

Roland Becker · Hans-Gerhard Buge ·
Wolfram Bremser · Irene Nehls

Mineral oil content in sediments and soils: comparability, traceability and a certified reference material for quality assurance

Received: 31 January 2006 / Revised: 10 March 2006 / Accepted: 13 March 2006 / Published online: 29 April 2006
© Springer-Verlag 2006

Abstract The performance of twelve laboratories with previously established proficiency in the determination of the mineral oil content in a fresh water sediment is described. The summation parameter total petrol hydrocarbon (TPH) is defined according to ISO 16703:2004 with regard to the sample preparation to be applied, the flame ionisation detection (FID) and the boiling range of C₁₀–C₄₀ to be integrated. Comprehensive tests of homogeneity and stability have been carried out on the candidate material using appropriate models. The outcome of the study served as the basis for the certification of the candidate reference material as ERM-CC015a. The certified mass fraction is 1,820±130 mg kg⁻¹ and traceability was established by using an appropriate calibration standard certified for the mass fraction of C₁₀–C₄₀. The interlaboratory scatter of measurement results in this exercise can largely be explained by the variability of the individual calibrations based on this common calibration standard.

Keywords Total petroleum hydrocarbons · Extraction · Gas chromatography · Calibration · Interlaboratory study

Introduction

Mineral oil hydrocarbons are among the pollutants most often found on contaminated sites. Mineral oils are composed of a vast mixture of hydrocarbons displaying a boiling range dependent on the origin and processing of the base mineral oil. The quantification of the mineral oil content is reasonably done as a summation parameter that covers a defined boiling range and is often referred to as total petroleum hydrocarbons (TPH). Extraction procedures [1–3] and the choice of detection method [4, 5] for

the quantification of TPH in various environmental matrices such as soil, water and waste have recently become a matter of ongoing concern in the literature and are the subject of international standardisation activities.

Due to its variable nature, TPH is a parameter that is defined by using a method which must be followed strictly. The comparability of TPH quantification results in soil and similar matrices from different laboratories is ensured by following the international standard ISO 16703:2004 that prescribes the extraction procedure, clean-up of the extract and the boiling range to be quantified. It offers guidance on the use of an appropriate calibrant and stipulates flame ionisation detection (FID), which (compared to e.g. mass-selective detection) displays linearity of the signal response over six orders of magnitude. At the same time, the FID response of CH-equivalents is independent of the respective distribution of alkane chain isomers [6, 7]. The method is therefore suitable for the overall quantification of hydrocarbon mixtures regardless of their respective composition of congeners.

Routine laboratories conduct the vast majority of environmental analyses and participate in external quality assurance measures such as proficiency testing [8]. In addition, there is a need for reference materials to be applied as a tool for internal quality control. Here, we describe the certification of a respective matrix reference material produced from a real-world sediment and its traceability to a stated reference.

Materials and methods

Chemicals and reagents

n-Heptane, *n*-hexane and acetone for residual analysis were purchased from Promochem (Wesel, Germany). Retention time markers *n*-decane (C₁₀) and *n*-tetracontane (C₄₀) were purchased from Fluka (Basel, Switzerland) and used as a solution in *n*-heptane (about 50 mg L⁻¹).

Calibration standard BAM-K010, a certified mixture of additive-free diesel and lubricating oil as specified in ISO

R. Becker (✉) · H.-G. Buge · W. Bremser · I. Nehls
Federal Institute for Materials Research and Testing (BAM),
Richard-Willstätter-Strasse 11,
12489 Berlin, Germany
e-mail: roland.becker@bam.de

16703:2004, was used to cover the range 1,000–5,000 mg kg⁻¹ equidistant with five calibration solutions prepared on a weight/weight basis.

Two control solutions of BAM-K010 in *n*-hexane were prepared (0.713 and 1.619 mg mL⁻¹) and filled in 4.5 Certan containers (Promochem, Wesel, Germany). Magnesium silicate (Florisil, 60–100 mesh) from Merck (Darmstadt, Germany) was heated for 16 h at 140°C and allowed to cool in a stream of dry nitrogen.

Sample extraction and gas chromatography

Following ISO 16703:2004, 5 g of soil was shaken with 20 mL of acetone and after addition of 10 mL of *n*-heptane containing C₁₀ and C₄₀, and the mixture was sonicated for 1 h at 60°C. A 45 mL aliquot of water was added followed by shaking (10 min) and centrifugation (10 min, 3,500 rpm). The organic phase was collected and centrifuged (10 min, 3,500 rpm) after addition of 60 mL of water. A 5 mL sample of the organic phase was submitted to clean-up (Florisil/Na₂SO₄ column).

TPH was quantified using an Agilent 6890 gas chromatograph (GC) equipped with an Agilent 7683 autosampler (Agilent, Waldbronn, Germany), a FID (detector temperature 370°C) and an HT 8 capillary column (15 m × 0.25 mm × 0.25 μm); the injection volume was 3 μL. N₂ (5.0) was used as carrier gas (3 mL min⁻¹). The GC oven program started at 50°C, was held for 10 min, increased to 360°C at 30°C min⁻¹ and then held for 10 min; this enabled a complete run within 20 min (see Fig. 1).

Preparation of the candidate material

A representative fresh water sediment contaminated with mineral oil from industrial and municipal input over decades was sampled from River Weisse Elster near Leipzig, Germany. The material was air-dried to constant

weight and fractionated by sieving after crushing of agglomerates. A total of 29.47 kg of sample was collected after passing through a 125-μm mesh and this was then homogenised using a drum hoop mixer for 6 h. Bottling was performed by means of a spinning riffler using a service-proofed procedure of partitioning and back-mixing to yield 360 units containing 81.5±0.5 g sediment in brown glass bottles which were stored at -20°C in the dark. Table 1 lists the characterisation data of the matrix, and Fig. 1 depicts the TPH pattern. It should be noted that the boiling range above C₄₀ is not included in the mass fraction of TPH.

Characterisation

Homogeneity

The variance of the respective analytical procedure s_{method}^2 and the variance of the TPH content within a bottle s_{wb}^2 are initially unknown and cannot be investigated independently.

This is achieved here by reducing the heterogeneity as far as technically feasible and by choosing a sample intake for one determination such that its further increase does not significantly decrease the repeatability standard deviation of TPH quantification. At this point s_{wb}^2 —the variability due to within-bottle heterogeneity—is considered negligible and the observed variability of the TPH content s_r^2 is dominated by s_{method}^2 .

$$s_r^2 = s_{wb}^2 + s_{method}^2 \quad (1)$$

Every 24th unit was taken from the bottling process and the TPH content was determined five times in these 15 units under repeatability conditions. Processed extracts were analysed by GC/FID using the method outlined above

Fig. 1 Gas chromatogram of an extract of ERM-CC015a

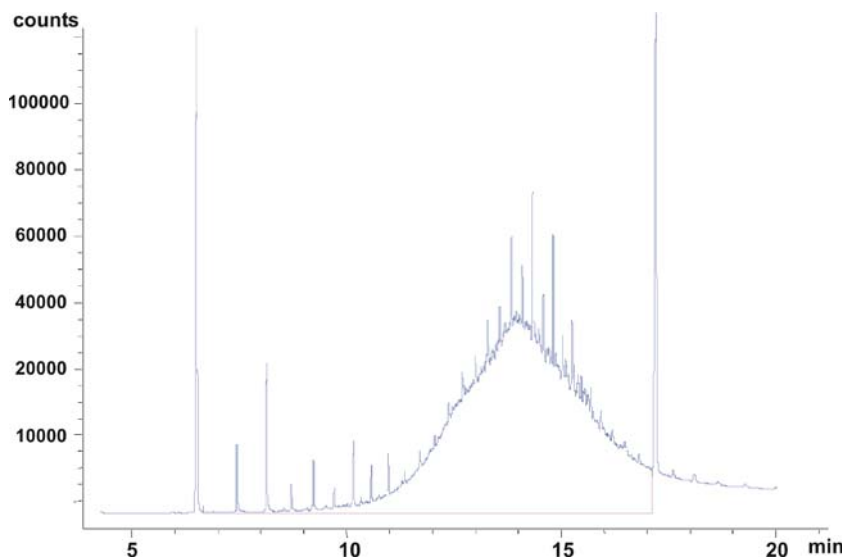


Table 1 Characterisation of the sediment matrix

Property	Value	Method
Particle size range (μm)	0–125	Sieving classification DIN 18123
Water content (%)	5.72 \pm 0.10	Karl–Fischer titration
Drying loss (%)	5.60 \pm 0.10	Drying to constant weight at 105°C (ISO 11465)
Total organic carbon (g kg^{-1})	84 \pm 8	ISO 10694
Total inorganic carbon ($\mu\text{g kg}^{-1}$)	0.60 \pm 0.04	ISO 10694
CHN analysis (%)	C: 8.45 \pm 0.02 H: 1.81 \pm 0.01 N: 0.62 \pm 0.01	Combustion
pH	5.78 \pm 0.01	E DIN ISO 10390

under repeatability conditions such that all 75 extracts were quantified against one calibration after randomisation.

No evidence suggesting a rejection of the hypothesis that the material is homogeneous was observed and the uncertainty of the TPH content between the bottles u_{bb} was estimated as 26.97 mg kg^{-1} for 5 g sample intake according to Eq. (2) [9] (see also Table 2).

$$u_{bb} = \sqrt{\frac{MS_{\text{between}} - MS_{\text{within}}}{5}} \quad (2)$$

Stability

From experience a temperature-driven deterioration of the TPH content was to be expected for this material. Therefore, a total of 11 units of the candidate material were submitted to a so-called isochronous study [10] with accelerated ageing at constant temperatures between 4 and 60°C over periods of three weeks to 12 months as depicted in Fig. 2a. After the respective periods of time, respective units were stored again at the reference temperature of –20°C. All units were analysed for TPH using the method

described above together with three reference samples which had been kept at –20°C since bottling. Each extract was measured three times together with procedural blanks. The whole set of measurements included 64 runs conducted under repeatability conditions.

A significant downwards trend for the TPH content is observed for higher temperatures. The dependence of the thermal degradation of the TPH content on time is modelled using the basic assumption of proportionality between the amount of analyte deteriorated (dm) during a time interval dt and the total amount of the analyte available in the sample. The proportionality coefficient is the temperature-dependent reaction rate $k_{\text{eff}}(T)$. Thus, from

$$dm = -mk_{\text{eff}}(T)dt$$

where m is the mass of the analyte (total amount) in the sample, one obtains an exponential dependence on time of the remaining amount of analyte (see also e.g. [11]). Although this assumption is not universal, it merely describes a large number of degradation processes. At a given temperature, the effective pseudo-first-order reaction rate $k_{\text{eff}}(T)$ may consist of more than one and even concurring degradation processes all having their individual dependence on temperature.

Table 2 Results of the statistical evaluation of the homogeneity study and interlaboratory exercise

Quantity	Abbreviation	Unit	Value	Source
Mean square error within bottles	MS_{within}	$\text{mg}^2 \text{kg}^{-2}$	11,817.03	ANOVA ^a
Mean square error between bottles	MS_{between}	$\text{mg}^2 \text{kg}^{-2}$	15,451.65	ANOVA ^a
Between-bottle uncertainty	u_{bb}	mg kg^{-1}	26.97	Eq. (2)
Correction factor for calibrant purity	f_{pur}	–	0.967	BAM-K010
Relative uncertainty of calibrant purity	$u_{\text{pur},r}$	–	0.009	BAM-K010
Mean of the intercomparison	w_{char}	mg kg^{-1}	1,886.36	Interlaboratory
Standard uncertainty of w_{char}	u_{char}	mg kg^{-1}	57.77	Study
Relative standard uncertainty	$u_{\text{char},r}$	%	3.06	
Certified mass fraction of TPH	w_{cert}	mg kg^{-1}	1,820	Eq. (3)
Uncertainty of w_{cert}	u_c	mg kg^{-1}	65	Eq. (4)
Relative uncertainty of w_{cert}	$u_{c,r}$	%	3.57	
Expanded uncertainty of w_{cert}	U	mg kg^{-1}	130	$2 u_c$
Relative between-laboratory calibration uncertainty	$u_{\text{calibr},r}$	%	4.45	Eq. (6)

^a1-way analysis of variance

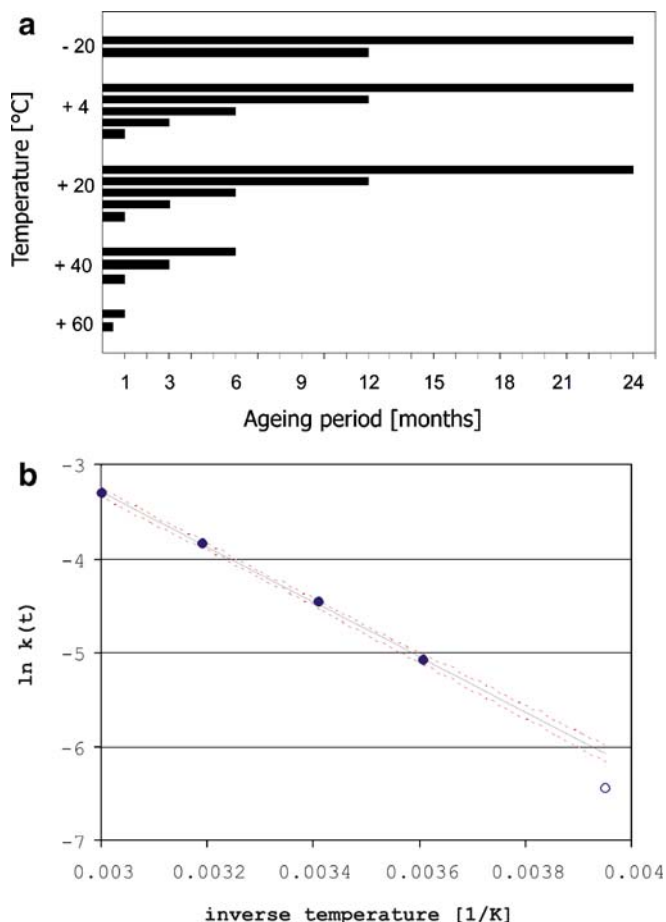


Fig. 2 a Accelerated ageing, temperatures and periods. The isochronous study is done after 12 months. The first post-certification monitoring is planned after 24 months. **b** Arrhenius plot for ERM-CC015a. Regression of TPH degradation rate $k(T)$ over the inverse temperature and 95% one-sided confidence interval. Estimated shelf lives are 2.3, 0.9, 0.5 and 0.3 years at -20°C , $+4^{\circ}\text{C}$, $+20^{\circ}\text{C}$ and $+40^{\circ}\text{C}$, respectively

Modelling of the dependence of $k_{\text{eff}}(T)$ on temperature is therefore quite complicated since degradation may be fed by different sources, especially in natural matrices where all accompanying substances and their influence on the degradation process are not normally known. Thus, a simple Arrhenius model was assumed and tested against experimental data. The reasoning behind this approach is outlined in detail elsewhere [12]. A plot of the reaction rates $k_{\text{eff}}(T)$ over the inverse temperature is depicted in Fig. 2b.

Obviously, the temperature dependence can indeed be approximated by a straight line in the range from $+4^{\circ}\text{C}$ to $+60^{\circ}\text{C}$. The estimated $k_{\text{eff}}(T)$ at -20°C does not fit well with the model, a fact which may be attributed to mainly two reasons: either a second degradation process which fully dominates the high-temperature region is activated at temperatures above -20°C , or the discrepancy between prediction and experiment is due to the influence of the innate method variability. Although the absolute value of the latter is approximately constant for all sampling points

(in temperature and time) involved, its relative influence is largest at the lowest temperature where no or only a very small (insignificant in a statistical sense) degradation is both expected and observable in the experiment.

Within-laboratory method variability can substantially be reduced by partially taking out drift effects. For all temperatures above reference, this can simply be achieved by referring the absolute value measured for a stressed sample to the value measured for the reference sample which is closest in the measurement sequence (normalisation). The raw data from the exercise described here were normalised, then subjected to the procedure as described above.

Data for the samples stored at reference temperature cannot be normalised. Formal evaluation of the raw reference sample values obtained in this exercise revealed a positive (increasing) trend mainly due to drift in the measurement sequences. These drift effects were investigated using an iterative fit procedure on all data measured in one sequence, appropriate corrections were derived also for the references, and a reasonable reaction rate estimate obtained.

This observed effective reaction rate for samples stored at -20°C is below the rate predicted from the Arrhenius model and was not included in the regression of $k_{\text{eff}}(T)$ versus $1/T$ (Fig. 2b). By using the upper one-sided confidence bound of the predicted value for $k_{\text{eff}}(T)$, one overestimates the degradation process and thus underestimates the real validity of the material. Shelf life estimates are in this sense worst-case estimates to the safe side.

The activation energy ΔE_a estimated for the high-temperature region is $24.4 \pm 0.4 \text{ kJ mol}^{-1}$. By using these data and the assumed model, an estimate can be obtained when degradation will presumably force the certified mineral oil content to fall outside the certified expanded uncertainty bounds (normally the lower bound) for any required temperature. This allows the assessment of both the shelf life at storage temperature and any limitations during transportation to the end user. Note that, for this assessment, the upper one-sided confidence bound of the predictions for $k_{\text{eff}}(T)$ is used, thus the uncertainty of its determination is included in the evaluation.

Although the preliminarily estimated shelf life at a storage temperature of -20°C is satisfactory and the stability over short periods at ambient temperatures is sufficient for dispatch and use, any prolonged exposure to room or even higher temperatures will drastically reduce the period of validity of the certified values. Therefore, their validity is set to one year after delivery provided that the transport conditions meet the restrictions set by the estimated shelf life and the material is stored at -20°C by the user upon receipt. The first rough estimation of stability will be updated over the period of availability of the material by post-certification monitoring of units stored at 4°C and 20°C . Measurement intervals will be about half of the shelf life estimated on basis of all earlier measurements.

Certification study

Study design

Nine laboratories selected on grounds of their performance in two recent proficiency testing rounds on TPH analysis in soil operated by BAM were invited to participate in the certification exercise along with three laboratories from BAM. Selection criteria included the consistency of documentation of extraction, clean-up, calibration and instrumental analysis according to ISO 16703:2004 and the declaration of commitment to comply with these requirements during the certification analyses.

Two units of the candidate material were to be analysed by each laboratory in triplicate. A unit of the certified calibration standard BAM-K010 was provided for each participant along with the information that the level of content was expected to be between 1,000 and 4,000 mg kg⁻¹.

In addition, each participant received the two control solutions whose concentrations of TPH in *n*-heptane were unknown (to the participants). These were prepared such that they corresponded to the lower and upper region of the appropriate range of calibration. The standard procedure according to ISO 16703:2004 had to be followed strictly and was to be documented by means of a detailed questionnaire.

Fig. 3 Results of the certification study. **a** Laboratory means and standard deviations (sediment): certified value w_{cert} and expanded uncertainty U . **b** Relative deviation of laboratory means from mass fraction of BAM-K010 (control solutions)

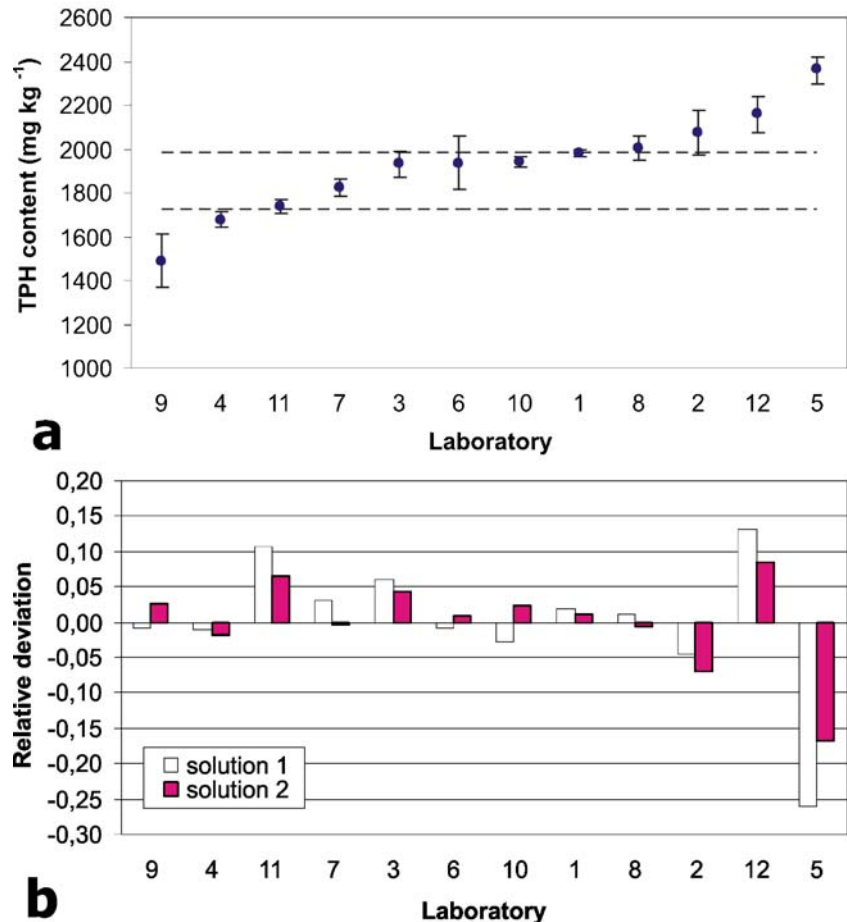


Table 3 Comparison of calibrations

No.	Δ_b^a (%)	Δ_{TPH}^b (%)
1	0.1	-0.5
2	1.7	7.0
3	0.2	3.8
4	0.2	-2.8
5	5.1	-23
6	1.5	-1.5
7	0.1	0.2
8	5.0	-0.9
9	1.2	-0.5
10	0.7	-1.0
11	10.1	6.5
12	2.5	8.0

^aRelative deviation of slopes of calibration and control solutions 1+2

^bRelative deviation of reference and measured TPH content in control solutions 1+2

Evaluation of certified value and its uncertainty

Results for the mineral oil content were to be reported on the basis of total mass intake, and no dry mass determinations were asked for in order to exclude another source of error and since it had previously been established

that the water content is not altered significantly if the material is handled as specified. Results of the TPH determination were scrutinised for consistency and a few obvious transcription errors were corrected after clarification with the respective laboratories.

As can be seen from Fig. 3 most laboratories show positive or slightly negative correlation between the sample and the control value. Only laboratory number 5 is distant from the main group, and additionally shows a considerable anti-correlation between sediment sample and control solutions.

Although no further investigations into the reasons for this discrepancy have been conducted, it seems justifiable to exclude, for technical reasons, the values of this particular laboratory from further processing. This is also supported by the comparison of the slopes of the calibration curves of the participants and their measurements of the control solutions (Table 3). Since the control solutions are made from BAM-K010 they should fit into the calibration curves done with BAM-K010. This is reasonably true for the majority of participants, whereas laboratory 5 is among the three exceptions with a high relative deviation Δ_b of their calibration slopes from those obtained from the two control solutions.

Δ_{TPH} is the relative difference of the value for the TPH mass fraction in the sediment extract either calculated from the reported concentrations for the control solutions or from the respective reference values. As expected, laboratory 5 deviates significantly from the other participants. The reason for these discrepancies is most likely due to the manual integration of TPH range which is more prone to variation caused by the operator than in the case of an ordinary gaussian peak.

Due to the lack of variance homogeneity (Snedecor F-test, Scheffé test and Bartlett test) laboratory means were used for further evaluation rather than the whole set of single value after pooling. No indication for outlying observations (Cochran, Grubbs, Nalimov) or a non-normal distribution were found.

Therefore, the mean of laboratory means was taken as the best estimate for the TPH mass fraction w_{char} and the standard uncertainty of the mean of laboratory means as the uncertainty contribution u_{char} from the intercomparison.

The estimate for the certified TPH mass fraction w_{cert} is derived by correction for the purity f_{pur} of the calibration standard used for determination according to Eq. (3).

$$w_{cert} = w_{char} \times f_{pur} \quad (3)$$

The corresponding combined uncertainty u_c must appropriately be composed from the uncertainty contributions from the certification intercomparison u_{char} , a possibly undetected inhomogeneity u_{bb} , and the uncertainty of the purity correction u_{pur} according to Eq. (4).

$$u_{c,r}^2 = u_{char,r}^2 + u_{bb,r}^2 + u_{pur,r}^2 \quad (4)$$

The expanded uncertainty is $U=ku_c$ and the expansion factor is $k=2$. The certified value and the expanded uncertainty are rounded according to the published recommendations [13] to be $1,820 \pm 130 \text{ mg kg}^{-1}$.

The standard uncertainty of the intercomparison u_{char} is composed of a contribution due to the variability of calibration among the laboratories u_{calibr} , and other components u_{rest} including those from sample preparation Eq. (5).

$$u_{char}^2 = u_{calibr}^2 + u_{rest}^2 \quad (5)$$

An estimate for u_{calibr} may be obtained from the combination of the relative deviations of reference value and measurement results (Table 3) according to Eq. (6).

$$u_{calibr}^2 = \frac{\sum \Delta_{TPH}^2}{n} \quad (6)$$

where n is the number of laboratories. The average calibration-caused deviation u_{calibr} amounts to 4.45%. The average lab deviation from the mean of laboratory means for the samples under certification is 9.69% (not to be confused with the standard uncertainty of the mean $u_{char,r}$). Thus, approximately 21% of the overall variance of the certification exercise is explained by calibration variability. Although other sources of uncertainty, namely sample preparation, contribute the major part of the study variability, calibration remains a significant factor and should be controlled. This finding is opposite to the commonly stressed argument about the negligibility of all “equipment-related” influences. The provision of reliable, gravimetry-based synthetic control samples is a convenient way of keeping calibration influences under control.

Conclusions

The development of a reference material with a unique matrix/analyte combination and the traceability of the certified value to an internationally accepted stated reference is presented. The significance of any variability of the TPH content was derived from a comprehensive homogeneity study in accordance with ISO Guide 35 on the certification of reference materials. The long-term stability was estimated on grounds of a sound model whose validity has successfully been tested against the data. Worst-case estimation of the expiry date and regular post-certification monitoring of the material make sure that the user obtains a CRM with reliable certified value plus corresponding uncertainty. Somewhat unexpectedly, the calibration contributed significantly to the interlaboratory variability. As a result, it may be concluded that even the relative tedious sample preparation in question has reached a conclusive state of the art among the selected laboratories.

Acknowledgements The authors wish to acknowledge the efforts of and the contributions made by the laboratories participating in this interlaboratory certification study.

References

1. Xie GB, Barcelona MJ, Fang JS (1999) *Anal Chem* 71:1899–1904
2. Buddhadasa SC, Barone S, Bigger SW, Orbell JD (2003) *J Jpn Petrol Inst* 46:28–34
3. Ventura K, Adam M, Dostalek JJ (2003) *Liq Chromatogr R T* 26:247–259
4. Krupcik J, Oswald P, Oktavec D, Armstrong DW (2004) *Water Air Soil Poll* 153:329–341
5. Miller MW, Wright MS, Allen AM, Dudek B (2005) *Soil Sediment Contam* 14:105–113
6. Jorgensen AD, Picel KC, Stamoudis VC (1990) *Anal Chem* 62:683–689
7. Kallai M, Veres Z, Balla J (2001) *Chromatographia* 54: 511–517
8. Becker R, Koch M, Wachholz S, Win T (2002) *Accred Qual Assur* 7:286–289
9. ISO Guide 35: general and statistical principles for certification (approved revision, April 2005). ISO, Geneva, Switzerland
10. Lamberty A, Schimmel H, Pauwels J (1998) *Fresenius J Anal Chem* 360:359–361
11. Heydorn K (1998) *Accred Qual Assur* 3:111–114
12. Bremser W, Becker R, Kipphardt H, Töpfer A, Lehnik-Habrink P (2006) *Accred Qual Assur* (in press)
13. Guide to the expression of uncertainty in measurement (1993) ISO, Geneva, Switzerland