transferred quantitatively to a volumetric vial and treated with 0.5 mL of 1.5 M HNO_3, according to the procedure described by Valcarcel et al. [16]. After that, the measurements of fluorescence of these micellar extracts were done by using a modification of the international official method.

2.4.2. Experimental designs for the optimization of the UAME procedure

As the UAME is not a standardized method, it was necessary first to optimize the experimental condition of the extractions. The micellar extraction method used was based on a previously described method for some selected PAH in spiked marine sediments [27]. However, in this work, three different surfactants, POLE, Triton X-100 and SDS, were used to select the most suitable extraction micellar solution for total PAH from the IHRM. The best micellar solution was selected according to the higher analytical signal. Before that, it was necessary to optimize the UAME conditions with each surfactant media. For this purpose, total fluorescence area was used as analytical signal and a global response for the total PAH content in the IHRM. This optimization process for each micellar medium was initially planned with a two-level full factorial design, 2^3, involving eight runs. These initial screening designs were used to detect those variables having the most influence on the experimental response. The variables studied as well as their values ('+' representing the maximum and '−' the minimum level) are shown in Table 2. Maximum and minimum levels were chosen according both to previous experience with the application of ultrasound for PAH extraction and from the available data in the literature. The minimum level of the surfactant concentration factor was selected as the lowest critical micellar concentration values of the studied surfactants [28].

Then, in order to obtain more accurate optimization for each micellar system, a central composite rotary design (CCRD) with $\alpha = 1.682$ and three central points was applied, considering the same three factors studied in the screening design. Table 3 lists the values given to each factor. All the experimental treatments (screening and optimization designs) were undertaken in a random order according to the arrangement table calculated for the used software (Unscrambler 9.7).

After this optimization process, five additional extractions were performed under conditions of repeatability for each optimized UAME. Finally, statistical analysis was conducted by evaluating the mean values of the global responses of these extractions.

2.4.3. Fluorescence determination of total PAH

The total PAH content of the IHRM extracts was determined by fluorimetric analysis using a SLM Aminco Bowman Series 2 luminescence spectrometer equipped with a xenon discharge light source (1,500 W). Wavelength accuracy and wavelength repeatability were ± 0.5 nm and ± 0.25 nm, respectively. Throughout the instrumental measurements the excitation and emission slits widths were adjusted at 8 nm and 4 nm, respectively; PMT voltage was kept at 600 V. Fluorescence measurements were performed using a standard 1 × 1-cm quartz cell. Calibrations for the micellar system and for the official method were accomplished by external standardization at eight concentration levels in the range from 0.60 to 9.00 mg L^{-1} of total PAH in terms of chrysene equivalent units. All calibration solutions

Table 2
Factor levels in the screening design (2^3).

Factor	Key	Low (−)	High (+)
Time (min)	A	5	30
Surfactant concentration (10^{-3} M)	B	0.1	100
Surfactant volume:amount of IHRM (ml/g)	C	6.25	18.75

Table 3
Factor levels used in the CCRDs for each micellar system.

Factor		Low (−)	High (+)	Centre	Axial (−α)	Axial (+α)
A	(min)	8.0	30.0	19.0	0.5	37.5
B	(10^{-3} M)	33	100	66.5	10.2	122.8
C	(ml/g)	13.0	50.0	31.5	0.4	62.6

were prepared with the standard solution of chrysene and analyzed five times under intermediate precision conditions.

The fluorescence of the Soxhlet extracts was measured at λ_{ex} = 310 nm, and the emission was recorded at λ_{em} = 360 nm. However, to measure the fluorescence of the micellar extracts it was necessary to optimize the official method taking into account each of the selected extraction medium. This modification involved verifying the wavelengths of excitation and emission with respect to the fluorescence of the surfactant medium in each UAME. For this purpose, a set of micellar solutions of the standard chrysene solution were prepared with concentrations of each surfactant at similar level of their critical micellar concentrations (CMC) and approximately equal to their solubility values in water. Ordinary least square (OLS) method was applied to calculate the calibration curves. Linear relationships with r > 0.99 were always obtained and validated (ANOVA at 95% of confidence level), and detection limits were between 1.24 mg L^{-1} and 2.48 mg L^{-1} of total PAH for the official method and for calibration curves of standard solutions in the different surfactants, respectively.

2.5. Statistics and data validation

All of the parameters of the calibration lines have been calculated with ULC 2.0 (Univariate Linear Calibration) computer software [29]. Validations of the recoveries data were performed by following the

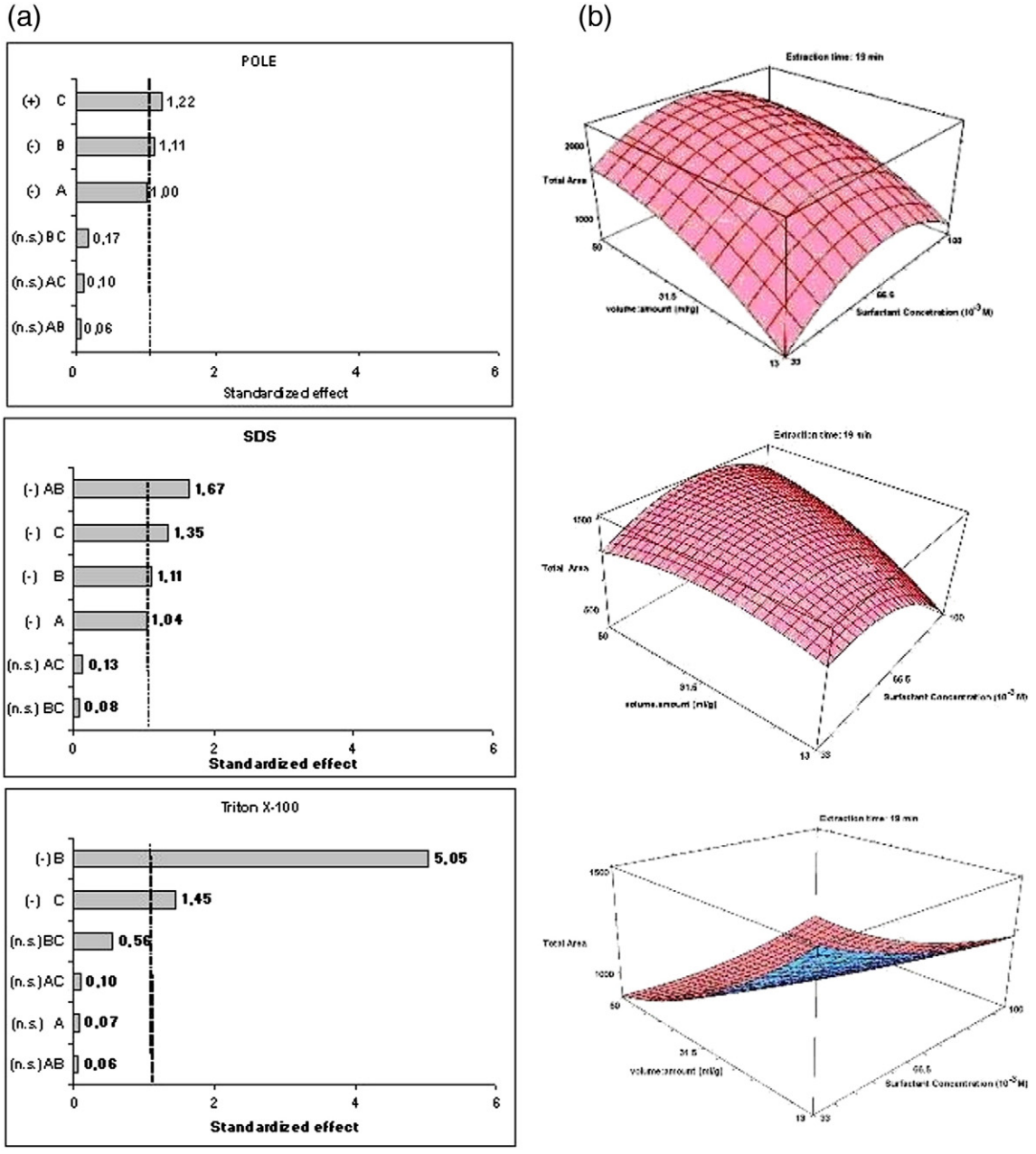

Fig. 1. (a) Pareto chart for the standardized main effects in the factorial 2^3 experimental designs and (b) estimated response surfaces for total fluorescence area using the central composite designs obtained by plotting surfactant concentration vs. volume:amount ratio. The vertical lines indicate the statistical significance bound of the effects.

Table 4
Optimal values for each UAME.

Factor	Key	Micellar system		
		POLE	SDS	Triton X-100
Time (min)	A	21	22	21
Surfactant concentration (10^{-3} M)	B	65	44	30
Surfactant volume:amount of IHRM (ml/g)	C	47	40	13

Table 5
Predicted and experimental values of response under optimal conditions.

Optimal values	Total area		
	POLE	SDS	Triton X-100
Actual (average, n = 5)	1972	1469	1479
Predicted	2081	1497	1507

acceptance criteria established by USEPA [30]. Parametric statistical test were used to check outlying results and the normality of the data [31–33]. ANOVA statistical calculations and Pareto charts were performed using Microsoft Office Excel ® 2010 (Microsoft, Redmond, WA, USA). For the analysis of the experimental design data, the Unscrambler 9.7 (CAMO SA, Oslo, Norway) was used.

3. Results and discussion

3.1. Optimization of the UAME procedure

The data obtained in the screening designs were evaluated by ANOVA at the 95% confidence level in order to estimate the significance of the main effects of the extraction variables and their interactions for each micellar system. These results can be shown in Fig. 1(a). In these graphics, the standardized effects are presented in rank order.

The Pareto chart shows that the surfactant concentration and the volume of surfactant:amount of sample ratio were the statistical significant variables (at $p < 0.05$) having the most influence on the experimental response (total fluorescence area). Extraction time was scarcely significant, with modest effect, for total fluorescence area. As can be seen, in the same Fig. 1(a), the UAME efficiency was affected by interrelated variables only for the SDS micellar system. In this case, the interaction between the factors extraction time and SDS concentration appears statistically significant, with negative sign, for the total fluorescence area. The region where these experiments were carried out did not include the optimum. Therefore, for planning the optimization, the CCRDs shifted in the direction of the maximum gradient were established. After the CCRDs, as expected from the screening experiments, the surfactant concentration and surfactant volume:amount ratio appeared as statistically significant main effects ($p < 0.01$). The lack of fit for the models obtained for each surfactant was not significant ($p << 0.05$), and this fact demonstrates that the regression equation explained correctly the variability of the experimental data. The optimal value for the extraction time was similar for the three extraction systems. These results, in time extraction, were within the range of previous reports for ultrasonic extraction of selected PAH in soils and marine sediments [27,34,35]. Fig. 1(b) shows the response surfaces plots for the total area obtained by plotting surfactant concentration versus surfactant volume:amount ratio when the extraction time was fixed at 22 min. After evaluation of these fitted surface graphs, the best experimental responses for extracting the total PAH compounds of the IHRM were summarized in Table 4.

Then, the optimal micellar medium was selected by using the results of the additional experiments performed in repeatability conditions with each surfactant at the optimal experimental values presented in Table 4. The experimental results of each UAME and their corresponding predicted values are summarized in Table 5.

ANOVA analysis of these results indicated that POLE was the surfactant that showed the highest responses ($p < 0.01$). SDS and Triton X-100 presented similar response values ($p > 0.5$). Then, POLE was selected as the best micellar system to use for the extraction of total PAH content in the sludge samples by applying the UAME optimized method.

3.2. Assessment the accuracy of the extraction methods with the IHRM

The assessment of the accuracy of the traditional Soxhlet method and the UAME method was assessed against the developed IHRM. For this purpose, two-factor fully-nested designs were originally defined for both methods. The factors studied in each design were the repeatability and the runs on which measurement was performed; where the runs were in the higher rank than the repeatability [36].

3.2.1. Comparison of precision

From the requirements previously indicated, two measurements ($n_s = 2$ replicates) were made in repeatability conditions on seven different runs ($p_s = 7$ runs). Later, a series of statistical tests (Grubbs' tests, single and pair, and Shapiro–Wilks' test) was conducted in order to investigate presence of outlying results and the normality of the data, respectively. Additionally, USEPA acceptance criterion was applied, in order to check recoveries in each analysis [30]. After checking these conditions, the originally fully-nested designs became unbalanced, then ANOVA calculations were conducted in this way in order to obtained the different variance terms: repeatability variance, s_r^2, the run-different intermediate variance, s_{run}^2, and the variance of the means, $s_{\bar{X}}^2$. This variance, $s_{\bar{X}}^2$, of both methods was compared by F-tests. There was no evidence that both methods have different run-different intermediate precision at 95% significance, but the repeatability precision of the UAME method was better than the Soxhlet method at the same significance level of evaluation. The total PAH mean recoveries obtained were 79% (RSD = 6%) and 73% (RSD = 9%) for the UAME and for the Soxhlet extraction method, respectively.

3.2.2. Comparison of trueness

If the future routine samples have similar level of concentration to the IHRM (intervention reference level of 40.0 mg kg^{-1}), the trueness of the extraction methods can be evaluated by checking whether the

Table 6
Results obtained in the assessment of accuracy.

Method	Test	Experimental values	Critical value	Result
Soxhlet	Shapiro Wilk	0.877	0.859	Pass
	Levene	5.00	5.52	Pass
	Trueness $\bar{X}$ = 25.3 mg kg^{-1} ($\alpha = 0.05$; $\nu = 19$)	2.778	2.093	Not pass *bias* = 9.3 mg kg^{-1}
UAME	Shapiro Wilk	0.899	0.859	Pass
	Levene	0.84	5.52	Pass
	Trueness $\bar{X}$ = 27.3 mg kg^{-1} ($\alpha = 0.05$; $\nu = 23$)	2.298	2.069	Not pass *bias* = 7.4 mg kg^{-1}
Precision comparison	$F_{(repeatability)}$	1.40	5.05	Pass
	$F_{(runs)}$	3.20	6.39	Pass
	$F_{(intermediate\ precision)}$	2.45	2.82	Pass

Table 7
Uncertainty values estimated for UAME and Soxhlet extraction methods (mg kg^{-1}).

Method	Sample/Homogeneity	μ_{bias}	μ_{proc}	$\mu_{ssamp.step}$	$t_{eff,\alpha/2}$	U	U/$\overline{X}$ (%)
Soxhlet	Routine/Not pass	3.4	2.2	6.0	2.045	14.8	43%
	IHRM/Pass			1.7	1.960	7.9	25%
UAME	Routine/Not pass	3.3	1.6	6.0	2.042	14.3	41%
	IHRM/Pass			1.7	1.960	7.2	23%

difference between the grand mean for each method, $\overline{X}$, and the assigned value of the IHRM, "the experimental bias", is not statically significant, i.e:

$$bias = |\overline{X} - C_{IHRM}| < t_{eff,\,\alpha/2} \cdot \sqrt{\left[\frac{s_I^2}{p_{eff}}\right] + u_{IHRM}^2}$$

where: $s_I^2 = s_r^2 + s_{run}^2$ is the variance of the intermediate precision; $p_{eff} = 12$ were the effective number of results after the originally fully-nested designs became unbalanced; and t was the two-sided tabulated t-value for $\alpha = 0.05$ and its associated number of effective degrees of freedom, ν_{eff}, was estimated with the Welch–Satterthwaite approach:

$$\nu_{eff} = \frac{\left[u_{IHRM}^2 + \left(\frac{s_I^2}{\nu_{S_I}}\right)\right]^2}{\frac{u_{IHRM}^4}{\nu_{IHRM}} + \frac{s_I^4}{\nu_{S_I}}}$$

The results of these comparisons are shown in Table 6.

Therefore, since the bias was statistically significant for both methods, the results of the routine samples should be corrected for the experimental bias at the studied concentration level of the IHRM.

3.3. Uncertainty estimation

3.3.1. Uncertainty arising from the analytical procedure

The standard uncertainty arising from the experimental variation of future measurements was calculated with the following equation:

$$\mu_{proc} = \sqrt{\frac{s_{run}^2}{p_s.n_s} + \frac{s_r^2}{n_s}}$$

where: p_s is the number of runs on which the measurement is carried out and n_s is the number of replicates performed every run. In this study, the worst case was assumed, then the standard uncertainties have been calculated when the routine samples is analyzed once (i.e. for $n_s = 1$ and $p_s = 1$).

3.3.2. Uncertainty arising from the assessment of trueness

This term corresponds to the uncertainty of the experimental bias,

$$u(bias) = \sqrt{\left[\frac{s_I^2}{p_{eff}}\right] + u_{IHRM}^2}$$

3.3.3. Uncertainty arising from other sources

This term includes the contribution of the sources of uncertainty arising from the heterogeneity of the material (subsampling) and from preprocessing steps applied to future working samples. In this study, the homogeneity of the IHRM was checked as it was proved in Section "In house reference material production". However, future working samples are not expected to be homogeneous, and so this contribution of uncertainty arising from heterogeneity must be included. On the other hand, the uncertainty from the preprocessing steps (weighing, drying, UAME, centrifugation, filtration, nitric acidification, etc.) are not included in this term, because this preprocessing steps were carried out on the IHRM using in the assessing the accuracy. The uncertainty arising from the heterogeneity was estimated with the subsampling step standard deviation, $s_{ssamp.step}$. For this purpose, ten different portions of the authentic sample, $n_{ps} = 10$, preserved as it was described in Section "Samples", were assaying under repeatability conditions. After checking the normality and rejection of two outliers, the $s_{ssamp.step}$ of these results was 6.0 mg kg^{-1} with 7 degrees of freedom.

3.3.4. Quantification of the overall uncertainty

The overall uncertainty was estimated by combining the uncertainties arising from the analytical procedure, the assessment of accuracy and the subsampling step, by applying the following equation:

$$U = t_{eff,\alpha/2}\sqrt{\left(\mu_{proc}^2 + \mu_{bias}^2 + \mu_{ssamp.step}^2\right)}$$

where: $t_{eff,\alpha/2}$ is the two-sided tabulated t-value for $\alpha = 0.05$, and its effective degrees of freedom were estimated using the Welch–Satterwaite approach.

Table 7 shows the results of the uncertainty values calculated for both methods.

Table 7 also shows that the relative expanded uncertainties values are very high for the routine sample (41% and 42 % for UAME and Soxhlet extraction respectively) and decrease considerably when the analyzed sample was the IHRM. As can be seen from Fig. 2, this decrease is due to the different contribution of the subsampling step, i.e. from lack of homogeneity of the analyzed sample. This regard emphasizes the importance of the homogenization treatment of the laboratory sample prior to subsampling/splitting steps. Regulatory standard methods

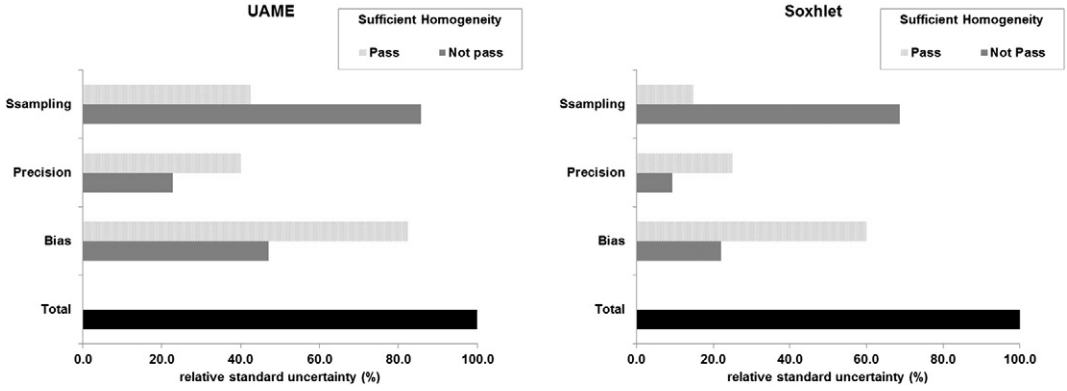

Fig. 2. Contribution of the analytical procedure, assessment of accuracy and subsampling to the measurement uncertainty in both extraction methods studied. These contributions are shown for the two homogeneity conditions of samples.

for the analysis of environmental complex matrices do not used to require a quality control of the subsampling in terms of the homogeneity. Moreover, the contribution of the uncertainty arising from heterogeneity (lack of homogeneity) is rarely evaluated to assess its effect on the overall uncertainty of the final results or on the decision-making process based on the analytical results. Here, it was demonstrated with relevant data, that the preparation of subsamples (especially of complex matrix) equally representative of the as-received sample submitted to the laboratory should be validated in order to ensure the "sufficient homogeneity"; e.g.: soil analysis [37], waste or wastewater analysis.

4. Conclusions

A full and practical methodology for assessing accuracy and calculating the measurement uncertainty was applied. This approach is based on the information from the precision and bias estimates calculated in assessing the accuracy with the IHRM prepared for its intended purpose. This methodology can be very valuable in absence of references of the highest metrological hierarchy (e.g. primary method, primary CRM).

An in-house reference material, IHRM, was developed for quality control purposes of total PAH determination in waste samples at the international level of reference. The IHRM demonstrated to be suitable for optimizing a non-standardized extraction method, and also it was useful to assess the accuracy of the results obtained with the two studied extraction methods.

The obtained overall uncertainties were thoroughly dependent on the assurance of the "sufficient homogeneity" of laboratory sample prior to the subsampling/splitting steps of the analytical process. This remark is particularly relevant when small sample quantities are required to be analyzed. The obtained uncertainty budgets show the essential role of the uncertainty in analytical quality controls and decision-making assessments.

Acknowledgements

This work was supported by funds provided by the Universidad Nacional del Sur, Argentina (PGI 24/Q050). Dra. Fernández Band acknowledges Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET).

References

[1] European Community, Commission Directive 99/31/EC of 26 April 1999, Off. J. Eur. Communities L182 (1999) 14–16.
[2] Ley No 24051 y Anexos, Residuos Peligrosos, Congreso de la Nación Argentina, Boletín Oficial de la Nación Argentina No 27307 del 17 de enero de 1992, Buenos Aires, Argentina, 1992.
[3] Ley No 11720, Generación, Manipulación, Almacenamiento, Transporte, Tratamiento y Disposición Final de Residuos Especiales, Congreso de la Provincia de Buenos Aires, Boletín Oficial de la Provincia de Buenos Aires No 23030 del 13 de diciembre de 1995, La Plata, Argentina, 1995.
[4] S.O. Baek, R.A. Field, M.E. Goldstone, P.W. Kirk, J.N. Lester, R. Perry, A review of atmospheric polycyclic aromatic hydrocarbons: sources, fate and behaviour, Water Air Soil Pollut. 60 (1991) 279–300.
[5] U. S. Code of Federal Regulations, Title 40, Part 423, Appendix A, U.S. Government Printing Office, Washington, DC, 2013.
[6] European Community, European Commission, DG Environment, Regulatory report, part 2, disposal and recycling routes for sewage sludge, Office for Official Publications of the European Communities, Luxembourg, 2001.
[7] BIPM, IEC, IFCC, ISO, IUPAP, OIML, ILAC, Guide to the expression of uncertainty in measurement, ISO, Geneva, 2008.
[8] S.L. Ellison, A. Williams (Eds.),Eurachem/CITAC guide: quantifying uncertainty in analytical measurement, edition 3, 2012.
[9] B. Magnusson, T. Näykki, H. Hovind, M. Krysell, Handbook for calculation of measurement uncertainty in environmental laboratories, NT Tech Report 537, Edition 3.1 Nordtest, Oslo, 2012.
[10] A. Maroto, J. Riu, R. Boqué, F.X. Rius, Estimating uncertainties of analytical results using information from the validation process, Anal. Chim. Acta. 391 (1999) 173–185.
[11] B. Brian, Guidelines for the in-house production of reference material. LGC/VAM/1998/090, Teddington, Middlesex, TW11 0LY, UK, 1998.
[12] UNESCO, Intergovernmental Oceanic Commission (IOC), The determination of petroleum hydrocarbons in sediments, Manuals and Guides No 11, Paris, 1982.
[13] C. Miège, J. Dugay, M.C. Hennion, Optimization and validation of solvent and supercritical-fluid extractions for the trace-determination of polycyclic aromatic hydrocarbons in sewage sludges by liquid chromatography coupled to diode-array and fluorescence detection, J. Chromatogr. A 823 (1998) 219–230.
[14] US EPA Method 3540, Soxhlet extraction, Revision 3, US Government Printing Office, Washington, DC, USA, 1996.
[15] US EPA method 8310, polynuclear aromatic hydrocarbons, US Government Printing Office, Washington, DC, USA, 1986.
[16] A. Criado, S. Cárdena, M. Gallego, M. Valcárcel, Direct automatic screening of soils for polycyclic aromatic hydrocarbons based on microwave-assisted extraction/fluorescence detection and on-line liquid chromatographic confirmation, J. Chromatogr. A 1050 (2004) 111–118.
[17] ATSDR (Agency for Toxic Substances and Disease Registry), Toxicological profile for polycyclic aromatic hydrocarbons, US Public Health Service, Atlanta, Georgia, USA, 1995.
[18] H.A. Al Darwish, E.A. Abd El-Gawad, F.H. Mohammed, M.M. Lotfy, Assessment of organic pollutants in the offshore sediments of Dubai, United Arab Emirates, Environ. Geol. 48 (2005) 531–542.
[19] R. Venkatachalapathy, S. Veerasingam, T. Ramkumar, Petroleum hydrocarbon concentrations in marine sediments along Chennai Coast, Bay of Bengal, India, Bull. Environ. Contam. Toxicol. 85 (2010) 397–401.
[20] R. Venkatachalapathy, S. Veerasingam, P. Raja, S. Sudhakar, V. Rajeswari, R. Mohamed Asanulla, R. Mohan, P. Sutharsan, Petroleum hydrocarbon concentrations in ten commercial fish species along Tamilnadu coast, Bay of Bengal, India, Environ. Sci. Pollut. Res. 18 (2011) 687–693.
[21] M.P. Llompart, R.A. Lorenzo, R. Cela, K. Li, J.M.R. Bélanger, J.R. Jocelyn Pare, Evaluation of supercritical fluid extraction, microwave-assisted extraction and sonication in the determination of some phenolic compounds from various soil matrices, J. Chromatogr. A 774 (1997) 243–251.
[22] T. Fearn, M. Thompson, A new test for 'sufficient homogeneity', Analyst 126 (2001) 1414–1417.
[23] K.A. Brownlee, Statistical theory and methodology in science and engineering, John Wiley & Sons, Inc., New York, 1960. 236.
[24] E.M.J. Verbruggen, R. Posthumus, A.P. van Wezel, Ecotoxicological serious risk concentrations for soil, sediment and (ground)water: updated proposals for first series of compounds, RIVM report 711701 020, National Institute of Public Health and the Environment, 3720 BA Bilthoven, Netherlands, 2001.
[25] J. Riu, R. Boque, A. Maroto, F.X. Rius, Trazabilidad en medidas físicas mediante calibración directa: Calibración de material volumétrico, Técnicas de Laboratorio 269 (2002) 112–114.
[26] US EPA Method 3650B, Acide-base partition cleanup, Revision 2, US Government Printing Office, Washington, DC, USA, 1996.
[27] V. Pino, J.H. Ayala, A.M. Afonso, V. González, Ultrasonic micellar extraction of polycyclic aromatic hydrocarbons from marine sediments, Talanta 54 (2001) 15–23.
[28] I.T. Yeom, M.M. Ghosh, C.D. Cox, K.G. Robinson, Micellar solubilization of polynuclear aromatic hydrocarbons in coal tar-contaminated soils, Environ. Sci. Technol. 29 (1995) 3015–3021.
[29] R. Boque, F.X. Rius, D.L. Massart, Straight line calibration: something more than slopes, intercepts, and correlation coefficients, J. Chem. Educ. 71 (1994) 230–232.
[30] Region 1 EPA-NE, data validation functional guidelines for evaluating environmental analyses, part II, section VOA/SV — IX, U.S. Environmental Protection Agency, Office of Science and Technology, Washington, DC, 1996.
[31] Association of Official Analytical Chemist (AOAC), Guidelines for collaborative study. Procedures to validate characteristics of a method of analysis, J. AOAC Int. 78 (1995) 143A–160A.
[32] 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. 452–454.
[33] S.S. Shapiro, M.B. Wilk, An analysis of variance test for normality (complete samples), Biometrika 52 (1965) 591–611.
[34] E. Aamot, J. Kraine, E. Steinnes, Determination of trace amounts of polycyclic aromatic hydrocarbons in soil, Fresenius' Z. Anal. Chem. 328 (1987) 569–571.
[35] F. Sun, D. Littlejohn, M.D. Gibson, Ultrasonication extraction and solid phase extraction clean-up for determination of US EPA 16 priority pollutant polycyclic aromatic hydrocarbons in soils by reversed-phase liquid chromatography with ultraviolet absorption detection, Anal. Chim. Acta. 364 (1998) 1–11.
[36] ISO, Accuracy (trueness and precision) of measurement methods and results. ISO 5725-3:1998 (E), International Organization for Standardization, Geneve, Switzerland, 1998. 1–56.
[37] F. Risso, R. Narizzano, M. Ottonelli, A. Magherini, M. Giampieria, E. Carlinia, Validation of subsampling procedures and granulometry distribution uncertainty assessment in environmental soil analysis, Anal. Methods 4 (2012) 570–574.