Current Topics in Analytical Chemistry
Vol. 14, 2022

Original Article

Using analytical quality by design principles in the analysis of petroleum compounds in the soil

Jelena Jurjeva[1,*] and Mihkel Koel[2]

[1]Eurofins Environment Testing Estonia OÜ, Paavli 5-3, 10412 Tallinn, Estonia;
[2]Department of Chemistry and Biotechnology, Tallinn University of Technology, Akadeemia tee 15, 12618 Tallinn, Estonia.

ABSTRACT

Analytical quality by design (AQbD) principles are well introduced in pharmaceutical industry and have big influence on analytical methods because of the high importance of quality control in this industry. These principles are adaptable in environmental analysis also and can be integrated into environmental quality management system. In the present work an attempt is made to use the principles of Quality-by-Design in the process of analysis of oil compounds in soil. The analytical target in this case is oil compounds that are spilled into soil, namely 16 parent polycyclic aromatic hydrocarbons (PAHs) and their alkylated analogs, 7 polychlorinated biphenyls (PCBs), and three aliphatic fractions from decane to pentatriacontane. The method used was gas-chromatography mass spectrometry (GC-MS) where the design of the experiment (DoE) was applied, in order to choose the optimal chromatographic conditions. The chemometric methods (Pearson product-moment correlation, Shapiro-Wilk test, and the principal component analysis) were applied for a comparison of methods used for analysing PAHs in soil, and the identification of the soil samples that were polluted with diesel oil and motor oil. The developed method was evaluated with respect to linearity, the limit of detection (LoD), and the limit of quantification (LoQ), within laboratory precision and the measurement uncertainty. Cross-validation was undertaken to compare the common laboratory method of PAH analysis and a novel analysis using the t-test. The Z scores were calculated when using the results of the proficiency test for 16 PAHs and 7 PCBs. A combined modified procedure based on ISO 18287:2006 for the soil quality, together with the SPIMFAB method as an instruction for analyses of the aromatics fraction > C16-C35, was proposed for the two matrices (soil and sediment). The results show that the recovery rates for the spiked samples were nearly 100% for both the soil and the sediment.

KEYWORDS: method validation, gas chromatography, experimental design, principal component analysis, analytical quality by design.

1. INTRODUCTION

The analytical quality by design (AQbD) approach contributes to achieving an improved method performance to ensure analytical procedures are well understood, fit for purpose, and robust. For development of the analytical process related to quality-by-design, it is important to follow Guideline Q8 (R2) on Quality by Design from the International Conference on Harmonization (ICH), which is defined as “a systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management” [1]. It emphasises that the method of analysis and the performance characteristics of the analytical process have to be

*Corresponding author: jurjeva.jelena@gmail.com

chosen to fulfil the specific objectives of the study. The process of developing an analytical method in a QbD environment contains several steps - starting with setting of analytical target (what, when and where to measure) followed by selection of proper method and related analytical procedure, accompanied with definition of necessary robustness of analytical procedure and design of experiments [2].

In environmental analysis, setting of analytical target profile depends on soil threshold values of contaminants established by authorities. Compliance control requires reliable and reproducible methods of sampling, sample pre-treatment prior to analysis, and analytical measurements to produce results that are valid for legal purposes. This raises the importance of the methods of validation [3, 4] that are natural part of QbD also.

Oil products are a complex mixture of several organic compounds, including aliphatic hydrocarbons (n-alkanes, iso-alkanes, and cycloalkanes), PAHs and their more toxic alkylated derivatives, and different polar compounds [5, 6]. Polychlorinated biphenyls (PCBs) comprise a class of 209 individual compounds. PCB concentrations are often reported as the sum of seven congeners (PCB7, IUPAC No. 28, 52, 101, 118, 138, 153, 180) or as "total PCB" [3, 7, 8]. The analysis of a group of strongly adsorbing non-volatile organic compounds (PCBs, PAHs) that are present in oil products were considered in this work.

In environmental analysis of dangerous substances in the soil and sediment matrices the high performance of instrumentation with very low detection limits is needed for accurate determination [9, 10]. Still, very often, the accurate determination of the alkylated PAHs in complex environmental samples is difficult because of the unresolved GC chromatograms [11]. Several studies on the quantification and identification of the PAHs [12, 13], PCBs [14], and the aliphatic and aromatic fractions [15] in the soil are available in the literature. In this study the analysis of 16 parent polycyclic aromatic hydrocarbons (PAHs) and their alkylated analogs, 7 polychlorinated biphenyls (PCBs), and three aliphatic fractions from decane to pentatriacontane in the soil and the sediment was undertaken using a combined method based on ISO 18287:2006, the SPIMFAB method, and the instruction for analyses of the aromatics fraction > C16-C35 [16, 17, 18]. In the present work the optimal chromatographic conditions were found through the full factorial design using three quantitative critical process parameters (CPPs). Also, this study aimed to establish a set of guidelines for the in-house validation of this kind of combined method because there is no information about a validated combined method for the analysis of all of these compounds together. The main emphasis was on assessing the parameters of validation, such as linearity, LoD, LoQ, accuracy, and precision. Additionally, the expanded uncertainty of the presented method was estimated when using the Nordtest approach.

The selection of chemometric and statistical methods (Pearson product-moment correlation (PPMC), Shapiro-Wilk test), was made to compare the different analytical methods for analyses of the soil samples that were polluted with the PAHs. It was also possible to group the samples according to the oil type, with the use of the principal component analysis (PCA).

To the best knowledge of the authors there are no studies analyzing oil compounds following the principles of AQbD. In this study AQbD was applied in three phases: the method design (DoE); the validation of the method; quality control (recovery tests, participation in inter-laboratory ring tests). The analytical target profile (ATP) in this study was set to develop a precise and accurate method to determine oil compounds in soil and sediment with GC-MS technique to support uncertainty estimation throughout the range of analyte concentrations and achieve the target LoQs. The peak areas (to increase sensitivity as much as possible) and the resolution between each peak were selected as the critical method attributes (CMAs) due to the most important ATPs. The strict threshold values established by authorities on the content of contaminants [19, 20] in the soil, the quantitative analysis of 35 compounds with different boiling points (from 174 ^{o}C for decane to 536 ^{o}C for indeno(1,2,3,c,d)pyrene), and possible interferences from the soil matrix make a choice in favor of the GC–MS for development of the method because of its high sensitivity, specificity, and reproducibility towards the analyzed compounds.

2. MATERIALS AND METHODS

The validation as described in this study was applied to the GC-MS method for the quantitative analysis of 35 compounds or groups of compounds, including 16 EPA PAHs (*naphthalene, acenaphthylene, acenaphthene, fluorene, phenanthrene, anthracene, fluoranthene, pyrene, benz(a)anthracene, chrysene, benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(a)pyrene, indeno (1,2,3,c,d)pyrene, dibenz(a,h)anthracene, benzo(g,h,i) perylene*), the alkylated PAHs *(methylnaphthalenes, dimethylnaphthalenes, trimethylnaphthalenes, 1,4,6,7-tetramethylnaphthalene, 1,2,5,6-tetramethylnaphthalene, 1,2-methylanthracene/-methylphenanthrene, methylpyrenes/fluoranthenes, methylchrysenes/benz(a)anthracenes)*, 7 PCBs (*PCB28, PCB52, PCB101, PCB118, PCB153, PCB138, PCB180*), three aliphatic fractions (*aliphatic >C_{10}-C_{12}, aliphatic >C_{12}-C_{16}, aliphatics >C_{16}-C_{35}*), and the biphenyl. Since the method was intended to quantify more than one analyte, each analyte was tested, in order to ensure that there was no interference from the matrix. The reported method validation data and the determination of accuracy and precision included all outliers that were possible from the experiment.

2.1. Materials and reagents

The stock solutions for the preparation of the calibration and the internal standards were purchased from AccuStandard, New Haven, USA, Sigma-Aldrich/Merck, Darmstadt, Germany, Dr. Ehrenstorfer GmbH Reference Materials, Augsburg, Germany, and Chiron Petroleum Reference Standards, Trondheim, Norway. Hexane was purchased from Honeywell International Inc., Charlotte, USA. The internal standards (ISTDs) of Naphthalene-D8, Pyrene-D10, and Perylene-D12 were used for the calculation of the concentrations of 16 PAHs, 7 PCBs, and the aromatic compounds. Ortho-Terphenyl was used for the calculation of the aliphatic compounds. The extraction solutions containing the standards were prepared in acetone with concentrations of 1000 µg/L for the PAH-ISTDs, and 15000 µg/L for O-Terphenyl.

For the validation of the method, two sets of calibration standards were prepared. The first one was for the PAHs and the PCBs, while the second one was for the aliphatic and aromatic compounds. The first calibration stock solution for the PCBs and the PAHs was prepared from PAH standard US-116N-1 (Agilent Technologies, Inc., Santa Clara, USA) and PCB Mix-3 (LGC Standards, Teddington, Middlesex, UK.). These two standards were mixed in hexane, and an ISTD solution was added to obtain PAH concentrations of 20 000 µg/L, PCB concentrations of 600 µg/L, and PAH ISTD concentrations of 2000 µg/L, in the final solution.

The second stock calibration solution for the aliphatic and aromatic compounds was prepared from Custom Mix Multi-Standard alkylated PAHs, aromatics in toluene (100 mg/L of each compound), and Calibration/Window Defining Hydrocarbon Standard (C_8-C_{40} 1000 mg/L of each compound). The Custom Mix Multi-Standard alkylated PAH (100 mg/l), the Window Defining Standard (1000 mg/l), and the ISTD solution were mixed in n-hexane. The aromatic compound concentration in the solution was 2000 µg/ml and the aliphatic concentration was 50000 µg/L. The concentration for the PAH ISTDs was 2000 µg/L and for O-Terphenyl, it was 30000 µg/L. All of the PAHs/PCBs and the aliphatic/aromatic calibration standards were prepared by a dilution of the stock calibration solution at different concentration levels. The calibration standards and the sample extract were mixed in hexane.

2.2. GC method selection and parameters optimisation

The finding of the optimal chromatographic conditions for the determination of the PAHs, the PCBs, the aliphatics, and the aromatics by GC-MS started with the choice of the GC column (from two options - Rxi-XLB and Rxi-17Sil). In the preliminary tests the PAHs/PCBs and the aliphatic/aromatic calibration standards were injected into both columns. The Rxi-17Sil column showed the best separation of the co-eluted PAHs (for instance, benzo(b)fluoranthene, benzo(k)fluoranthene, and benzo(j)fluoranthene) but Rxi-XLB had the highest responses for the aliphatic compounds. The last-mentioned column was chosen for further validation as the separation between benzo(b)fluoranthene, benzo(k)fluoranthene, and benzo(j)fluoranthene was not crucial in this study.

The temperature parameters, namely, the initial oven temperature, the final oven temperature, and

the final inlet temperature mostly influence the recovery of the aliphatic compounds in chromatographic analysis. The DoE was applied to achieve the optimum chromatographic conditions through the full factorial design with the above-mentioned three quantitative critical process parameters (CPPs) and optimize the ratio between pentatriacontane and decane (C35/C10).

2.3. Method performance testing and validation

The within-laboratory validation was performed according to Eurachem, the GUM Guide, and the Nordtest Guide [4, 21, 22]. The following parameters were evaluated: linearity, LoQ, LoD, accuracy, precision, and the measurement uncertainties for each component. The effect of the soil matrix was also evaluated [23].

2.3.1. Linearity and the working range

In order to check the linearity of the calibration curve, a set of calibration standards at concentrations of 10, 50, 500, 2000, 5000, 10000, and 20000 µg/L was prepared for the calculation of the PAHs and at concentrations of 0.3, 3, 15, 60, 150, and at 600 µg/L for the PCBs. The linearity of the aliphatic compounds was checked in the ranges of 5000 to 50000 µg/L for the alkylated PAHs (aromatic compounds), and in the ranges of 200 - 5000 µg/L for the individual compound. All measurements were performed in triplicate, in order to check the homoscedasticity and the heteroscedasticity of the data.

2.3.2. Determination of the LoD and the LoQ

The LoD and the LoQ were determined separately for each matrix (soil and sediment) using the following equations (Eq. (1, 2)):

$$LoD = Xavg+3*STDEV \quad (1)$$

$$LoQ = Xavg+10*STDEV \quad (2)$$

where *Xavg* is the average concentration of the replicates, and *STDEV* is the standard deviation of 10 replicate samples.

2.3.3. Accuracy, precision, and the measurement of the uncertainties estimation

Two approaches were used during the validation for the determination of bias and the measurement uncertainty: the analysis of the CRMs and the recovery tests (standard addition to a sample), as suggested by the Nordtest guide. The minimum number of replicates used for the calculation of precision and bias was 10 for each matrix. The recovery experiments were performed by comparing the analytical results for the extracted samples at three concentration levels (low, medium, and high), together with the unextracted standards that represented 100% recovery [22]. For the calculation of the expanded measurement uncertainty, the coverage factor k = 2 was used, providing a level of confidence of 95%.

2.4. Statistical methods used in validation

2.4.1. The lack-of-fit test

A linear regression model was expected for all of the compounds in this work. The relationship between x (the concentration of the analyte) and y (the instrument response) was a straight line $y = b_1x + b_0$, where b_0 was the y-intercept and b_1 was the slope of the line. The three replicates of the five expected levels of concentration values were measured by GC-MS for all of the compounds that were included in the method.

The validation of the linear calibration model was performed by the lack-of-fit test, in conjunction with a residual plot [24, 25]. The F-statistic was used to test the null hypothesis that the linear model was correct when used. The F value was obtained and the probability (P) associated with the F statistic was calculated using the Excel 2010 function =F.INV.RT (0.05 probability; DoF (degree of freedom) (numerator, n-2); DoF (denominator, n*(p-1)).

2.4.2. T-test statistics and null hypothesis

The main purpose of the validation was to demonstrate that the candidate method gave PAHs results that were equivalent to an existing method (the old laboratory method as described in [26]). The t-test was used to evaluate whether there was a statistically significant difference between the mean value for a series of determinations and the accepted reference value [27]. The sample used for the t-test was a real soil sample, which was analyzed 3 times with the old method and 3 times with the new method. For the t-test, 16 EPA PAH differences (the replicates average results) were calculated.

The t-test had the form of (Eq. (3)): the mean value of the differences, *Xavg dif*, divided by the mean value of the deviation of the differences, *STDEV dif*, multiplied by the square root of the resulting pairs.

$$t = \frac{Xavg\ dif}{STDEV\ dif} * \sqrt{n} \quad (3)$$

2.4.3. One-Way ANOVA testing

One-Way ANOVA was used to check the reproducibility of the methods. Three different chemists performed the analyses on different days. *F*, the critical *F*, and the *P* values allowed for direct conclusions to be drawn on whether the variations between the results obtained by the different chemists were significantly greater than the variation in the results that were obtained by one chemist.

2.4.4. Z-Score in the chemical proficiency testing

The final step during the validation of the analytical method was the evaluation of the performance characteristics through inter-laboratory ring tests, namely, the proficiency tests, in order to demonstrate that the method was fit for purpose. The results of the laboratories were assessed by converting them to a Z-score using the following equation (Eq. (4)):

$$Z = \frac{Xi - Xa}{\sigma} \quad (4)$$

where *Xi* is the result of the laboratory *i*, *X*a is the assigned reference value, and σ is the assigned standard deviation for the proficiency test. The Z-Scores were typically interpreted as $Z \leq 2$, which was acceptable, but questionable outside the range of ±2, and actionable outside the range of ±3 [28, 29]. The reported value (laboratory result) was compared with the assigned value (the value used to assess the proficiency), and the Z scores for the 16 PAHs and 7 PCBs were calculated by the organizer of the proficiency test [30].

3. RESULTS AND DISCUSSION

3.1. Optimisation of chromatographic conditions

The DoE was performed by using the Custom Design Platform in JMP software [31]. The design used in this work consisted of three factors or CPPs, each at three levels (3^3 design). Table 1 displays the design matrix, enlisting the selected factors during

Table 1. The 3^3 design matrix.

		Factor A (Initial Oven Temperature)		
Factor B (Final Oven Temperature)	**Factor C (Final Inlet Temperature)**	**0**	**1**	**2**
0	0	000	100	200
0	1	001	101	201
0	2	002	102	202
1	0	010	110	210
1	1	011	111	211
1	2	012	112	212
2	0	020	120	220
2	1	021	121	221
2	2	022	122	222
CPPs		**Levels of Factor Studied**		
		Low (0)	**Intermediate (1)**	**High (2)**
Initial Oven Temperature		40	50	60
Final Oven Temperature		300	320	340
Final Inlet Temperature		300	310	320

the screening studies, along with their respective low (0), intermediate (1), and high (2) levels.

For this model, 21 experimental runs were conducted in total. A standard concentration of 25 mg/L for each of the aliphatic compounds was used for all of the runs. They were analyzed for the critical analytical attributes (CAAs), namely, the C35/C10 ratios (response of C35 to the response of the C10 peak).

The influence of the second-order interactions between the initial and the final temperatures of the oven, the oven and the inlet final temperatures, and the oven initial and the inlet final temperatures were checked. The last two interactions (Table 2) were not statistically significant (p-value > 0.05) and they were removed from the model.

Fig. 1 shows the influence of the temperatures on the change of the ratios C35/C10. There was no difference in the ratios when the final temperature of the oven was changed from 300 ºC and 340 ºC. However, when the inlet final temperature increased from 300 ºC to 320 ºC, the ratios of C35/C10 changed significantly (the slope of the design curve increased).

The optimal chromatographic conditions were chosen based on the ratios of C35/C10 that were close to the value of 1. The optimized chromatographic solution was observed at an initial oven temperature of 40 ºC, with a final oven temperature of 340 ºC, and final inlet temperature of 320 ºC.

3.2. Analytical parameters of the method

The optimized method was ratified for linearity, LoD, LoQ, precision, accuracy, selectivity, and measurement uncertainty, according to the guidelines as recommended by the Nordtest Guide, the GUM Guide, and Eurachem.

3.2.1. The reproducibility, the accuracy, and the uncertainty assessment

The within-laboratory reproducibility, the accuracy, and the uncertainty were validated for the two matrixes. For the sediment matrix, the values of

Table 2. Summary of the model effects.

CPPs	p-value
Final inlet temperature	0.00221
Initial oven temperature*Final oven temperature	0.00662
Initial oven temperature	0.01329
Final oven temperature	0.04872
Final oven temperature*Final inlet temperature	0.11643
Initial oven temperature*Final inlet temperature	0.29325

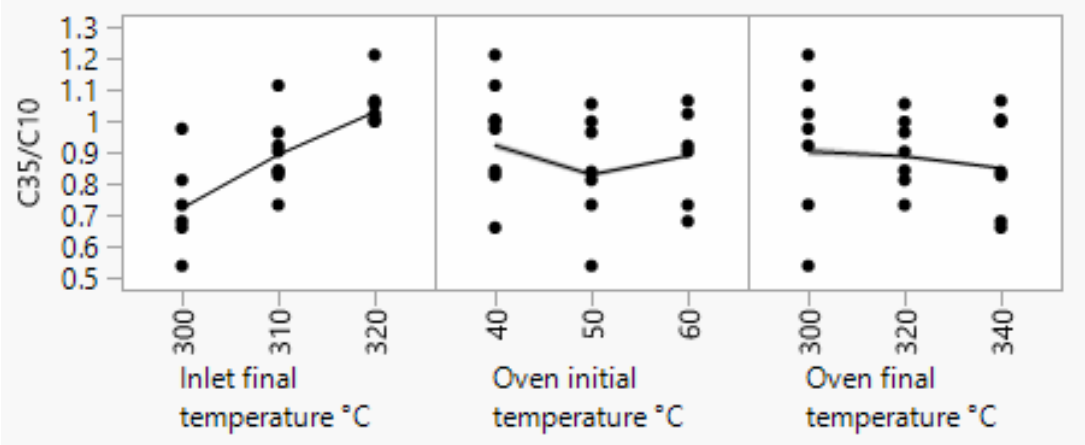

Fig. 1. The influence of temperatures on the change of the ratios of C35/C10.

the parameters were higher than they were for the soil. The higher values for the analyses of the oil compounds in the sediment met the expectations, as this matrix had more complicated mineral structures than did the soil.

For most of the compounds, the expanded measurement uncertainty was less than 25%. The outliers (the calculated uncertainty that was more than 25%) were for methyl-pyrene/fluoranthene (26%), methyl-chrysene/benz(a)anthracene (27%), and PCB 153 (26%). The reproducibility of the measurements for the PCBs and the PAHs was less than 10%. The aliphatic fractions and the alkylated PAHs demonstrated higher differences between the replicate measurements due to the interfering compounds in the matrix.

3.2.2. The linearity assessment

The linearity method was assessed by a visual inspection of the plot and that was supported by the statistics (the lack-of-fit test) and the residuals

Table 3. The target and the calculated LoQs for the PCBs, the PAHs, and the aliphatic compounds.

Compound	Target LoQ mg/kg	Calculated LoQ mg/kg
Naphthalene	<0.030	0.015
Acenaphthylene	<0.030	0.014
Acenaphthene	<0.030	0.016
Fluorene	<0.030	0.018
Phenanthrene	<0.030	0.014
Anthracene	<0.030	0.016
Fluoranthene	<0.030	0.014
Pyrene	<0.030	0.015
Benz(a)anthracene	<0.030	0.019
Chrysene	<0.030	0.017
Benzo(b)fluoranthene	<0.030	0.019
Benzo(k)fluoranthene	<0.030	0.02
Benzo(a)pyrene	<0.030	0.018
Indeno(1,2,3,c,d)pyrene	<0.030	0.014
Dibenz(a,h)anthracene	<0.030	0.016
Benzo(g,h,i)perylene	<0.030	0.016
Ali>C10-C12	<5	2
Ali>C12-C16	<5	2
Ali>C16-C35	<10	9
PCB28	<0.002	0.001
PCB52	<0.002	0.001
PCB101	<0.002	0.002
PCB118	<0.002	0.002
PCB153	<0.002	0.002
PCB138	<0.002	0.002
PCB180	<0.002	0.002
Methyl-pyr/fluorant	<0.5	0.2
Methyl-chry/benz(a)a	<0.5	0.1

plot from the linear regression. Based on that, for each analyte, the calculations of concentrations were performed according to the linear calibration models. A linear calibration curve was confirmed in the range of 50-5000 µg/L for all of the parent PAHs; 400-5000 µg/L for the alkylated PAHs compounds; 15-600 µg/L for the PCBs, and 10000-50000 µg/L for the aliphatic compounds, with a coefficient of determination $R^2 \geq 0.998$. The expected concentrations and the obtained concentrations did not differ by more than 20%. The lowest calibration point was equal to the LoQ. At this level, each analyte was determined with an acceptable uncertainty.

3.2.3. The LoD and the LoQ assessment

The 10 blank samples were spiked with analytes below the expected LoD. The target (the values that were the requirement of the client) and the calculated LoQs are shown in Table 3.

3.2.4. Recovery tests

The blank matrix samples were spiked with known concentrations of the PAHs (50, 2000, 5000 µg/L) in the range of the calibration curve. At each level, the samples were prepared in triplicate. The criterion of recovery was selected according to the calculated measurements of uncertainty, which meant that the recovery must fall between 80% and 120%, in order to be acceptable (2/3 of the calculated MU 30%). In the present work, all of the compounds exhibited a recovery within the allowed range. Table 4 shows the recoveries in the soil and in the sediment at a concentration of 2000 µg/L.

3.3. Comparision of methods

The two PAH laboratory methods (the developed and the laboratory reference method [26]) were compared by using the t-test statistic. The mean value of the differences between the 16 PAHs was 0.0041 and the deviation was 0.8629. The t-test gave a reading of t = 0.02. When it was compared to the table value of 2.13, with a freedom level of 15 (16 PAHs -1) and 95% of probability, there were no differences in the results between the methods.

One-Way ANOVA was used for the comparison of the results by 3 different chemists, in order to control whether the variation between the results that were obtained on behalf of the different chemists was significantly greater than the variation

Table 4. The PAH recovery test for the soil and the sediment matrix.

Compound	Recovery % (in the soil)	Recovery % (in the sediment)
Naphthalene	93	93
Acenaphthylene	90	92
Acenaphthene	92	97
Fluorene	99	101
Phenanthrene	102	95
Anthracene	99	98
Fluoranthene	97	89
Pyrene	94	88
Benz(a)anthracene	97	98
Chrysene	96	96
Benzo(b)fluoranthene	110	102
Benzo(k)fluoranthene	99	104
Benzo(a)pyrene	112	112
Indeno(1,2,3,c,d)pyrene	105	102
Dibenz(a,h)anthracene	102	108
Benzo(g,h,i)perylene	103	101

in the results that were obtained by one chemist. Table 5 shows an example of the ANOVA results and there were no differences between the analysts in the analyses of the PAHs and the PCBs (Fcrit > F).

The proficiency test results for the 16 PAHs and the 7 PCBs were satisfactory. The Z-scores for the 14 PAHs and the 7 PCBs were in the range of between -1.5 to 1.7 (acceptable), and for the 2 PAHs, namely, Chrysene and Benzo(b)fluoranthene the Z score was 2.8, which is questionable. The separation between benzo(b)fluoranthene and benzo(k) fluoranthene was insufficient. Benzo(j)fluoranthene, which was not part of the EPA PAHs, eluted between the two analytes, and this could lead to higher concentrations than expected. The known interferences for the determination of chrysene indicated that triphenylene was much more resonance stable than its isomers. In those cases, where interferences cannot be avoided, the sum of the individual compounds can be reported, for instance, triphenylene/chrysene and benzo(b,j,k) fluoranthene.

The analysis of two methods showed comparable results. The major advantage of the developed method is that extraction time for sample preparation is less than 1 hour compared to 16 hours for reference method.

3.4. Comparison of developed methods with other methods

Table 6 shows the comparison of the validated method (Method 1) with previous studies from literature [9, 10, 12, 13, 26, 32-35]. There are no methods where the simultaneous analyses of the PAHs, the PCBs, the aliphatics, and the aromatics compounds are performed, and this way the comparison was performed for the PAHs' analysis methods only. The 10 different GC-MS, GC-FID, HPLC methods, with a different extraction type (Soxhlet extraction, sonication or ultrasonic treatment, mechanical agitation, accelerated solvent extraction (ASE), solid-phase microextraction (SPM)), and the pretreatment steps (sample clean up, filtration, concentration) were all compared. Method 2 was the old laboratory method, as described in [26].

Table 5. The F statistical, the p-value, and the F critical calculation when using One-Way ANOVA.

Source of Variation	SS	df	MS	F	p-value	Fcrit
Between Groups	1.7616	3	0.5872	0.9625	0.4359	3.2874
Within Groups	9.1506	15	0.6100			
Total	10.9122	18				

Table 6. A comparison of the validated method (Method 1) with the other 9 methods.

Method No	PCA Number	Equipment	Extraction Type	Clean-up	Filtration	Concentration
Method 1	Sit1	GC-MS	Mechanical	No	No	No
Method 2	Sit2	GC-MS	Mechanical	No	No	No
Method 3	Sit3	GC-MS	ASE	SPE	No	Yes
Method 4	Sit4	HPLC	Sonication	Silica gel	Yes	Yes
Method 5	Sit5	GC-MS	Soxhlet	Alumina/ Silica Gel	No	Yes
Method 6	Sit6	GC-MS	Sonication	Copper/Alumina/ Silica Gel	Yes	Yes
Method 7	Sit7	GC-FID	SPM	Sorbent	No	No
Method 8	Sit8	GC-MS	Sonication	Copper	No	Yes
Method 9	Sit9	HPLC	Sonication	No	No	Yes
Method 10	Sit10	GC-MS	Mechanical	No	No	No

The comparison of the different methods was conducted, in order to find the influence of the extraction type, the clean-up, the filtration, and the concentration procedure on the following parameters: extraction time in min (Time), the number of method preparation steps (Stages), the minimum and the maximum values that were obtained for LoQ when using these methods (LoQ.min and LoQ.max), the average repeatability of the method in %, the average recovery of the control samples, the weight of the sample (in grams), the volume of the extract solutions, the number of extracts, and the internal standards (ISTDs). Fig. 2 shows the score and the loading plots of the 10 methods that were used for the comparison, as well as the parameters that influenced the grouping of the methods. The score and the loading plot helped to find the correlation between the parameters.

The extraction time contributed the most to the grouping of the methods. For Method 5 (Soxhlet extraction), the volume of the extract (120 ml when compared with the average value of 35 ml) influenced the most; for Method 10 and 2 (mechanical extraction), the extraction time 720 and 960 min influenced the most, respectively; for Method 9 (Sonication), maximum value was obtained for LoQ; for Method 7 and 8 (SPM and Sonication, respectively), the recovery of the control samples was more than 110%. The validated method (Method 1) did not differ too much from Methods 3 and 6. All of these 3 methods were used for the GC-MS analysis. Acetone and hexane were used as the extraction solvent. The extraction time for these methods was less than 60 min, the volume of the extract was less than 50 ml, and the repeatability was less than 12%. The validated method had the highest LoQs for the PAHs (0.014 - 0.020) when compared with 0.00007 - 0.00017 (Method 3), and with 0.0014 - 0.0119 mg/kg (Method 6). The validated method was used for the

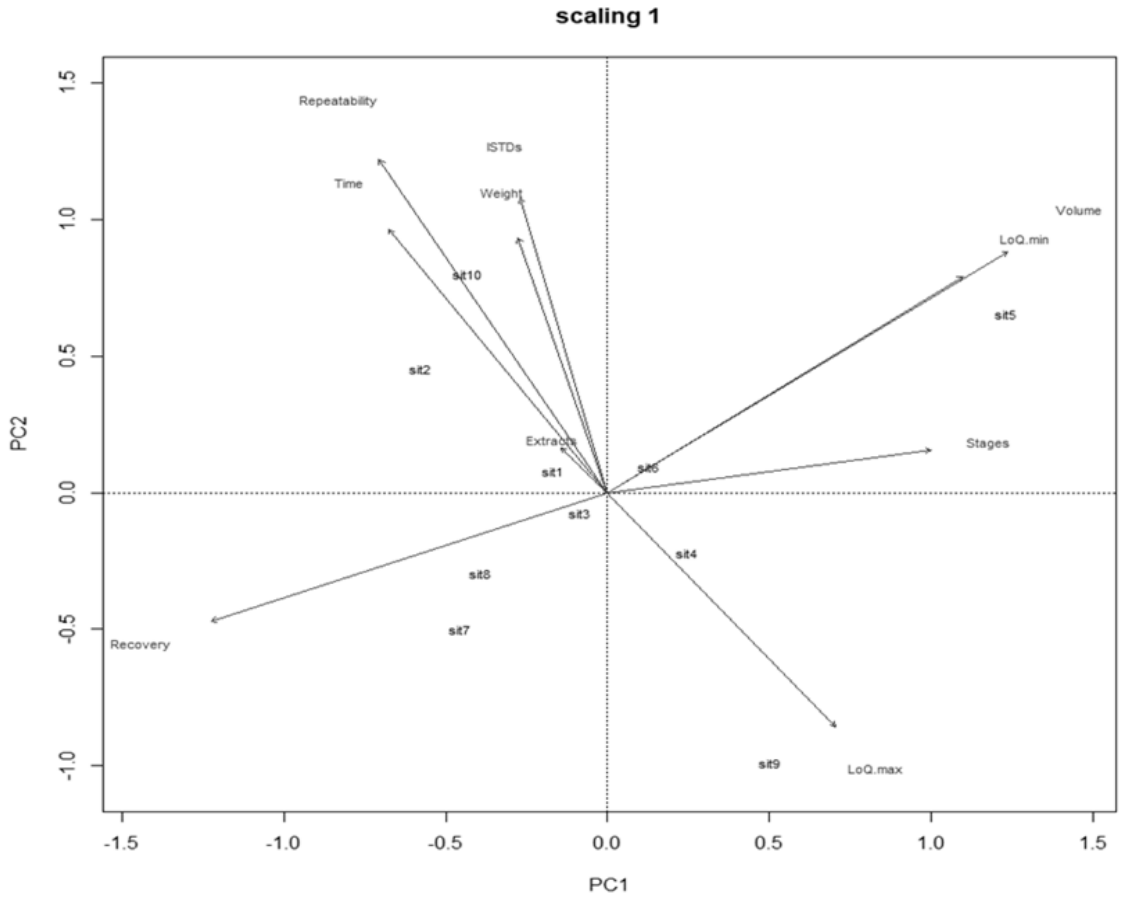

Fig. 2. The scaling plot of the 10 methods and the 10 parameters that were used for the grouping of the methods.

mechanical extraction, Method 3 for ASE and for ultrasonic Method 6 was used. According to some research [36], ultrasonic agitation has shown low recoveries for the lower molecular weight PAHs, like naphthalene, acenaphthene, and acenaphthene, and less efficiency when compared with, for example, Soxhlet extraction. The mechanical extraction method did not need the use of expensive equipment, compared to microwave extraction and ASE. When the extraction time was less than 1 hour for the validated method, it did not need an extra step for the purification of the samples.

The Pearson product-moment correlation (PPMC) coefficients between the parameters were calculated with the use of the R statistic, in order to find out how strong a relationship was between the parameters. The Shapiro-Wilk test was used to calculate the p values and to test the null-hypothesis. A high degree of positive correlation (from 0.5 to 1) was obtained for Weight and Repeatability (0.5364); Stages and Volume (0.6467); ISTDs and Time (0.6255); and Volume and LoQ min (0.8403). A high negative correlation (from -0.5 to -1) was obtained for Volume and Recovery (-0.7857); and Recovery and LoQ min (-0.6534). No correlation (less than ±0.1000) was found for Weight and ISTDs (0.0450); Weight and LoQ min (0.0706); the number of Extracts and Volume (-0.0714); the number of Extracts and Recovery (-0.0335); and the number of ISTDs and Recovery (-0.0676).

Fig. 3 shows the linear relationship between the two sets of data and the plot with a normal distribution error: Volume vs. Recovery (r = -0.7857 and p-value = 0.9712), Volume vs. LoQ min (r = 0.8403

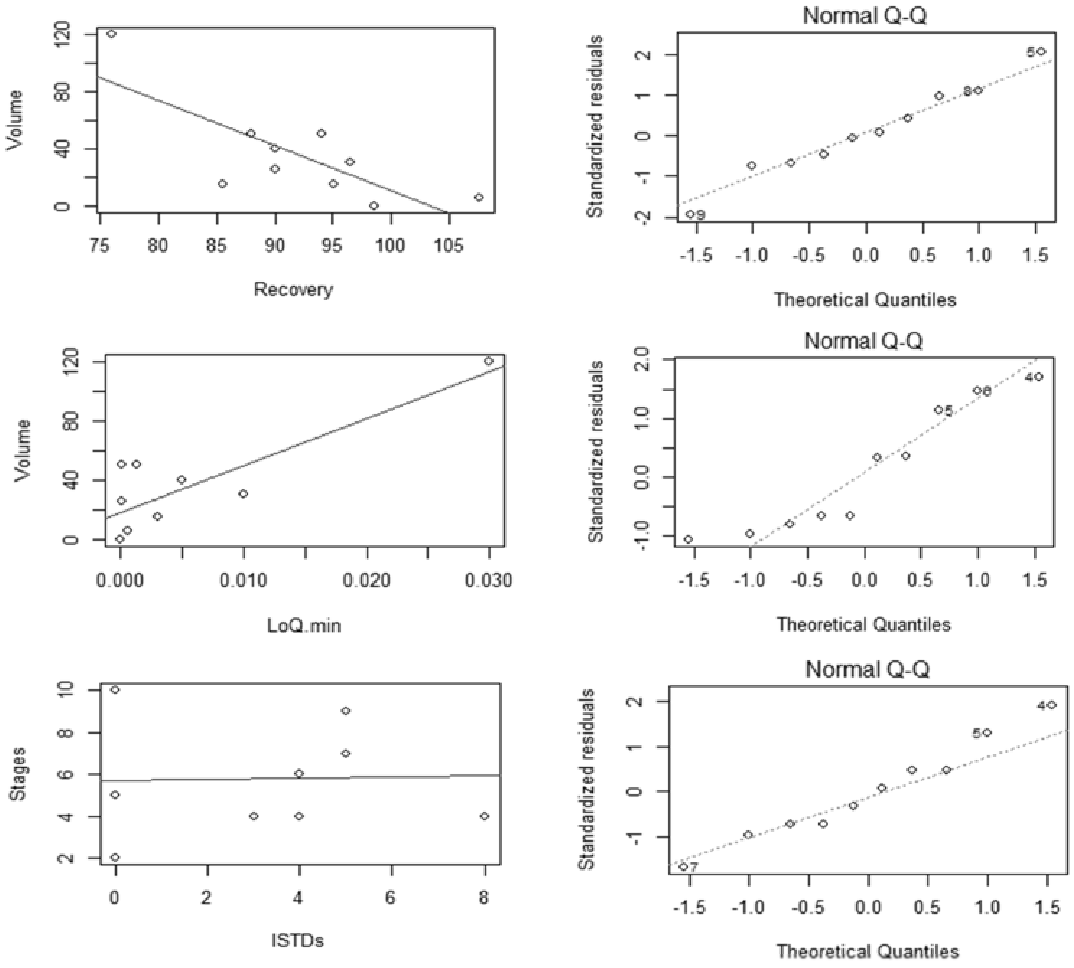

Fig. 3. The linear relationship between the two sets of data and the plot with a normal distribution error.

and p-value = 0.09836), Stages vs. ISTDs (r=0.030 and p-value = 0.767). If the volume of the extract increased the recovery of the PAHs decreased, together with a decrease in the volume of the extract, and a lower LoQ was obtained. The scaling plot (Fig. 2) concentrates on the associations between the variables (parameters). The variables that are positively correlated have arrows in the same direction (Volume vs. LoQ min), while the variables that are negatively correlated have arrows in opposite direction (Volume vs. Recovery). Fig. 3 (Stages vs. ISTDs) shows an example where the correlation between the parameters is absent.

The linear relationship between the volume and the minimum LoQ was significant because the correlation coefficient was significantly different from 0 and the line may be used for a prediction of the model. In the developed method extraction volume was 30 ml and minimum LoQ was 0.014 mg/kg; laborotory reference method has extraction volume 15 ml with LoQ 0.003 mg/kg (the number of pretreatment stage for both method is four, and the number of ISTDs is 3 and 8 respectively). Comparing the threshold values established by authorities on the content of contaminants [19, 20] in the soil, our developed method was sensitive enough.

3.5. The oil type identification with the PCA

The developed method was applied for controlled testing of the soil and the sediment samples. The 7 diagnostic ratios (Ant/Phe, Fla/Pyr, Ant/(Ant + Phe), Fla/(Fla + Pyr), Ind/(Ind + B(g,h,i)P, B(a)Ant/(B(a)Ant + Chr), HMW/LMW) and PCA were used for the identification of the oil source in 45 samples that arrived from different locations of Sweden and Norway. The PCA method and the PAH ratio calculations are described in a previous study [26]. In this study, 3 aliphatic fraction ratios C_{10}-C_{12}/total aliphatic, C_{12}-C_{16}/total aliphatic, and C_{16}-C_{35}/total aliphatic were also included in the PCA. The total of the aliphatics was all of the aliphatic compounds between decane and pentatriacontane that were analyzed with GC-MS and they had an ion target of m/z = 71. In the control samples that were polluted with fresh oils (diesel oil and motor oil), the concentrations of the PAHs that were compared with the aliphatic compounds were very small.

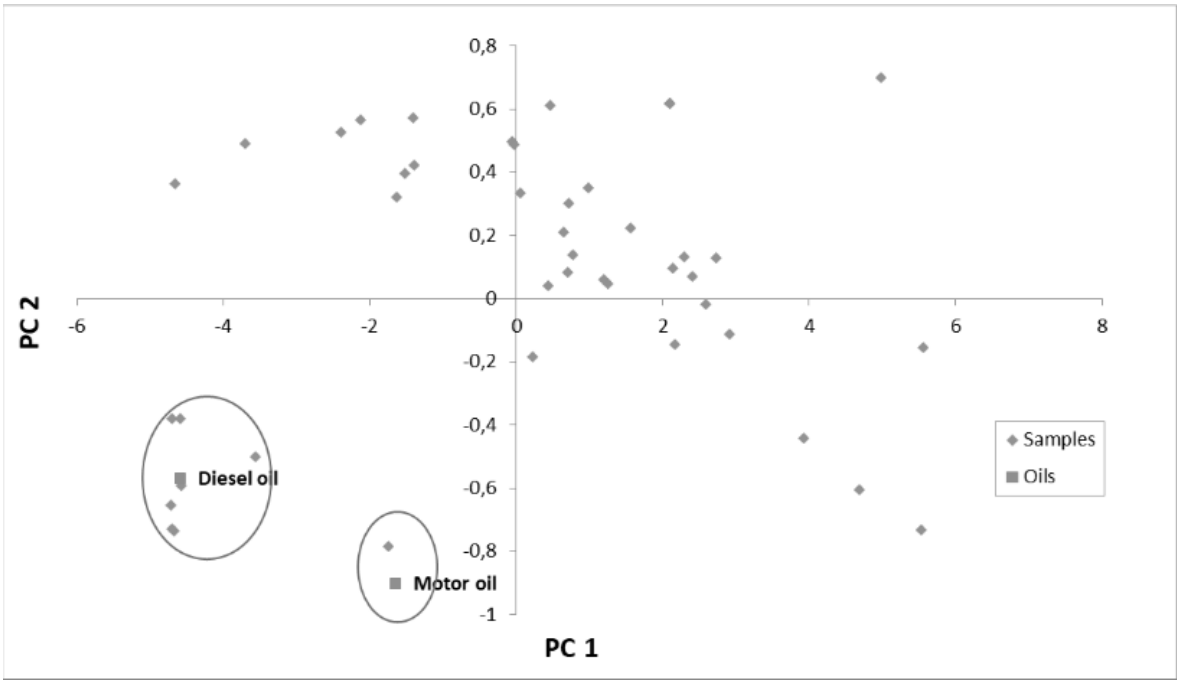

Fig. 4. The score plot of the binary ratios of the 45 samples and the control samples that were polluted with diesel oil and motor oil.

Table 7. The binary ratios of the samples that were polluted with diesel oil and motor oil.

Ratio	Diesel Oil	Motor Oil
Flu/Pyr	0.0554	0.0373
Ant/(Ant+Phe)	0.0373	0.0411
Flu/(Pyr+Flu)	0.0525	0.0359
HMW/LMW	0.0652	3.0787
Al 10-12/Al total	0.2444	0.0015
Al 12-16/Al total	0.4175	0.0018
Al 16-35/Al total	0.3381	0.9967

The score plot of the first and the second PCs (Fig. 4) shows 45 samples from different locations in Sweden and Norway. The score plot reveals that 6 samples were polluted with diesel oil and 1 sample with motor oil. The most influenced ratios for the classification of the samples into groups were HMW/LMW and 3 aliphatic ratios; benz(a) anthracene, chrysene, indeno(1,2,3,c,d)pyrene, and benzo(g,h,i)perylene were not detected in the samples that were polluted with diesel oil and motor oil. Table 7 shows the binary ratios for the control samples that were polluted with motor oil and diesel oil.

The HMW/LMW ratios for the diesel samples were in the range 0.043-0.083. The Ant/(Ant + Phe) were in the ratios of 0.011-0.042. In the fresh diesel oil samples, the ratios were 0.065 and 0.037, respectively. The aliphatic fraction percentages for the diesel oil were approximately 24% C10-C12, 42% C12-C16, and 34% C16-C35. In motor oil, the percentages were 0.15%, 0.18%, and 99.67% for the same fractions. The HMW/LMW ratio in the motor oil sample was approximately 3.0 and in the analyzed soil sample that was polluted with this type of oil, the ratio was 2.99.

Most of the samples were not similar to the analyzed standard oils. In most of the samples, all 16 PAHs and the samples were probably polluted by mixtures of different oils, or by the weathered oils, which were difficult to identify without oil standards and a history background about the location place of the samples.

4. CONCLUSIONS

The AQbD approach was applied for the simultaneous identification and the quantification of the 16 EPA PAHs, the 7 PCBs, the alkylated PAHs, and the aliphatic compounds in the soil by the GC-MS analytical technique with mechanical extraction method for sample preparation. The design of the experiment helped to find the optimal GC conditions for the method. With a proper GC column, the oven program was possible to get a sufficient resolution and recovery. Another advantage of this method was the non requirement of an extra purification step of the samples before GC analysis and this made it simpler and fast (the extraction time was less than 1 hour). The reduction of the different pretreatment steps on a single sample and reduced amount of the consumables make this method greener. The validation of the developed method demonstrated that it could be used for controlled testing of the soil and the sediment. Satisfactory accuracy, precision, linearity, and recovery rates in the investigated concentration range were achieved. Linear calibration curves with sufficient correlation coefficients (>0.998) were obtained. The obtained results of the inter-laboratory comparison (Z-scores) confirmed that the method developed in this study was appropriate for the analyses of the PAHs, the PCBs, the alkylated PAHs, and the aliphatic compounds in the soil and the sediment. The GC-MS data were very good source for further chemometric approach. The PCA analyses helped to identify the soil samples that were polluted with diesel oil and motor oil; the PPMC coefficient helped to find the relationship between the parameters, in order to predict the linear model and to compare the different methods.

ACKNOWLEDGMENTS

This research was supported by the Estonian Center of Analytical Chemistry (ECAC) funded by the Estonian Research Council (TT4).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES

1. International Conference on Harmonization (ICH), Tripartite guidelines,' ICH Q8 (R2): Pharmaceutical Development', 2009, London.

2. Tome, T., Žigart, N., Časar, Z. and Obreza, A. 2019, Org. Process Res. Dev., 23, 1784-1802.
3. Kalbe, U., Lehnik‑Habrink, P., Bandow, N. and Sauer, A. 2019, Environ. Sci. Eur., 31, 29. https://doi.org/10.1186/s12302-019-0211-3.
4. Eurachem Guide, The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics. Second Edition, 2014.
5. Bayona, J. M., Domínguez, C. and Albaigés, J. 2015, Trends in Environmental Analytical Chemistry 5, 26-34. https://doi.org/10.1016/j.teac.2015.01.004.
6. Ali, M. F. and Abbas, S. 2006, Fuel Processing Technology, 87(7), 573-584. https://doi.org/10.1016/j.fuproc.2006.03.001
7. Choi, H. M., Veriansyah, B., Kim, J., Kim, J. D. and Lee, Y. W. 2009, Journal of Environmental Science and Health Part A, 44, 494-501. https://doi.org/10.1080/10934520902719936.
8. Buah-Kwofie, A., Yeboah, P. O. and Pwamang, J. 2011, Chemosphere 82(1), 103-106. https://doi.org/10.1016/j.chemosphere.2010.09.063.
9. Han, M., Kong, J., Yuan, J., He, H., Hu, J., Yang, S., Li, S., Zhang, L. and Sun, C. 2019, Talanta 205. https://doi.org/10.1016/j.talanta.2019.120128.
10. Kumar, B., Verma, V. K., Gaur, R., Kumar, S., Sharma, C. S. and Akolkar, A. B. 2014, Advances in Applied Science Research, 5(1), 201-209.
11. Zhao, Y., Hong, B., Fan, Y., Wen, M. and Han, X. 2014, Ecotoxicology and Environmental Safety, 100, 242-250. https://doi.org/10.1016/j.ecoenv.2013.10.018.
12. Yamada, T. M., Souza, D. A., Morais, C. R. and Mozeto, A. A. 2009, Journal of Chromatographic Science, 47(9), 794-799. https://doi.org/10.1093/chromsci/47.9.794.
13. https://www.agilent.com/Library/applications/5990-8414EN.pdf, (accessed December 2022).
14. Krzemień-Konieczka, I. and Buszewski, B. 2015, Pol. J. Environ. Stud., 24(5), 2029-2033. https://doi.org/10.15244/pjoes/41589.
15. Adeniji, A. O., Okoh, O. O. and Okoh, A. I. 2017, Journal of Chemistry. https://doi.org/10.1155/2017/5178937.
16. ISO 18287:2006. Soil quality - Determination of polycyclic aromatic hydrocarbons (PAH) - Gas chromatographic method with mass spectrometric detection (GC-MS), 2006.
17. https://www.ivl.se/download/18.2aa26978160972788071d154/1530276751684/C305.pdf, (accessed December 2022).
1 8. https://docplayer.se/4764417-Instruktion-for-analys-av-fraktionen-aromater-c16-c35.html, (accessed December 2022).
19. https://www.riigiteataja.ee/akt/731616, (accessed December 2022).
20. https://www.miljodirektoratet.no/globalassets/publikasjoner/klif2/publikasjoner/andre/1691/ta1691.pdf, (accessed December 2022).
21. GUM JCGM 100:2008 Evaluation of measurement data — Guide to the expression of uncertainty in measurement, first edition, 2008.
22. Nordtest project 1589-02. Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories. Version 3.1, 2012.
23. Zhou, W., Yang, S. and Wang, P. G. 2017, Bioanalysis, 9, 23. https://doi.org/10.4155/bio-2017-0214.
24. Moosavi, S. M. and Ghassabian, S. 2018, IntechOpen. https://doi.org/10.5772/intechopen.72932.
25. Thompson, M., Ellison, S. L. R. and Wood, R. 2002, Pure Appl. Chem., 74(5), 835-855.
26. Jurjeva, J. and Koel, M. 2019, Oil Shale, 36(3), 410-430. https://doi.org/10.3176/oil.2019.3.04.
27. Cherry, S. R., Sorenson, J. A. and Phelps, M. E. Nuclear Counting Statistics in: Physics in Nuclear Medicine, Fourth Edition, 2012, 125-140.
28. Hibbert, D. B. Quality assurance for the analytical chemistry laboratory, Oxford University Press, 2007.
29. Analytical Methods Committee. 2016, Anal. Methods 28. https://doi.org/10.1039/c6ay90078j.
30. https://www.sigmaaldrich.com/analytical-chromatography/proficiency-testing.html, (accessed December 2022).
31. https://www.jmp.com/en_us/home.html, (accessed December 2022).

32. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/755603/pahinsoil.pdf, (accessed December 2022).
33. Dalvand, K. and Ghiasvand, A. 2019, Analytica Chimica Acta, 1083, 119-129. https://doi.org/10.1016/j.aca.2019.07.063.
34. Dong, C. D, Chen, C. F. and Chen, C. W. 2012, Int. J. Environ. Res. Public Health, 9(6), 2175-2188. https://doi.org/10.3390/ijerph9062175.
35. https://cdnmedia.eurofins.com/Microsites/media/1148/metode_4_2008_2_udg.pdf, 2008 (accessed December 2022).
36. Lau, E. V., Gan, S. and Ng, H. K. 2010, International Journal of Analytical Chemistry. https://doi.org/10.1155/2010/398381.