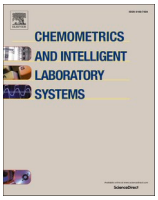
Chemometrics and Intelligent Laboratory Systems 243 (2023) 105008

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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



Analytical quality by design using a D-optimal design and parallel factor analysis in an automatic solid phase extraction system coupled to liquid chromatography. Determination of nine PAHs in coffee samples

L. Valverde-Som [a], M.M. Arce [a], L.A. Sarabia [b], M.C. Ortiz [a,*]

[a] Dep. Química, Universidad de Burgos, Facultad de Ciencias, Plaza Misael Bañuelos s/n, 09001, Burgos, Spain
[b] Dep. Matemáticas y Computación, Universidad de Burgos, Facultad de Ciencias, Plaza Misael Bañuelos s/n, 09001, Burgos, Spain

ARTICLE INFO

Keywords:
Analytical quality by design
D-optimal design
SPE
HPLC-FLD
Polycyclic aromatic hydrocarbon
Coffee

ABSTRACT

Optimizing a multi-residue analysis when using an automatic SPE (solid phase extraction) system and complex matrices becomes a difficult problem because of the large number of experimental factors that can influence the recovery of the analytes. Furthermore, in most cases, the conditions of the factors that enhance the response of one analyte are in conflict with those suitable for some others.

In this work, AQbD (Analytical Quality by Design) is applied to the development of an analytical procedure based on automatic SPE coupled to HPLC-FLD in the determination of nine polycyclic aromatic hydrocarbons (PAHs) in coffee samples.

Focussing on the SPE, the elution volume, the dry time, and volume in the wash stage, and the organic solvent (at two, three, three, and four levels, respectively) were considered.

The first problem is to handle these four factors (control method parameters, CMPs) at different levels to optimize responses (critical quality attributes, CQAs). This task has been carried out using a D-optimal design that, starting from a full factorial design of four factors with 72 experiments, reduced this number to 19, maintaining the precision of the estimates, saving time and costs in the laboratory.

The second problem is related to the choice of CQAs to apply the AQbD methodology. A complex matrix such as coffee contains impurities that interferes with the target analytes and may even coelute in the chromatographic determination. A PARAFAC decomposition allows avoiding this problem and uses the "second order advantage" to unequivocally identify each analyte. Then, the obtained sample loadings were used as responses. Specifically, each CQA is the difference between spiked and blank coffee samples. All these CQAs must be maximized.

Once the experimental data were obtained, two alternatives were posed: on the one hand, the classical optimization based on the estimation of the effects of CMPs on the CQAs, and on the other hand, applying the AQbD methodology to construct the design space that allows to increase the knowledge of the automatic SPE system.

Because the experimental domain of CMPs is discrete and the SPE system performs differently for each analyte, it is not possible to obtain the maximum of all CQAs at the same factor levels. Therefore, the design space of the CMPs is obtaining through the Pareto front of the non-dominated values of CQAs.

The nine PAHs selected were phenanthrene (PHE), anthracene (ANT), fluoranthene (FLN), pyrene (PYR), chrysene (CHR), benzo[a]anthracene (BaA), perylene (PER), benzo[b]fluoranthene (BbF) and benzo[a]pyrene (BaP). European regulations amending foodstuff set maximum levels for BaP and the sum of the content of four compounds (PAH4): BaP, BaA, BbF and CHR.

1. Introduction

Polycyclic aromatic hydrocarbons (PAHs), compounds that have two

* Corresponding author.
E-mail address: mcortiz@ubu.es (M.C. Ortiz).

https://doi.org/10.1016/j.chemolab.2023.105008
Received 21 July 2023; Received in revised form 5 October 2023; Accepted 15 October 2023
Available online 17 October 2023

Abbreviations

EPA	American Environmental Protection Agency
AQbD	analytical quality by design
ANT	anthracene
BaA	benzo[a]anthracene
BaP	benzo[a]pyrene
BbF	benzo[b]fluoranthene
CHR	chrysene
CMP	Control Method Parameters
CORCONDIA index or CORE	core consistency diagnostic
CQA	Critical Quality Attributes
CCα	decision limit
DS	design space
CCβ	detection capability
FLN	fluoranthene
GC-MS	gas chromatography coupled to mass spectrometry
HPLC-FLD	high performance liquid chromatography with fluorescence detector
HLB	hydrophilic-lipophilic-balanced
IARC	International Agency for Research on Cancer
MCR-ALS	multivariate curve resolution by alternating least squares
PARAFAC	parallel factor analysis
PER	perylene
PHE	phenanthrene
PAHs	polycyclic aromatic hydrocarbons
PYR	pyrene
SPE	solid phase extraction
VIF	variance inflation factor

or more condensed aromatic rings, are formed by fragments of organic compounds produced by pyrolysis (at high temperatures) [1]. PAHs can be a source of contamination for foods, on the one hand, they appear ubiquitous in the environment such as in water, soil and air, and on the other hand, they arise in cooking practices such as baking, drying, grilling, smoking, roasting, barbecuing, toasting, heating and frying [2]. Roasting is a key step in coffee production carried out at temperatures between 120 and 230 °C and gives properties of aroma, colour and flavour [3].

Between PAHs, 16 have been considered as high priority pollutants by the American Environmental Protection Agency (EPA) [4] due to their potential toxicity to humans and other organisms and because of their prevalence and persistence in the environment. Among these 16 PAHs, the most harmful is benzo[a]pyrene (BaP) which is included in group 1 "carcinogenic to humans" in the International Agency for Research on Cancer (IARC) list [5]. Looking forward, that classification could change, since the last update of July 2023 is preparing the moving of anthracene (ANT) from group 3 to 2B.

For this reason, PAHs are limited in foodstuff through European Regulation (EU) 2023/915 [6], which repeals Regulation (EC) No 1881/2006. Maximum levels are established for the amount of BaP and for the sum of the content of four compounds named PAH4 (BaP, benzo [a]anthracene (BaA), benzo[b]fluoranthene (BbF), and chrysene (CHR)) for different kinds of foodstuff. As regulation mentions, instant or soluble coffee are excluded from the maximum level for powders of food of plant origin for the preparation of beverages category since negligible amount of PAHs has been found. However, coffee ground is not taken into account in any of the food categories.

In the literature, there are several extraction methods and two main analytical techniques, gas chromatography coupled to mass spectrometry (GC-MS) and high performance liquid chromatography with fluorescence detector (HPLC-FLD), in order to determine PAHs in coffee [1,3,7]. Although lower limits of detection and quantification could be achieved by means of GC-MS regarding HPLC-FLD, the first one cannot be afforded by every laboratory. For this reason, in the present paper, HPLC-FLD was selected since a quick method for the determination of PAHs has previously been optimized [8]. In terms of extraction, the solid phase extraction (SPE) method was selected to extract and pre-concentrate in a single stage [9,10]. It is common to carry out an alkaline saponification with KOH or NaOH in coffee samples in order to remove the lipid fraction, pigments, and other organic contaminants that can interfere during the analysis. However, the benefit of the stage is questioned due to the possible degradation of some PAHs and the interference in the fluorescent signal [11,12].

Different sorbents such as C18 [10,13], silica [14,15] and author-synthetized [16,17] have been used in the SPE of PAHs from coffee samples. However, for this work, the Oasis® HLB cartridge was selected, which contains a water-wettable sorbent. HLB refers to Hydrophilic-Lipophilic-Balanced, which means that the sorbent is made from a specific ratio of two monomers, hydrophilic *N*-vinyl-2--pyrrolidone and the lipophilic divinylbenzene, which allows working with acidic, neutral, and basic compounds [9,18].

Starting from Refs. [9,18], it can be seen that there are a lot of steps and several parameters for each of them that can be optimized for the SPE process. In the reviewed literature, parameters such as pH and volume of sample in load step; volume and type of solvent in the elution step have been optimized [10,13,16,17]. They usually did it through a "one-variable-at-a-time" approach, that means parameters are analysed by changing one factor at a time keeping the other ones constant, without addressing interactions among parameters which could expected to be significant.

In this work, the design space (DS), defined as in Analytical Quality by Design (AQbD) [19–21], has been employed in the analysis of the performance of the automatic SPE system in a multi-residue analysis. To obtain the CMPs (Control Method Parameters) that maximize the desired Critical Quality Attributes (CQAs) a D-optimal design and parallel factor analysis (PARAFAC) have been used. Once the data were collected according to the chosen design, the results (loadings of PARAFAC decomposition) were analysed using a multiple linear regression analysis to establish the multivariate relationship, f, between CMPs and CQAs. From this point onwards, classical optimization and AQbD differ between them. In the former case, f was used to determine the specific values of CMPs that provided the optimal CQAs. However, in the AQbD approach, the inverse function of f was considered to determine the region in the experimental domain of CMPs where the values of CQAs are maintained within the range that the researcher deems appropriate for the purpose of the analysis. Since the design space is an n-dimensional region, its study allows estimating synergies or antagonisms between CMPs, thereby increasing knowledge of the analytical method. Optimizing a multi-residue analysis when using an automatic SPE system and complex matrices becomes a difficult problem due to the large number of experimental factors that can influence the extraction of the analytes (CQAs). Furthermore, in most cases, the conditions that enhance the response for one analyte are in conflict with those suitable for some other.

The first problem is to handle several factors (CMPs) with different levels to optimize the responses (CQAs). This task has been carried out using a D-optimal design to reduce the experimental efforts. Four parameters to be optimized have been selected, being three of them: volume in elution step; dry time with nitrogen and volume in the wash stage. Furthermore, the type of organic solvent used in the preparation of coffee samples was considered as the fourth factor in the design of the experiments, since the addition of organic solvents, such as methanol or acetonitrile, avoid the adsorption of PAHs into the glassware walls [7,

10].

In order to develop the optimization, two alternatives were posed: on the one hand, the classical optimization based on the estimation of the effects of the levels of the factors on the responses and, on the other hand, exploration of the discrete version of the DS which allows increasing knowledge of the automatic SPE procedure.

The second problem is related to the choice of the type of response (CQAs) to optimize in the experimental design. A complex matrix such as coffee contains impurities that can interfere with the target analytes and may even coelute in the chromatographic determination. A PARAFAC decomposition allows avoiding this problem and uses the "second order advantage" to unequivocally identify each analyte [22,23].

In the present paper, phenanthrene (PHE), ANT, fluoranthene (FLN), pyrene (PYR), CHR, BaA, perylene (PER), BbF, and BaP were selected for their determination and quantification with HPLC-FLD in several coffee samples (100 % natural roast coffee of two intensity levels and a mixture of 50 % natural roast coffee and 50 % torrefied roast coffee).

2. Material and methods

2.1. Chemicals and reagents

Anthracene (ANT ≥98 %, CAS no. 120-12-7), fluoranthene (FLN ≥98 %, CAS no. 206-44-0), perylene (PER ≥99 %, CAS no. 198-55-0) and benzo[b]fluoranthene (BbF 98 %, CAS no. 205-99-2) were acquired in Sigma-Aldrich (Steinheim, Germany). Phenanthrene (PHE 98 %, CAS no. 85-01-8), pyrene (PYR 98 %, CAS no. 129-00-0) and benzo[a]pyrene (BaP 96 %, CAS no. 50-32-8) were purchased by Alfa Aesar (Kandel, Germany). Benzo[a]anthracene (BaA 99 %, CAS no. 56-55-3) was acquired in Acros Organic (Geel, Belgium). Chrysene (CHR ≥95 %, CAS no. 218-01-9), acetonitrile (CAS no. 75-05-8), methanol (CAS no. 67-56-1) and acetone (CAS no. 67-64-1) were supplied by Merck (Darmstadt, Germany), the three solvents were LiChrosolv® for liquid chromatography. Dichloromethane (CAS no. 75-09-2; for HPLC HiPerSolv CHROMANORM) was acquired in VWR Chemicals (Fontenay sous Bois, France). Deionized water was obtained by using the Milli-Q gradient A10 water purification system from Millipore (Bedford, MA, USA).

2.2. Instrumental

In order to extract the nine PAHs from the liquid coffee, an automatic SPE model ASPEC GX-271 from Gilson, Inc. World Headquarters (Middleton, WI, USA) and Oasis® HLB 6 cc (200 mg) extraction cartridges from Waters (Milford, MA, USA) were used. The SPE system consists of a tip groove probe, a Verity® 4060 Syringe Pump equipped with a 10 mL syringe, a four polyethylene bottles rack from which the system aspirates the coffee samples (see details of preparation in Section 2.3.). An additional GX rinse pump system was employed to clean the probe with acetone.

For the evaporation of the dichloromethane solvent of the extract obtained from the SPE, Multivap™ nitrogen evaporation system model 11364, Organomation Associates Inc. (Berlín, MA, USA) was used. In the next step, an Ultrasonic Cleaner USC 1200 THD (VWR International, Leuven, Belgium) and a vortex stirrer LBX Instruments V05 series with speed control (Barcelona, Spain) were employed for homogenizing the samples.

The determination of the nine polycyclic aromatic hydrocarbons, PHE, ANT, FLN, PYR, CHR, BaA, PER, BbF, and BaP, was carried out using an Agilent 1260 Infinity HPLC chromatograph (Santa Clara, CA, USA) consisting of a quaternary pump (G1311C), a sampler (G1329B), a thermostatic column compartment (G1316 A), and a fluorescence detector (G1321B). A Kinetex EVO-C18 column (150 mm × 4.6 mm, 5 μm) coupled to a pre-column (SecurityGuard™ ULTRA cartridges sub-2 μm and core-shell columns with 4.6 mm internal diameters), both from Phenomenex (Torrance, CA, USA), were used for the separation.

The chromatographic separation was carried out following the optimized conditions in Ref. [8], using an isocratic mobile phase of 38 % water, 19 % methanol and 43 % acetonitrile, at a flow rate of 1.5 mL min^{-1}. The temperature of the column compartment was fixed at 42 °C and the injection volume was 10 μL. An excitation wavelength of 274 nm was programmed in the fluorescence detector software to record the emission spectrum between 290 and 550 nm, each 1 nm, during the analysis, for 30 min.

2.3. Standard solutions and samples

Individual standard stock solutions of 100 mg L^{-1} were prepared by dissolving each standard in acetonitrile and stored frozen and protected from light. Intermediate individual solutions of 10, 40, 6, 3.5, 2, 0.8, 4, 3.5, and 0.7 mg L^{-1} for PHE, ANT, FLN, PYR, CHR, BaA, PER, BbF, BaP, respectively, were prepared from the individual stock solutions by dilution with acetonitrile. A first intermediate mixture solution of 333 and 500 μg L^{-1} for PHE and ANT respectively, of 83 μg L^{-1} for FLN, PYR, CHR, PER and BbF, and of 33 μg L^{-1} for BaA and BaP, was prepared from the intermediate individual solutions by dilution with acetonitrile. A second intermediate mixture solution of 50 and 75 μg L^{-1} for PHE and ANT respectively, of 12.5 μg L^{-1} for FLN, PYR, CHR, PER and BbF, and of 5 μg L^{-1} for BaA and BaP, was prepared from the first mixture solution by dilution with acetonitrile. The first intermediate mixture solution was used to prepare eleven calibration standards by dilution with acetonitrile, which were ranged between 0 – 40 and 0 – 60 μg L^{-1} for PHE and ANT respectively, between 0 – 10 μg L^{-1} for FLN, PYR, CHR, PER and BbF, and between 0 – 4 μg L^{-1} for BaA and BaP. All the solutions were stored protected from light at 4 °C.

In order to take into account possible matrix interferences, in the optimization of the automatic SPE system, capsules filled with 100 % natural roast coffee and intensity level 7, named (V), were used. Once the automatic SPE system was optimized, three different types of coffee from three new brands were considered for their analysis in this work: i) capsules filled of 100 % natural roast coffee and intensity level 7, named (M1, M4, M7); ii) capsules filled with 100 % natural roast coffee and intensity level 11, named (M2, M5, M8); and iii) capsules filled at laboratory with 5 g of a mixture of 50 % natural roast coffee and 50 % torrefied roast coffee (coffee roasted in the presence of sugar), named (M3, M6, M9). All the types of coffee were bought in local supermarkets.

The procedure to obtain the liquid coffee consisted in putting two capsules of each type and each brand of coffee in the capsule coffee machine with Milli-Q water, and a volume of 150 mL was collected in individual beakers. Between each type and each brand of coffee, the coffee machine was cleaned with Milli-Q water without the coffee capsule, passing through the system an approximate volume of 130 mL. For each type and each brand of coffee, a blank coffee sample and a spiked coffee sample were prepared in 250-mL volumetric flasks adding: i) 50 mL of liquid coffee; ii) 2.25 mL of acetonitrile for the blank coffee sample or 2.25 mL of the second intermediate mixture solution for spiked coffee sample to have a concentration of 450 and 675 ng L^{-1} for PHE and ANT respectively, 112.5 ng L^{-1} for FLN, PYR, CHR, PER and BbF, and 45 ng L^{-1} for BaA and BaP; iii) 25 mL of acetonitrile and 25 mL of methanol (experimental factor optimized in Sections 4.2. and 4.3.); iv) completed to the mark with Milli-Q water.

Blank and spiked coffee samples were prepared and transferred to polyethylene bottles for SPE system. The analysis of both types of coffee samples must be carried out in order to calculate the quantity of analytes added which was extracted after applying the SPE procedure, considering the quantity of analytes already present in the blank coffee samples.

2.4. Solid phase extraction and preparation of extracts

The optimized SPE procedure was based on Refs. [9,18] where the selected cartridge was used in the extraction of PAHs. The procedure consisted in six steps: two conditions, load, wash, and two elutions. First,

the Oasis® HLB cartridges were conditioned in two steps: 5 mL of methanol and 5 mL of Milli-Q water, both with a result flow rate of 4 mL min^{-1}.

The full load of the coffee sample from polyethylene bottles was 40 mL, volume which was taken in five individual loads of 8 mL with a result flow rate of 1 mL min^{-1}. At that point, the probe was cleaned with acetone during 0.5 min.

In wash stage, 1 mL (experimental factor optimized in Sections 4.2. and 4.3.) of 5 % methanol in Milli-Q water was applied with a result flow rate of 4 mL min^{-1}. Then, cartridges were dried under a nitrogen flow of 1 bar during 7.5 min (experimental factor optimized in Sections 4.2. and 4.3.). Again, the probe was cleaned with acetone during 0.5 min.

Finally, the elution step was carried out in two tasks: 5 mL (experimental factor optimized in Sections 4.2. and 4.3.) and 1 mL of dichloromethane were added with a result flow rate of 2 mL min^{-1}. The volume of both tasks was collected inside a glass test tube in which 0.5 mL of Milli-Q water was previously added. The cartridges were dried under a nitrogen flow of 1 bar during 1 min. The source flow rates of all the steps of the automatic SPE system were fixed at 10 mL min^{-1}.

The dichloromethane of the final extracts was evaporated with a nitrogen evaporation system which was installed at 2 bar of pressure. The test tubes were moved away when the final volume reached 0.5 mL, and then, 1 mL of acetonitrile was added in each test tube. For homogenizing the samples, the ultrasonic cleaner at a power level of 9 during 3 min and a vortex stirrer during 25 s were employed. These volumes of 1.5 mL were transferred to amber HPLC vials using 0.22 μm pore size polypropylene filters.

2.5. Software

The automatic SPE system was controlled by Trilution LH software. OpenLab CDS ChemStation software was used for acquiring HPLC data. PLS_Toolbox [24] for use with MATLAB [25] was employed in PARAFAC decomposition. The D-optimal experimental design was selected with NEMRODW [26]. The program COO-FRO [27] was used to obtain the parallel coordinates plot. The calibration and accuracy lines were fitted and validated applying STATGRAPHICS Centurion 19 [28]. Decision limit (CCα) and detection capability (CCβ) were calculated using the DETARCHI program [29].

3. Theory

3.1. Design of experiments: D-optimal design

For the design of experiments, four parameters were chosen as factors: volume (X_1) in the first elution step; dry time with nitrogen (X_2) and volume (X_3) in the wash stage; and type of organic solvent (X_4) used in the preparation of the coffee samples. The possible interaction between X_2 and X_3 has also been considered.

A saturated design, such as a Plackett-Burman design, usually used as a screening design, does not allow handling more than two levels per factor. In addition, these designs do not allow seeing the possible interactions between factors that are always confused with the main factors [30,31]. For this reason, in this work a full factorial design as initial experimental domain was used. The full factorial design supposed a total of $N = 2 \times 3 \times 3 \times 4$ experiments.

The levels of each factor and their code are shown in Table 1a). The first stage of the design of experiments was to fit a model which relates, from the experimental data, the change in each response considered Y_i, $i = 1, \ldots,9$ (a response for each PAH) when varying the levels of the factors. For a response Y, the presence-absence model that includes all the effects and interactions is written by means of coefficients of indicator variables, which is shown in Eq. (1).

$$\begin{aligned} Y = {} & \beta_0 + \beta_{A1}X_{A1} + \beta_{A2}X_{A2} + \\ & + \beta_{B1}X_{B1} + \beta_{B2}X_{B2} + \beta_{B3}X_{B3} + \\ & + \beta_{C1}X_{C1} + \beta_{C2}X_{C2} + \beta_{C3}X_{C3} + \\ & + \beta_{D1}X_{D1} + \beta_{D2}X_{D2} + \beta_{D3}X_{D3} + \beta_{D4}X_{D4} + \\ & + \beta_{B1C1}X_{B1}X_{C1} + \beta_{B1C2}X_{B1}X_{C2} + \beta_{B1C3}X_{B1}X_{C3} + \\ & + \beta_{B2C1}X_{B2}X_{C1} + \beta_{B2C2}X_{B2}X_{C2} + \beta_{B2C3}X_{B2}X_{C3} + \\ & + \beta_{B3C1}X_{B3}X_{C1} + \beta_{B3C2}X_{B3}X_{C2} + \beta_{B3C3}X_{B3}X_{C3} \end{aligned} \quad (1)$$

In Eq. (1), the binary variables X_{ij}, $i = A, B, C, D$, $j = 1, \ldots, n_i$ (being n_i the number of the levels of the factor i-th) takes the value of 1 if the factor i is at level j and the value of 0 in the other cases. In addition to the intercept β_0, this model has 21 coefficients β_{ij}, that estimate the effect on the response when the i-th factor is fixed at level j. In the case of the interactions, the coefficient β_{ijkl} corresponds to the effect on the response when the factor i is at level j and the factor k is at level l simultaneously. This model will be used in the analysis of the effect of the levels to establish the combination of the levels which lead to the optimal solution. However, the estimation of the coefficients carried out from the experimental data is unfeasible since for each factor i the sum of the binary variables is constant, that is $X_{i1} + X_{i2} + \ldots + X_{in_i} = 1, i = A, \ldots, D$. Therefore, the reference-state model, in Eq. (2), is used instead for the estimation of effects.

Table 1
a) The four factors and their levels in the SPE optimization procedure. b) Levels of the factors that maximize each response $Y_{analyte}$.

a)	Code	Factor	Number of levels	Code Level
	X_1	Elution volume	2	A1: 2 mL A2: 5 mL
	X_2	Dry time with nitrogen in the wash stage	3	B1: 5 min B2: 7.5 min B3: 10 min
	X_3	Wash volume	3	C1: 1 mL C2: 3 mL C3: 5 mL
	X_4	Organic solvent	4	D1: 0 % D2: 20 % MeOH D3: 20 % ACN D4: 10 % MeOH +10 % ACN

b)	Y_{PHE}	Y_{ANT}	Y_{FLN}	Y_{PYR}	Y_{CHR}	Y_{BaA}	Y_{PER}	Y_{BbF}	Y_{BaP}	Level chosen
X_1	A2	A2	A2	A1	A2	–	A2	A2	A2	5 mL
X_2	B3	B1	B2	B2	B2	B2	B2	B2*	B2	7.5 min
X_3	C2	C1	C1	C1	C1	C1	C1	C1	C1	1 mL
X_4	D1*	D4	D4	D4*	D4*	D4*	D4	D4	D4*	10 % MeOH +10 % ACN

*significant at 0.05.

$$\begin{aligned} Y = {} & \beta'_0 + \beta'_{A1}X'_{A1} + \\ & + \beta'_{B1}X'_{B1} + \beta'_{B2}X'_{B2} + \\ & + \beta'_{C1}X'_{C1} + \beta'_{C2}X'_{C2} + \\ & + \beta'_{D1}X'_{D1} + \beta'_{D2}X'_{D2} + \beta'_{D3}X'_{D3} + \\ & + \beta'_{B1C1}X'_{B1}X'_{C1} + \beta'_{B1C2}X'_{B1}X'_{C2} + \\ & + \beta'_{B2C1}X'_{B2}X'_{C1} + \beta'_{B2C2}X'_{B2}X'_{C2} \end{aligned} \quad (2)$$

In addition to the intercept β'_0, the reference-state model has only 12 coefficients β'_{ij} which correspond to the effect of changing the level of a factor regarding a reference level. For Eq. (2) the reference level considered was the last one for each factor. In the model in Eq. (2), the levels of each factor are codified jointly by the binary variables which intervene in this factor. For example, when the factor X_2 is at level 1, $X'_{B1} = 1$ and $X'_{B2} = 0$; if it is at level 2, $X'_{B1} = 0$ and $X'_{B2} = 1$; and if it is at level 3 (the reference one) $X'_{B1} = -1$ and $X'_{B2} = -1$.

Once calculated the coefficients of the reference-state model (β'_{ij}), the ones of the presence-absence model (β_{ij}) are obtained because both models are mathematically related between them (as can be seen in the Supplementary Material of Ref. [32]).

Taking into account the number of factors and the number of levels of each one (Table 1a)), the full factorial design supposes a total of 72 experiments when considering all the available combinations. However, not as many experiments are needed to estimate the model in Eq. (2), since only 13 experiments are theoretically needed. In order to reduce the number of considered experiments, the D-criterion is applied, which gives a selection of experiments that can be carried out. A technical explanation about D-optimal design can be seen in Ref. [30]. The D-optimal design selected has 19 experiments which are shown in Table 2. In addition to the reduction of a 73.6 % of the work at the laboratory, the quality of the estimates is guaranteed since the VIF (variance inflation factor) values of the coefficients of the model in Eq. (2) vary between 1.04 and 2.09, which are lightly higher than the ones obtained with the 72 experiments of the full factorial design, varying between 1 and 1.78, but wide lower than the required threshold of 4 or 7 [30].

3.2. Pareto front and parallel coordinates plot

The issue of reaching the maximum of the several variables (Y_i, $i = 1$, …,9), which are functions of experimental conditions, is a multi-objective task. In addition, in this case, the domain of the experimental conditions is discrete. Because of that, the mathematical tools of the response surface analysis such as optimal path and canonical analysis are impossible to be applied. Moreover, the conditions that maximize the Y_i related to an analyte can be non-adequate for the Y_i related to another. For these reasons, Pareto front of the non-dominated solutions is used in this paper in order to discard the experimental conditions that do not need to be studied (dominated solutions). Considering two different experimental conditions (w,z), defining each of them by the corresponding level of the factors, the Y_i(z) for condition z is said to dominate the Y_i(w) for condition w (that is, z dominates w) when a) and b) are true:

a) For each PAH, $Y_i(z) \geq Y_i(w)$ is fulfilled
b) At least one of the previous inequalities is strict

The Pareto front is formed by the non-dominated solutions. Obviously, the remaining solutions have no interest. Refs. [8,33] show that Pareto front is a discrete estimate of the DS in AQbD.

In this work, Pareto front are points of a curve in the nine-dimensional space. Each point corresponds to an experimental condition and is made up of the nine loadings (Y_i) (one for each PAH) obtained in that condition. In the practise, the multi-dimensional character of the Pareto front requires its visualization by using a parallel coordinates plot [34], which has been introduced in analytical applications [35] and was of great use in AQbD [8,33]. The vector of the CQAs, nine responses (Y_1, Y_2, …,Y_9), and the corresponding CMPs, experimental conditions (X_1, X_2, X_3, X_4) are represented by means of a parallel coordinates plot. This graph consists in drawing four parallel vertical lines on of which the levels of the factor are marked X_i, $i = 1$, …,4; and next, another nine parallel lines where the corresponding values of Y_i, $i = 1$, …,9 are marked. Finally, the marked points of the parallel vertical lines are joint through broken lines. That broken line shows the performance of the nine responses at the same time as the conditions of the automatic SPE from which have been generated.

3.3. Second order advantage

Several chemometric techniques can be applied to analyse three-way or higher-order multivariate data with the aim of obtaining unequivocal identification and quantification of analytes when measured in complex matrices. The most common ones are multivariate curve resolution by alternating least squares (MCR-ALS) and parallel factor analysis (PARAFAC/PARAFAC2) [36,37]. The solution of the applied decomposition must be unique because it implies achieving the second order advantage which permits analyte quantification in the presence of unknown and uncalibrated interferents [38]. Among both techniques mentioned

Table 2
Codified experimental conditions and responses for the experiments of the D-optimal design.

X_1[a]	X_2[a]	X_3[a]	X_4[a]	Y_{PHE}	Y_{ANT}	Y_{FLN}	Y_{PYR}	Y_{CHR}	Y_{BaA}	Y_{PER}	Y_{BbF}	Y_{BaP}
A1	B1	C1	D1	0.1060	0.0648	0.0759	0.0787	0.0854	0.0790	0.0925	0.0197	0.0731
A2	B2	C1	D1	0.0694	0.0707	0.0651	0.0606	0.0917	0.0766	0.0934	0.0755	0.0828
A2	B3	C1	D1	0.0601	0.1161	0.1138	0.0554	0.0617	0.0652	0.0798	0.0648	0.0691
A1	B1	C3	D1	0.0483	0.0393	0.0479	0.0456	0.0564	0.0601	0.0512	0.0405	0.0521
A2	B3	C3	D1	0.0586	0.0163	0.0341	0.0453	0.0594	0.0504	0.0626	0.0139	0.0513
A1	B1	C2	D2	0.0689	0.0421	0.0813	0.0682	0.0774	0.0876	0.0648	0.0652	0.0795
A1	B2	C2	D2	0.0703	0.0578	0.0801	0.0706	0.0780	0.0840	0.0749	0.0731	0.0752
A2	B3	C2	D2	0.0938	−0.0950	0.0287	0.0450	0.0581	0.0579	0.0813	0.0460	0.0751
A2	B1	C3	D2	0.0400	0.0697	0.0669	0.0551	0.0526	0.0658	0.0681	0.0497	0.0652
A1	B3	C3	D2	0.0309	0.0532	0.0302	0.0437	0.0339	0.0416	0.0212	0.0245	0.0298
A2	B1	C1	D3	0.0067	0.1172	0.0130	0.0048	0.0464	0.0319	−0.0100	0.0230	0.0352
A1	B3	C1	D3	0.0004	0.0859	0.0000	0.0100	0.0259	0.0291	−0.0019	0.0120	0.0246
A1	B1	C2	D3	0.0070	0.0456	−0.0071	0.0011	0.0157	0.0137	−0.0189	0.0103	0.0103
A1	B3	C2	D3	0.0295	0.0648	0.0221	0.0269	0.0285	0.0319	0.0148	0.0148	0.0352
A2	B2	C3	D3	0.0375	0.0339	0.0473	0.0186	0.0390	0.0408	0.0510	0.0547	0.0465
A1	B2	C1	D4	0.1132	0.0792	0.1217	0.1094	0.1184	0.1211	0.1191	0.0883	0.1155
A2	B1	C2	D4	0.0987	0.0783	0.1022	0.0905	0.1087	0.1086	0.1044	0.0954	0.1112
A2	B2	C2	D4	0.0718	0.0380	0.0664	0.0637	0.0812	0.0822	0.0884	0.0723	0.0886
A1	B2	C3	D4	0.0668	0.0551	0.0775	0.0612	0.0783	0.0771	0.0734	0.0569	0.0785

[a] A – D refer to the codes of each level of the factors from the design of experiments (see Table 1a)).

above, some authors prefer PARAFAC due to its second order advantage [23,39]. If the loss of trilinearity is significant, MCR-ALS is a useful alternative, since its signal requirements are weaker than those required by PARAFAC2. MCR-ALS has been widely used in analytical chemistry. Its major limitation is the presence of rotational ambiguity and then, the lack of uniqueness in the solution. However, in some cases, the non-uniqueness problem can be alleviated or completely avoided through intelligent use of the data structure and appropriate constraints in its application [40]. In particular, important progress in evaluating rotational ambiguity can be seen in Refs. [41–43].

Next, in order to an easier reading of the paper, several aspects of the PARAFAC decomposition are detailed.

A three-way data array $\underline{\mathbf{X}}$ of dimension $I \times J \times K$ is formed by x_{ijk} numbers, $i = 1, \ldots, I$; $j = 1, \ldots, J$; $k = 1, \ldots, K$. HPLC-FLD data can be arranged in a three-way array $\underline{\mathbf{X}}$ and analysed with the PARAFAC decomposition technique, where for each of the K samples analysed, the intensity measured at J wavelengths is recorded at I elution times around the retention time of every analyte. The trilinear PARAFAC model of rank F for the data array $\underline{\mathbf{X}} = (x_{ijk})$ is written [44,45] as Eq. (3):

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}, i = 1, 2, \ldots, I;\ j = 1, 2, \ldots, J;\ k = 1, 2, \ldots, K \qquad (3)$$

where e_{ijk} are residuals of the fitted model, F is the number of factors and $\mathbf{a}_f$, $\mathbf{b}_f$ and $\mathbf{c}_f$, ($f = 1, 2, \ldots, F$) are the loading vectors of the chromatographic, spectral and sample profiles, respectively.

Chromatographic data are trilinear if the experimental data array is compatible with the structure in Eq. (3). The core consistency diagnostic (CORCONDIA index or CORE) [46] measures the trilinearity degree of the experimental three-way array when $F \geq 2$. If the three-way array is trilinear, then the maximum CORE value of 100 % is achieved. Additionally, the trilinearity is verified by using partitions in the data set (similarity or split-half analysis), the variance explained and the chemical coherence of the three profiles [46,47]. In the construction of the PARAFAC model, constraints on the profiles can be imposed, for example, non-negativity and/or unimodality.

The PARAFAC solution is unique when the three-way array is trilinear and the appropriate number of factors has been chosen to fit the PARAFAC model [47]. The uniqueness property, also known as "second order advantage" makes it possible to identify compounds unequivocally by their chromatographic and spectral profiles as laid down in some official regulations and guidelines [48,49], even in the presence of a coeluent that appears with the analyte of interest.

4. Results and discussion

4.1. PARAFAC decomposition to obtain the responses for the D-optimal design

As it was mentioned before, in order to take into account possible matrix interferences, in the optimization of the automatic SPE system coffee capsules were used. The reason is that the coffee matrix will differently perform from a standard solution. Since the blank coffee samples could already contain certain quantity of PAHs, it is necessary taking into account the difference between spiked and blank coffee samples in order to carry out the optimization procedure. That difference was obtained through the PARAFAC decomposition sample loadings and was considered as response ($Y_{analyte}$) for the D-optimal design (see columns 5 to 13 in Table 2).

Table 3a) shows characteristics of PARAFAC models for design of experiments, from which the difference of sample loadings was obtained. As can be seen in columns 1 and 2, the retention time window and the spectral range were specifically selected for each analyte, because of that the size of I and J in the third column varies between PAHs. However, the K dimension is always 66, which corresponds to the 19 experiments of the D-optimal design (n = 38 since there were blank and spiked coffee samples); 12 control standard solutions analysed by duplicate (n = 24) to guaranteed sample profile of PARAFAC models; and the remaining 4 were control solutions prepared in coffee matrix.

The constraints imposed for the best fitting of the unique PARAFAC decomposition model were different for each analyte, as can be seen in column 4 in Table 3a). All the PARAFAC models were fitted with two factors, which correspond to the analyte and the baseline, except for the PARAFAC model for BbF whose third factor was associated with another interferent. The two characteristics to guarantee that PARAFAC models were coherent are shown in columns 6 to 7 in Table 3a), as can be seen the CORCONDIA index varied between 91 and 100 %, the variance between 97.99 and 99.53 %. Moreover, the mean similarity calculated by the split-half analysis was 84.22 %.

Table 3
Characteristics of the PARAFAC decomposition models obtained for the determination of PAHs for: a) design of experiments and b) coffee samples.

Analyte	Retention time (min)	Wavelength (nm)	I × J × K	Constraints[a]			Factors	CORE (%)	Variance (%)	Spectral correlation coefficient
				P1	P2	P3				
a) Design of experiments										
PHE	4.75–5.08	325–550	36 × 226 × 66	U	U	N	2	100	97.99	0.9966
ANT	5.13–5.50	355–410	40 × 56 × 66	U	U	N	2	94	99.53	0.9945
FLN	6.36–6.74	380–520	41 × 141 × 66	U	U	N	2	99	99.38	0.9990
PYR	6.80–7.28	340–550	52 × 211 × 66	U	U	N	2	100	98.00	0.9994
CHR	9.50–10.03	340–500	58 × 161 × 66	U	U	N	2	100	99.10	0.9987
BaA	10.08–10.70	360–530	67 × 171 × 66	U	U	N	2	100	99.10	0.9990
PER	13.20–13.90	420–530	75 × 111 × 66	U	U	N	2	100	99.53	0.9957
BbF	13.95–14.65	370–520	75 × 151 × 66	U	U	U	3	91	99.17	0.9972
BaP	14.80–15.70	380–550	96 × 171 × 66	U	U	N	2	100	99.10	0.9973
b) Coffee samples										
PHE	4.75–5.08	325–550	36 × 226 × 94	U	U	N	2	100	98.53	0.9974
ANT	5.17–5.60	370–440	46 × 71 × 94	U	U	N	2	99	99.46	0.9963
FLN	6.30–6.76	380–520	50 × 141 × 94	U	U	N	2	98	99.48	0.9962
PYR	6.80–7.40	340–550	65 × 211 × 94	U	U	N	3	96	96.98	0.9968
CHR	9.50–10.30	360–480	86 × 121 × 94	UN	U	N	3	94	99.13	0.9997
BaA	10.12–10.70	360–530	62 × 171 × 94	U	U	N	2	100	99.15	0.9966
PER	13.25–14.00	420–530	81 × 111 × 94	U	U	N	2	100	99.43	0.9914
BbF	14.05–14.75	350–520	75 × 171 × 94	U	U	N	3	91	99.21	0.9862
BaP	14.90–15.80	380–550	96 × 171 × 94	U	U	N	2	100	99.07	0.9971

[a] Constraints used for chromatographic (P1), spectral (P2) and sample (P3) profiles are codified as: (N) non-negativity, (U) unconstrained, (UN) unimodality and non-negativity.

The unequivocal identification of each PAH was done by comparing the chromatographic and spectral profiles, obtained with the PARAFAC decomposition, with those of a reference sample analysed in the laboratory. In the case of the chromatographic profile, the strict criteria of many European Regulations on veterinary residues and/or pesticides [48,49] has been followed. Therefore, the retention time obtained with PARAFAC decomposition, must correspond to the retention time of a reference sample, admitting a tolerance of ±0.1 min. In this work, all the chromatographic profiles fulfil the aforementioned premise. In the case of the spectral profile, the unequivocal identification has been carried out through the correlation coefficient between the spectral profile obtained from PARAFAC decomposition and the one from a reference standard solution. The values of the correlation coefficient obtained for each PAH are shown in column 8 in Table 3a), being all of them close to 1, what guarantees the identity of the PAH.

4.2. Analysis of the effects and selection of optimal condition

The first step is to fit the reference-state model in Eq. (2) to the experimental data in Table 2. During the procedure of least squares regression, a Box-Cox transformation [50] was necessary for the responses Y_{PHE}, Y_{ANT} and Y_{FLN}. The transformed variables were $(Y_{PHE})^{0.5}$, $(Y_{ANT})^{2}$ and $(Y_{FLN})^{0.8}$. Each transformation was an increasing monotone function, so the maximum is found in the same experimental condition for the original response and the transformed one.

Next, the coefficients of the presence-absence model in Eq. (1) were obtained from the ones of the reference-state model in Eq. (2). The graphical analysis of the effect of the levels of the factors on the different studied responses is shown in Fig. 1. For each response, the coefficients of the model in Eq. (1) are shown by means of bars, which can be positive or negative. The positive coefficients make the response higher, while the negative coefficients reduce it. Being for this work all the responses the difference between loadings of spiked and blank coffee samples, these have to be maximized because is closely related to the extraction of the automatic SPE procedure. Each coefficient is identified by the corresponding code in Table 1a) and Eq. (1): A1, A2, B1, B2, B3, C1, C2, C3, D1, D2, D3 and D4 for the factors, and B1–C1, B1–C2, B1–C3, B2–C1, …, B3–C3 for the interactions. In Fig. 1, the dotted vertical lines mark the limits of the critical region of the significance test considering a level of 0.05. The bars that exceed those limits correspond to significant standardized coefficients (standardization was carried out so that the critical region of the test for all coefficients had the same length), that means, that level of factor has to be chosen as the best condition for that response considered. However, also the selected levels for the individual factors have to be revised with the interaction. Although none interaction is significant, it is necessary to consider them in order to choose the optimal condition of the automatic SPE.

As an example, the analysis for PHE is described: for X_1, level A2 was selected as it shows a positive coefficient. In the case of X_2 and X_3 as individual factors, level B2 and C2 would have to be selected respectively. However, taking into account the interaction, level B3 and C2 have to be chosen for X_2 and X_3, respectively. This is because the sum of the three coefficients (the two of the individual factors and the one of the interaction) is higher when level B3 and C2 are chosen for X_2 and X_3, respectively. Finally, for X_4 level D1 was selected since is significant. Then, for PHE, levels A2, B3, C2 and D1 were selected for X_1, X_2, X_3 and X_4, respectively. The same procedure was followed to analyse the coefficients obtained for the eight remaining responses. Table 1b) collects the levels selected for each individual response.

The last column in Table 1b) shows the optimal condition finally chosen, following as the criterion of selection the level that appears in a higher number of responses. In this way, volume in the first elution step was 5 mL; dry time with nitrogen and volume in the wash stage were 7.5 min and 1 mL, respectively; and type of organic solvent used in the preparation of the coffee samples was 10 % methanol and 10 % acetonitrile, which means adding 25 mL of methanol and 25 mL acetonitrile.

4.3. Obtaining the design space by means of Pareto front and parallel coordinates plot

From the presence-absence model, the values for the nine responses (CQAs) were obtained for all conditions (CMPs) of the full factorial design. In this way, the matrix of the estimated responses $\mathbf{Y}_{72\times9}$ collects the estimated performance of the automatic SPE for all the analytes in all the levels of the factors considered in Table 1a). Fig. 2 is the joint graphical representation of the matrix $\mathbf{Y}$ and the experimental conditions. In the graph, each broken line represents a single vector, whose coordinates jointly contain the coordinates of the method variables (CMPs) and the quality characteristics (CQAs). In order to make easier the visualization of all the possible solutions, all the variables (four CMPs and nine CQAs) have been set on a common scale between 0 and 1, although in Fig. 2 the minimum and maximum values of each CQA are shown. The Pareto front, sets of CMPs and CQAs that are non-dominated, is depicted in green, red and black lines. The remaining solutions (depicted in blue lines) are worse, at least for one of the PAHs.

As can be seen in Fig. 2, there are sets of CMPs that lead to maximum

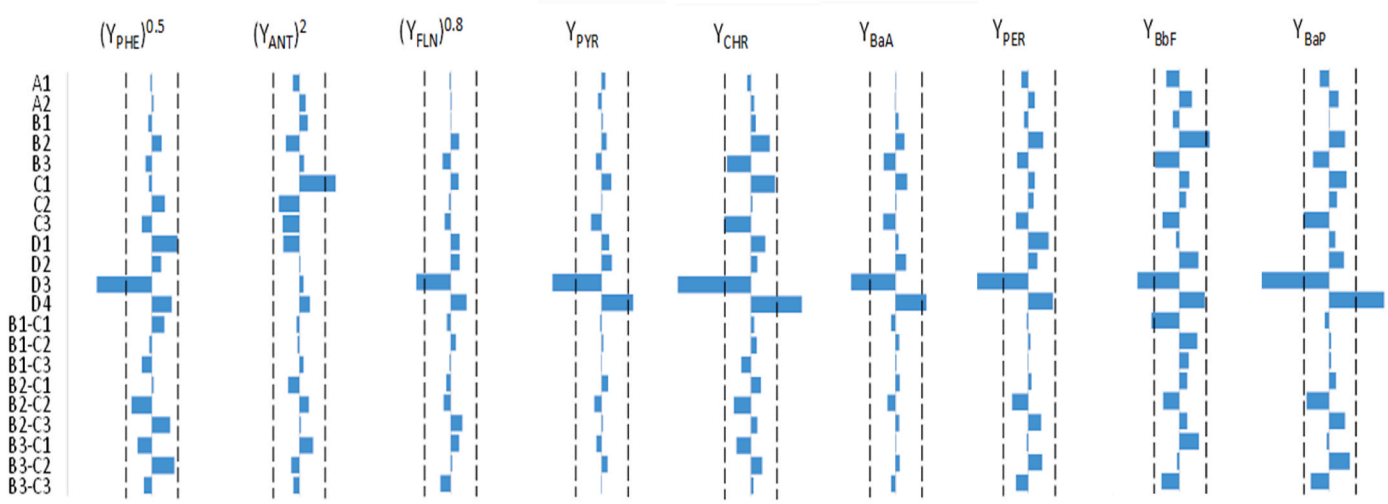


Fig. 1. Graphical analysis of the effects of different levels of the studied experimental factors on the responses (differences between sample loadings of spiked and blank coffee samples for the nine PAHs). On the ordinate axis, the labels correspond to the codes assigned to each level and each factor shown in Table 1a). The dotted vertical lines mark the limits of the critical region for selecting which standardized coefficients are significant (5 % significant level).

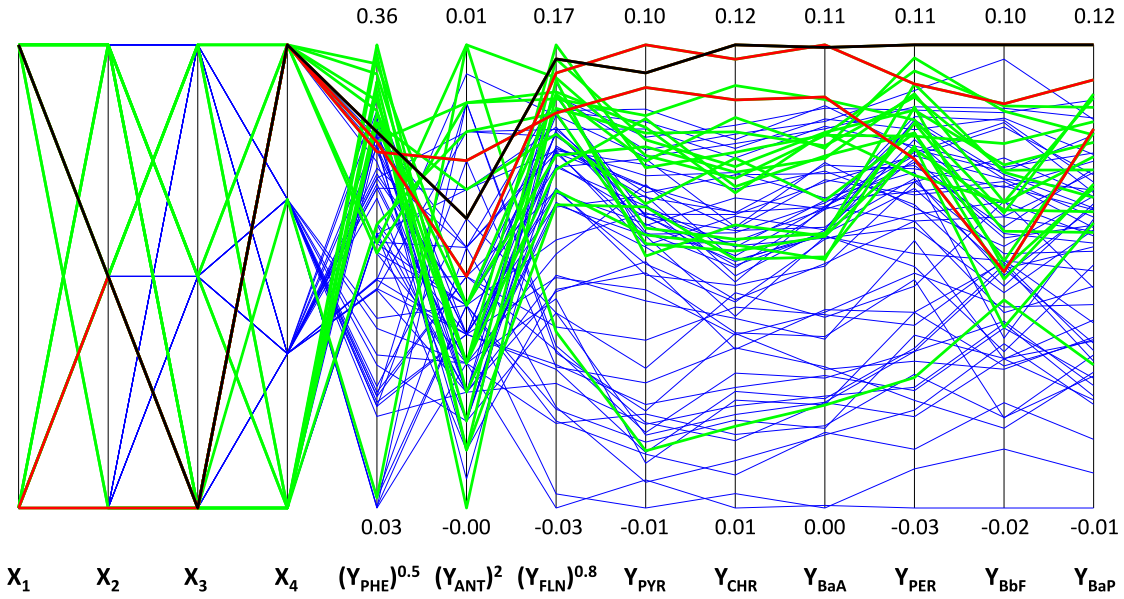


Fig. 2. Parallel coordinates plot for the experimental domain: elution volume, X_1 (two levels), dry time with nitrogen in the wash stage, X_2 (three levels), wash volume, X_3 (three levels), and organic solvent, X_4 (four levels), according to Table 1a), and for the responses $Y_{analyte}$. In green, the non-dominated responses (Pareto front) found. In red two examples of good conditions but not the best one. In black the optimal solution chosen (Table 1b)). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

values of $(Y_{PHE})^{0.5}$ with high values of $(Y_{FLN})^{0.8}$ and low values of $(Y_{ANT})^2$. This means that there is not a set of CMPs for the automatic SPE in which the maximum extraction of PHE, ANT and FLN is achieved simultaneously. In addition, in case of chosen the condition where the extraction of PHE and FLN is the maximum, the loss of ANT can happen.

Another detail to highlight is that all the Pareto front solutions lead to D1 (0 % of organic solvents) and D4 (10 % methanol and 10 % acetonitrile) levels for the factor X_4. Neither there are non-dominated solutions with the combination of levels B3 (10 min) and C3 (5 mL) for the wash stage. Obviously, with the red and black solutions from the Pareto front, an acceptable compromise in relation to the maximum of all the responses is reached. Specially, the one depicted in black agrees with the optimal conditions of the previous section. These three solutions are the discrete estimate of the DS following the AQbD criteria and correspond to the conditions (A1,B1,C1,D4), (A1,B2,C1,D4) and (A2,B2, C1,D4). Since the more important constraint occurs to ANT, depending on the difference which the analyst is willing to take on in relation to the maximum of $(Y_{ANT})^2$, more solutions from the experimental domain can be included in the DS.

It is clear that the analysis of the experimental data by means of Pareto front and parallel coordinates plot increase the knowledge of the automatic SPE in relation to the nine PAHs considered. While the determination of the optimum through the analysis of the effects leads to a unique solution of the experimental conditions.

4.4. Coffee samples results

Once selected the optimal conditions of the automatic SPE system, validation using the same brand of coffee as in the experimental design, and analysis of another nine brands of coffee were carried out. As it was mentioned in Section 4.1., blank and spiked coffee samples were analysed.

4.4.1. PARAFAC decomposition

Table 3b) shows characteristics of PARAFAC models for coffee samples, from which the difference of sample loadings was obtained. As can be seen in columns 1 and 2, the retention time window and the spectral range were specifically selected for each analyte, because of that the size of *I* and *J* in the third column varies between PAHs. However, the *K* dimension is always 94, which corresponds to the 11 calibration standard solutions analysed by duplicate (n = 22); the validation of the SPE procedure by quintuplicate (n = 10); 3 replicates of each new brand of coffee, except for M1 in which one of the replicates the SPE system failed and since sufficient volume of samples remained, two more replicates were carried out (n = 56); and the remaining 6 were control solutions prepared in coffee matrix.

The constraints imposed for the best fitting of the unique PARAFAC decomposition model were different for each analyte, as can be seen in column 4 in Table 3b). All the PARAFAC models were fitted with two factors, which correspond to the analyte and the baseline, except for the PARAFAC model for PYR and BbF whose third factor was associated with another interferent, and except for CHR in whose selected retention time window part of the chromatographic peak of BaA was included (appearing as the third factor). The two characteristics to guarantee that PARAFAC models were coherent are shown in columns 6 and 7 in Table 3b), as can be seen the CORCONDIA index varied between 91 and 100 % and the variance between 96.98 and 99.48 %.

The unequivocal identification of each PAH through the chromatographic and spectral profiles were done following the same procedure as in Section 4.1., guaranteeing the identity of each PAH.

Figs. 3–5 show chromatographic and spectral profiles of the PARAFAC models for coffee samples obtained for each PAH, being the orange factor the analyte, the blue one the baseline, and the yellow one another compound. Fig. 6 shows two examples of sample profile of the PARAFAC models for coffee samples: a three-factor model for PYR and a two-factor model for BaA. As can be seen in Fig. 6, blue factor which corresponds to the baseline is nearly constant. The orange factor, which is related to the analyte, follows an increasing trend for the calibration standard solutions, and for the coffee samples, the blanks appeared below the spiked ones. In the case of PYR sample profile, is remarkable

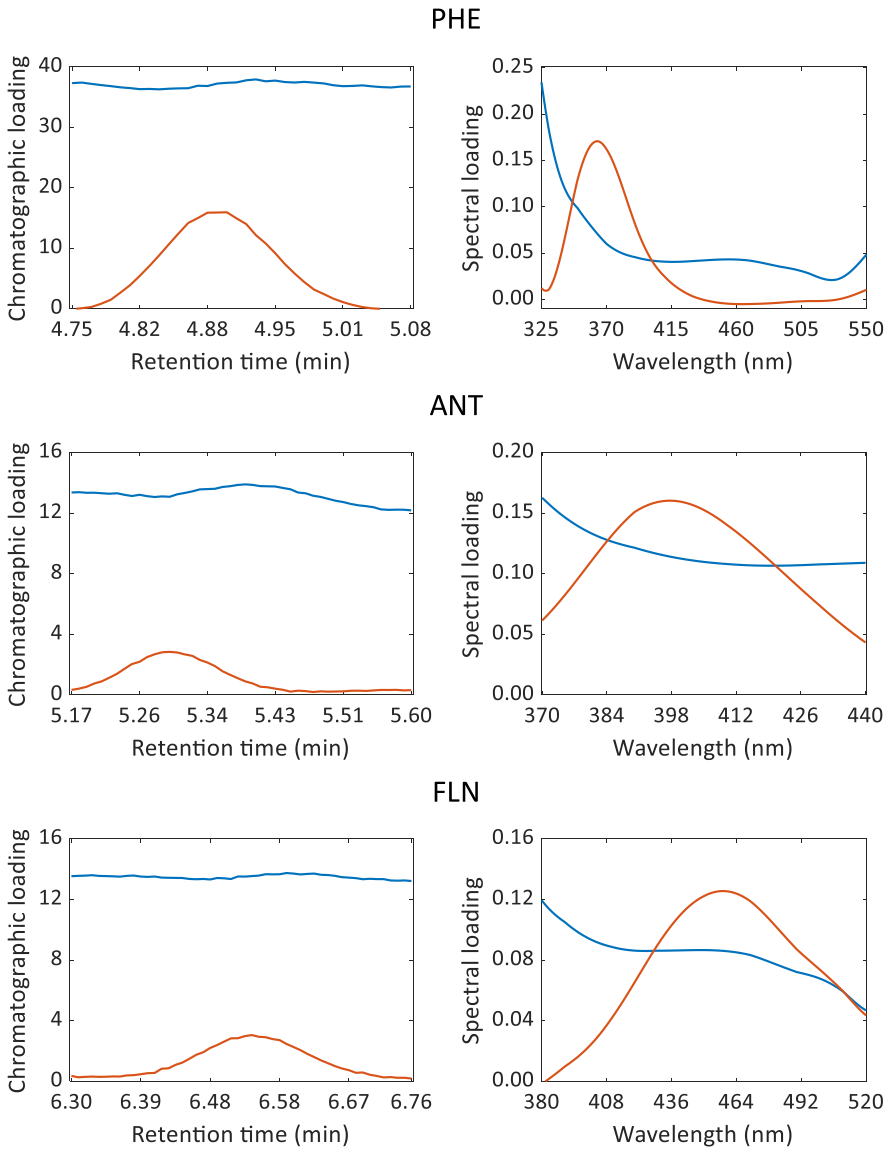


Fig. 3. Loadings of the PARAFAC models for coffee samples obtained for chromatographic (left) and spectral (right) for PHE, ANT and FLN, being the orange factor the analyte and the blue one the baseline. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

that the yellow factor does not appear in any of the new brands of coffee.

4.4.2. *Figures of merit of the analytical method*

The analytical procedure was validated in terms of linear range, accuracy (trueness and precision), decision limit (CCα) and detection capability (CCβ) for the nine PAHs under study.

Table 4 contains the details of calibration and accuracy lines. Calibration lines are fitted (one for each PAH) with the eleven standard solutions, whose range is shown in row 1 in Table 4 (all concentration levels are analysed by duplicate). Rows 2 to 7 in Table 4 show the parameters of the calibration lines, which are all statistically significant because the P-values in row 7 are less than 10^{-4} (H_0: the regression model is not significant).

Trueness and precision were checked using accuracy lines (predicted concentration *versus* true concentration). Their details are shown in rows 8 to 12 of Table 4, including the P-values of the joint hypothesis test (H_0: intercept equal to zero and slope equal to one) in row 12. As can be seen, there is no evidence to reject H_0 since the P-values are close to 1, therefore, the method is unbiased. The precision of the method can be estimated by the deviation standard, s_{yx} in row 11.

Once validated calibration lines, they were used in order to compute CCα and CCβ, which were determined with probabilities of false positive (α) and false negative (β) set at 0.05, following Refs. [51,52]. The values of CCα and CCβ are shown in rows 1 and 2 in Table 5, where it is seen that the analytical method enabled the quantification of 2.2, 3.8, 1.1, 0.3, 0.7, 0.2, 1.1, 1.1, and 0.2 μg L^{-1} for PHE, ANT, FLN, PYR, CHR, BaA, PER, BbF, and BaP, respectively.

4.4.3. *Validation and analysis of several coffee samples*

From the sample loadings obtained from PARAFAC models and the validated calibration lines, predicted concentration of each PAH were calculated for blank and spiked coffee samples. Then, the mean of the replicates was done, obtaining the values show in rows 1 and 2 of each type of coffee samples (V and M1 – M9) in Table 5. The code (V) refers to capsules filled with 100 % natural roast coffee and intensity level 7. (M1 – M3; M4 – M6 and M7 – M9) refer to three new brands of coffee. (M1, M4 and M7) are capsules filled of 100 % natural roast coffee and intensity level 7 of three new brands; (M2, M5 and M8) are capsules filled with 100 % natural roast coffee and intensity level 11; and (M3, M6 and M9) are capsules filled at laboratory with 5 g of a mixture of 50 %

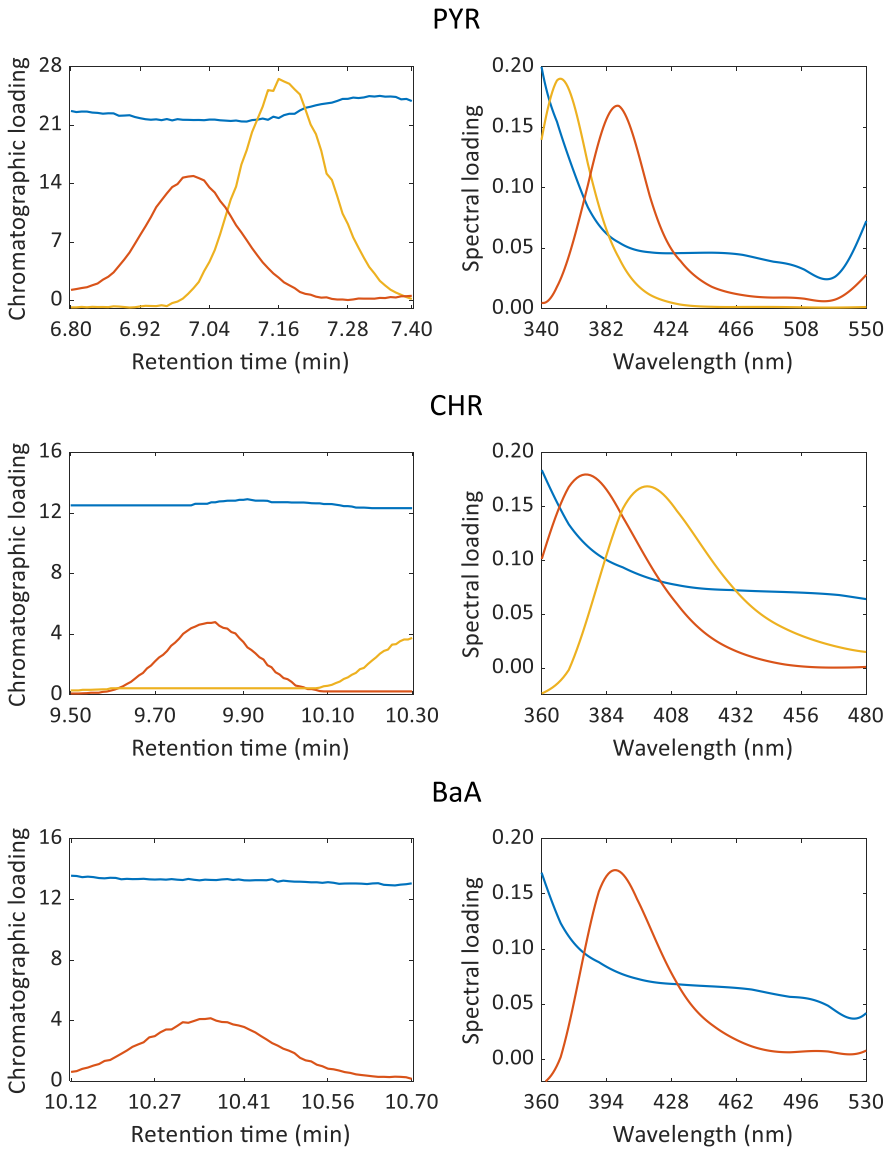


Fig. 4. Loadings of the PARAFAC models for coffee samples obtained for chromatographic (left) and spectral (right) for PYR, CHR and BaA, being the orange factor the analyte, the blue one the baseline, and the yellow one another compound. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

natural roast coffee and 50 % torrefied roast coffee.

The predicted concentration of each PAH for blank and spiked coffee samples was compared to CCα and CCβ values. If the predicted concentration was below CCα, "not detected" was assigned. Although, if it was between CCα and CCβ, the predicted concentration was highlighted with an asterisk (see Table 5). The difference between the predicted concentration of spiked and blank coffee samples was calculated in all cases, except when "not detected" appears in the blank ones, in which case only the spiked predicted concentration was taken into account.

As it was mentioned in Section 2.3., coffee samples were spiked with a standard solution which contained the nine PAHs. On the value of the concentration added in the 250-mL volumetric flask, the factor of preconcentration of the whole SPE process was applied. If the recovery of the SPE process was 100 %, the concentration value of each PAH in the vial to be injected would be the one shown in row 3 in Table 5.

The methodology applied not only for the validation coffee (named V) but also for the nine new different coffees (M1 - M9) consisted in: 1) calculate the concentration of each PAH present in the blank and spiked coffee sample (rows 1 and 2 of each type of coffee samples); 2) with the difference between both calculated concentrations and knowing the concentration added in the spiked coffee samples, the recovery of each PAH and each type of coffee is calculated (row 3 of each type of coffee samples in Table 5); 3) the obtained recovery is applied over the predicted concentration in the blank coffee sample in order to obtain the concentration of each PAH in liquid coffee and coffee ground (rows 4 and 5 of each type of coffee samples in Table 5). Obviously, the SPE preconcentration and the preparation of coffee samples conversion factors were applied. The amount of PAHs found in coffee samples cannot be compared to a maximum level, since regulation does not establish one for coffee ground.

When validation coffee sample (V) was analysed, recoveries nearly or above 50 % were obtained for all PAHs except for ANT and PER. Commission Regulation No 836/2011 in force establishes [53] the performance criteria for methods of analysis for polycyclic aromatic hydrocarbons, considering recoveries between 50 and 120 % for the four PAHs (BaP, BaA, BbF and CHR) in the food categories included in Commission Regulation 2023/915. As can be seen in Table 5, the proposed methodology fulfils the established values for these four PAHs,

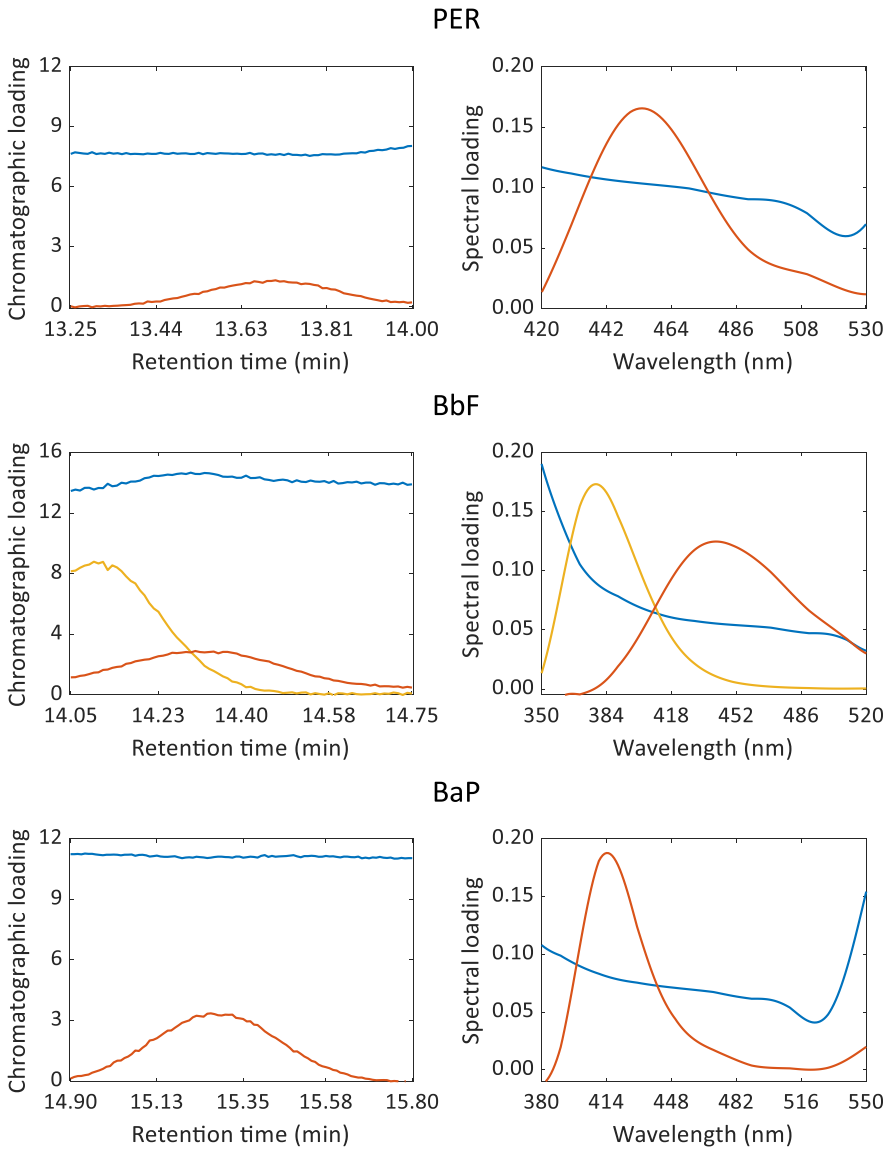


Fig. 5. Loadings of the PARAFAC models for coffee samples obtained for chromatographic (left) and spectral (right) for PER, BbF and BaP, being the orange factor the analyte, the blue one the baseline, and the yellow one another compound. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

being 57.8, 56.2, 50.6, and 54.8 % for BaP, BaA, BbF and CHR, respectively. In this brand of coffee, CHR was not detected and the highest concentration found was for PER (2.1 μg L^{-1} and 30.4 μg kg^{-1}).

This optimized automatic SPE system is applied to nine new different coffees. When an optimized method is applied to other different matrices of coffee, "a priori" it is unknown how recoveries of the nine PAHs can be affected by the components which form those matrices. At first sight, none patterns were found between same type of coffee in different brands, or between same brand in different types of coffee (M1 – M9). For that reason, details for each coffee sample are explained.

For M1, recoveries between 30.2 and 59.1 % were obtained. Except for PHE, concentration found was very low, even four analytes were not detected. In the case of PHE, which was the analyte found in the highest concentration for all samples M1 – M9, values between 0.8 and 3.4 μg L^{-1} in liquid coffee and between 11.7 and 50.5 μg kg^{-1} in coffee ground were obtained. For M2, the highest recoveries were found, between 51.8 and 78.1 %. The concentrations found were very low, even three analytes were not detected. For M3, recoveries between 34.3 and 62.1 % were obtained, except for BbF which was not detected in the spiked coffee sample. The concentrations found were very low, even four analytes were not detected. For M4, recoveries nearly or above 50 % were obtained for all PAHs except for BbF. The concentrations found were very low, even five analytes were not detected. For M5, recoveries between 43.6 and 59.8 % were obtained. The concentrations found were very low, even three analytes were not detected. For M6, the lowest recoveries were found, between 14.3 and 49.6 %. It could be possible the presence of other substances which may be interfering in the extraction capability of the automatic SPE system, decreasing recoveries of all the analytes regarding the coffee matrix used in the optimization stage (V). The concentrations found were very low, even four analytes were not detected. For M7, recoveries nearly or above 50 % were obtained for all PAHs except for FLN. The concentrations found were very low, even two analytes were not detected, except for FLN in which case 1.0 μg L^{-1} and 14.3 μg kg^{-1} was found. For M8, recoveries between 40.2 and 57.6 % were obtained. The concentrations found were very low, even three analytes were not detected. For M9, recoveries between 36.9 and 62.3 % were obtained. The concentrations found were very low, even two analytes were not detected. Related to PAH4, the highest value was

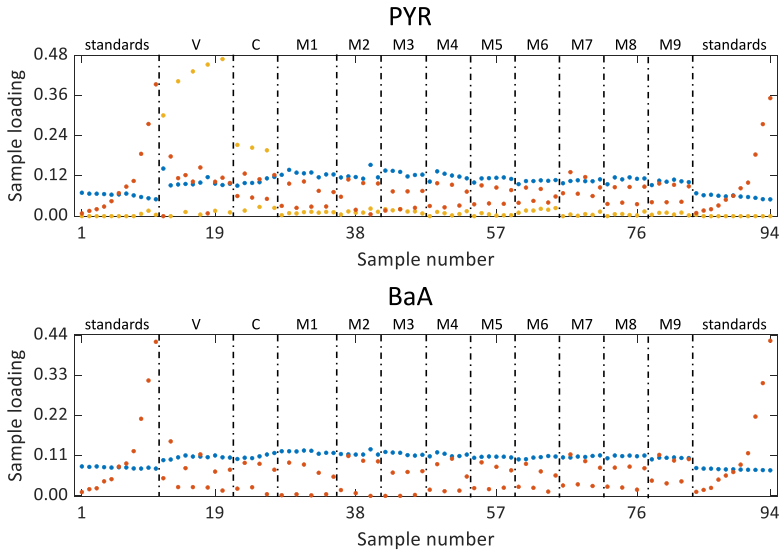


Fig. 6. Sample loadings of the PARAFAC models for coffee samples obtained for PYR and BaA, being the orange factor the analyte, the blue one the baseline, and the yellow one another compound. (Standards) refer to calibration standard solutions; (V) to validation coffee sample; (C) to control solutions prepared in coffee matrix; (M1 – M9) to the nine new coffee samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Calibration and accuracy lines parameters, fitted with the sample loadings from the PARAFAC models for the determination of PAHs in coffee samples.

		PHE	ANT	FLN	PYR	CHR	BaA	PER	BbF	BaP
Concentration range (μg L^{-1})		0 – 40	0 – 60	0 – 10	0 – 10	0 – 10	0 – 4	0 – 10	0 – 10	0 – 4
Calibration line	Outliers/standards	1/22	1/22	1/22	1/22	0/22	0/22	3/22	0/22	1/22
	Intercept	−0.0061	−0.0003	−0.0128	0.0128	0.0029	0.0125	−0.0121	0.0222	−0.0097
	Slope	0.0103	0.0079	0.0388	0.0346	0.0433	0.1016	0.0403	0.0376	0.1122
	Correlation coefficient	0.9988	0.9985	0.9943	0.9994	0.9983	0.9991	0.9939	0.9959	1.0000
	s_{yx}	0.0065	0.0084	0.0124	0.0034	0.0084	0.0059	0.0125	0.0114	0.0066
	P-value (H_0: Regression is not significant)	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$
Accuracy line	Intercept	$2.90\ 10^{-3}$	$3.37\ 10^{-3}$	$6.86\ 10^{-4}$	$-6.46\ 10^{-4}$	$-1.50\ 10^{-4}$	$-1.26\ 10^{-4}$	$6.01\ 10^{-4}$	$-2.61\ 10^{-4}$	$4.26\ 10^{-4}$
	Slope	0.9998	0.9999	0.9998	1.0005	1.0002	1.0004	0.9998	1.0000	0.9995
	Correlation coefficient	0.9988	0.9985	0.9942	0.9995	0.9983	0.9991	0.9939	0.9959	0.9990
	s_{yx}	0.6379	1.0696	0.3200	0.0980	0.1957	0.0575	0.3101	0.3025	0.0588
	P-value (H_0: Intercept equal to zero and slope equal to one)	0.9998	0.9999	1.0000	0.9975	0.9998	0.9988	1.0000	1.0000	0.9988

found for M9 (coffee roasted in the presence of sugar), being 0.5 μg L^{-1} and 7.8 μg kg^{-1}.

5. Conclusions

In this paper, the optimization of an automatic SPE system in the determination of nine PAHs from coffee samples was carried out. The D-optimal design led to saving time and costs in the laboratory maintaining the reliability of estimates, reducing the number of experiments from 72 to 19. As responses for this D-optimal design, difference of loadings from PARAFAC were considered, since PARAFAC is a tool which allows the unequivocal identification of analytes.

While the determination of the optimum through the analysis of the effects leads to a unique solution of the experimental conditions, the application of AQbD in the analysis of the experimental data allows exploring the full experimental domain. By this way, the knowledge of the automatic SPE system in relation to the nine PAHs considered is wider.

Once the method was developed, values of CCβ (detection capability) were calculated with the probabilities of false positive and false negative fixed at 0.05, being 2.2, 3.8, 1.1, 0.3, 0.7, 0.2, 1.1, 1.1, and 0.2 μg L^{-1} for PHE, ANT, FLN, PYR, CHR, BaA, PER, BbF, and BaP, respectively.

The highest values were found for coffee sample M9 (coffee roasted in the presence of sugar), with 0.1 μg L^{-1} in liquid coffee and 1.4 μg kg^{-1} in coffee ground for BaP, whereas for PAH4 were 0.5 μg L^{-1} and 7.8 μg kg^{-1}, respectively.

Credit author statement

L. Valverde-Som: Formal analysis, Investigation, Methodology; Writing - original draft, Writing - review & editing.

M.M. Arce: Formal analysis, Investigation, Methodology; Writing - original draft, Writing - review & editing.

L.A. Sarabia: Conceptualization, Methodology, Supervision; Writing - original draft, Writing - review & editing.

M.C. Ortiz: Conceptualization, Funding acquisition, Methodology, Supervision, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table 5
Results obtained for different coffee samples. CCα (decision limit) and CCβ (detection capability). (V) refers to validation coffee sample; (M1 – M9) to the nine new coffee samples.

Code coffee samples		PHE	ANT	FLN	PYR	CHR	BaA	PER	BbF	BaP	PAH4 #
V	CCα (μg L^{-1})	1.1	1.9	0.6	0.2	0.3	0.1	0.6	0.5	0.1	
	CCβ (μg L^{-1})	2.2	3.8	1.1	0.3	0.7	0.2	1.1	1.1	0.2	
	Spiked concentration added (μg L^{-1})	12.0	18.0	3.0	3.0	3.0	1.2	3.0	3.0	1.2	
	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 5)	4.2	2.2*	2.1	1.6	n.d.	0.1*	2.6	1.0*	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 5)	10.2	9.3	3.8	3.4	1.6	0.8	3.3	2.6	0.9	
	Recovery (%)	49.5	39.3	56.5	59.4	54.8	56.2	22.9	50.6	57.8	
	Concentration in liquid coffee (μg L^{-1})	1.6	1.1	0.7	0.5	n.d.	0.0(5)	2.1	0.4	0.1	0.5
	Concentration in coffee ground (μg kg^{-1})	23.2	15.4	10.2	7.3	n.d.	0.7	30.4	5.5	0.9	7.1
M1	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 4)	4.5	n.d.	1.0*	0.4	n.d.	n.d.	1.3	n.d.	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 4)	9.9	7.9	2.8	2.1	1.7	0.6	3.0	0.9*	0.8	
	Recovery (%)	45.2	44.0	59.1	56.7	55.4	49.8	57.7	30.2	55.3	
	Concentration in liquid coffee (μg L^{-1})	1.9	n.d.	0.3	0.1	n.d.	n.d.	0.4	n.d.	0.1	0.1
	Concentration in coffee ground (μg kg^{-1})	26.9	n.d.	4.5	2.1	n.d.	n.d.	6.2	n.d.	0.9	0.9
M2	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	3.5	n.d.	0.7*	0.4	0.4*	n.d.	1.2	n.d.	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	11.0	11.1	3.0	2.6	2.3	0.9	3.4	1.6	1.0	
	Recovery (%)	62.7	61.5	78.1	71.9	61.4	71.8	71.9	51.8	71.9	
	Concentration in liquid coffee (μg L^{-1})	1.0	n.d.	0.2	0.1	0.1	n.d.	0.3	n.d.	0.0(4)	0.2
	Concentration in coffee ground (μg kg^{-1})	14.9	n.d.	2.3	1.6	1.9	n.d.	4.6	n.d.	0.6	2.5
M3	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	6.2	n.d.	n.d.	0.2*	0.5*	n.d.	1.0*	n.d.	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	10.3	6.2	1.2	1.8	2.0	0.5	2.9	n.d.	0.8	
	Recovery (%)	34.3	34.6	41.6	51.7	50.0	43.6	62.1	–	58.0	
	Concentration in liquid coffee (μg L^{-1})	3.4	n.d.	n.d.	0.1	0.2	n.d.	0.3	n.d.	0.1	0.2
	Concentration in coffee ground (μg kg^{-1})	50.5	n.d.	n.d.	1.2	2.9	n.d.	4.7	n.d.	0.7	3.6
M4	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	2.8	n.d.	n.d.	0.5	n.d.	n.d.	1.1	n.d.	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	10.1	8.3	2.2	2.2	1.9	0.7	2.9	1.1	0.9	
	Recovery (%)	61.3	45.9	72.1	56.8	64.8	55.9	60.4	36.5	60.0	
	Concentration in liquid coffee (μg L^{-1})	0.8	n.d.	n.d.	0.2	n.d.	n.d.	0.4	n.d.	0.1	0.1
	Concentration in coffee ground (μg kg^{-1})	11.7	n.d.	n.d.	2.2	n.d.	n.d.	4.9	n.d.	0.8	0.8
M5	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	2.9	2.1*	1.7	0.7	n.d.	n.d.	1.4	n.d.	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	8.7	10.0	3.0	2.1	1.4	0.7	2.8	1.8	0.8	
	Recovery (%)	48.2	43.7	43.6	46.5	46.4	55.8	44.6	59.8	53.6	
	Concentration in liquid coffee (μg L^{-1})	1.1	0.9	0.7	0.3	n.d.	n.d.	0.6	n.d.	0.1	0.1
	Concentration in coffee ground (μg kg^{-1})	15.6	12.6	10.2	3.9	n.d.	n.d.	8.3	n.d.	0.9	0.9
M6	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	4.5	n.d.	0.8*	0.8	n.d.	n.d.	1.3	n.d.	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	8.1	2.6	1.3	1.8	1.5	0.6	2.7	1.0*	0.7	
	Recovery (%)	29.6	14.3	16.7	31.9	49.6	47.4	47.3	32.9	46.2	
	Concentration in liquid coffee (μg L^{-1})	2.9	n.d.	0.8	0.5	n.d.	n.d.	0.5	n.d.	0.1	0.1
	Concentration in coffee ground (μg kg^{-1})	42.5	n.d.	12.6	7.4	n.d.	n.d.	7.5	n.d.	1.0	1.0
M7	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	4.0	n.d.	2.0	1.5	n.d.	0.2*	1.2	0.6*	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	10.4	8.6	3.1	2.9	1.7	0.8	3.2	2.2	0.9	
	Recovery (%)	53.7	47.7	36.6	44.5	57.3	54.2	66.3	53.5	61.4	
	Concentration in liquid coffee (μg L^{-1})	1.4	n.d.	1.0	0.6	n.d.	0.1	0.3	0.2	0.1	0.3
	Concentration in coffee ground (μg kg^{-1})	19.2	n.d.	14.3	8.9	n.d.	0.8	4.6	2.9	0.7	4.4
M8	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	3.5	n.d.	1.3	0.7	n.d.	n.d.	1.2	0.6*	0.2*	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	8.9	9.2	2.5	2.2	1.5	0.6	2.9	1.9	0.8	
	Recovery (%)	45.6	51.3	40.2	47.9	50.2	53.5	57.6	41.8	55.2	
	Concentration in liquid coffee (μg L^{-1})	1.4	n.d.	0.6	0.3	n.d.	n.d.	0.4	0.3	0.1	0.3
	Concentration in coffee ground (μg kg^{-1})	19.7	n.d.	8.5	3.9	n.d.	n.d.	5.3	4.0	0.8	4.7
M9	Predicted concentration in blank coffee sample (μg L^{-1}) (n = 3)	3.7	n.d.	n.d.	0.8	0.4*	0.3	0.8*	0.6*	0.3	
	Predicted concentration in spiked coffee sample (μg L^{-1}) (n = 3)	9.2	6.6	1.5	2.3	2.3	0.9	2.3	2.3	1.0	
	Recovery (%)	45.7	36.9	49.8	49.5	62.3	53.0	48.1	55.9	58.0	
	Concentration in liquid coffee (μg L^{-1})	1.5	n.d.	n.d.	0.3	0.1	0.1	0.3	0.2	0.1	0.5
	Concentration in coffee ground (μg kg^{-1})	22.9	n.d.	n.d.	4.8	1.8	1.4	4.9	3.2	1.4	7.8

(n.d.) Not detected because value < CCα; (*) Value between CCα and CCβ; (–) Recovery cannot be calculated; (#) Sum of BaP, BaA, BbF and CHR.

Data availability

The authors do not have permission to share data.

Acknowledgements

The work is financed by the Consejería de Educación de la Junta de Castilla y León through BU052P20 project, co-financed with European FEDER funds. L. Valverde-Som and M.M. Arce wish to thank Junta de Castilla y León for their postdoctoral contract through BU052P20

project.

References

[1] P.L. Peng, L.H. Lim, Polycyclic aromatic hydrocarbons (PAHs) sample preparation and analysis in beverages: a review, Food Anal. Methods 15 (2022) 1042–1061, https://doi.org/10.1007/s12161-021-02178-y.

[2] L. Singh, J.G. Varshney, T. Agarwal, Polycyclic aromatic hydrocarbons' formation and occurrence in processed food, Food Chem. 199 (2016) 768–781, https://doi.org/10.1016/j.foodchem.2015.12.074.

[3] A. Sadowska-Rociek, M. Surma, E. Cieślik, Determination of polycyclic aromatic hydrocarbons in coffee and coffee substitutes using dispersive SPE and gas chromatography-mass spectrometry, Food Anal. Methods 8 (2015) 109–121, https://doi.org/10.1007/s12161-014-9876-9.

[4] United States Environmental Protection Agency, Toxic and priority pollutants under the clean water act. https://www.epa.gov/eg/toxic-and-priority-pollutants-under-clean-water-act, 2021. (Accessed 17 July 2023).

[5] International Agency for Research on Cancer, IARC Monographs On The Identification Of Carcinogenic Hazards To Humans, World Health Organization, 2023. https://monographs.iarc.who.int/list-of-classifications. (Accessed 17 July 2023).

[6] Commission regulation (EU) 2023/915 of 25 April 2023 on maximum levels for certain contaminants in food and repealing Regulation (EC) No 1881/2006, Off. J. Eur. Union L 119 (2023) 103–157.

[7] P. Plaza-Bolaños, A.G. Frenich, J.L.M. Vidal, Polycyclic aromatic hydrocarbons in food and beverages. Analytical methods and trends, J. Chromatogr. A 1217 (2010) 6303–6326, https://doi.org/10.1016/j.chroma.2010.07.079.

[8] M.M. Arce, S. Sanllorente, S. Ruiz, M.S. Sánchez, L.A. Sarabia, M.C. Ortiz, Method operable design region obtained with a partial least squares model inversion in the determination of ten polycyclic aromatic hydrocarbons by liquid chromatography with fluorescence detection, J. Chromatogr. A 1657 (2021), 462577, https://doi.org/10.1016/j.chroma.2021.462577.

[9] Waters Corporation, Polycyclic aromatic hydrocarbons in drinking water. https://www.waters.com/webassets/cms/library/docs/oasis127.pdf, 2003. (Accessed 17 July 2023).

[10] J.K. Houessou, C. Benac, C. Delteil, V. Camel, Determination of polycyclic aromatic hydrocarbons in coffee brew using solid-phase extraction, J. Agric. Food Chem. 53 (2005) 871–879, https://doi.org/10.1021/jf048633a.

[11] J.K. Houessou, C. Delteil, V. Camel, Investigation of sample treatment steps for the analysis of polycyclic aromatic hydrocarbons in ground coffee, J. Agric. Food Chem. 54 (2006) 7413–7421, https://doi.org/10.1021/jf060802z.

[12] S. Catena, S. Sanllorente, L.A. Sarabia, R. Boggia, F. Turrini, M.C. Ortiz, Unequivocal identification and quantification of PAHs content in ternary synthetic mixtures and in smoked tuna by means of excitation-emission fluorescence spectroscopy coupled with PARAFAC, Microchem. J. 154 (2020), 104561, https://doi.org/10.1016/j.microc.2019.104561.

[13] S.H. Loh, P.E. Neoh, C.T. Tai, S. Kamaruzaman, Simple μ-solid phase extraction using C18 film for the extraction of polycyclic aromatic hydrocarbons in coffee beverage, Malaysia J. Anal. Sci. 22 (2018) 1–7, https://doi.org/10.17576/mjas-2018-2201-01.

[14] A. Jimenez, A. Adisa, C. Woodham, M. Saleh, Determination of polycyclic aromatic hydrocarbons in roasted coffee, J. Environ. Sci. Health B 49 (2014) 828–835, https://doi.org/10.1080/03601234.2014.938552.

[15] M.S. García-Falcón, B. Cancho-Grande, J. Simal-Gándara, Minimal clean-up and rapid determination of polycyclic aromatic hydrocarbons in instant coffee, Food Chem. 90 (2005) 643–647, https://doi.org/10.1016/j.foodchem.2004.03.059.

[16] N. Rattanakunsong, S. Jullakan, J. Płotka-Wasylka, O. Bunkoed, A hierarchical porous composite magnetic sorbent of reduced graphene oxide embedded in polyvinyl alcohol cryogel for solvent-assisted-solid phase extraction of polycyclic aromatic hydrocarbons, J. Separ. Sci. 45 (2022) 1774–1783, https://doi.org/10.1002/jssc.202200041.

[17] Y. Shi, H. Wu, C. Wang, X. Guo, J. Du, L. Du, Determination of polycyclic aromatic hydrocarbons in coffee and tea samples by magnetic solid-phase extraction coupled with HPLC–FLD, Food Chem. 199 (2016) 75–80, https://doi.org/10.1016/j.foodchem.2015.11.137.

[18] Waters Corporation, Oasis HLB cartridges and 96-well plates. https://www.waters.com/webassets/cms/support/docs/715000109.pdf, 2014. (Accessed 17 July 2023).

[19] P. Ramalingam, B. Jahnavi, Chapter 5 - QbD considerations for analytical development, in: S. Beg, M.S. Hasnain (Eds.), Pharmaceutical Quality by Design, Academic Press London Wall, London, 2019, pp. 77–108, https://doi.org/10.1016/B978-0-12-815799-2.00005-8.

[20] S. Beg, J. Haneef, M. Rahman, R. Peraman, M. Taleuzzaman, W.H. Almalki, Chapter 1 - introduction to analytical quality by design, in: S. Beg, M.S. Hasnain, M. Rahman, W.H. Almalki (Eds.), Handbook of Analytical Quality by Design, Academic Press, 2021, pp. 1–14, https://doi.org/10.1016/B978-0-12-820332-3.00009-1.

[21] S. Beg, M. Rahman, Chapter 5 - analytical quality by design for liquid chromatographic method development, in: S. Beg, M.S. Hasnain, M. Rahman, W. H. Almalki (Eds.), Handbook of Analytical Quality by Design, Academic Press, 2021, pp. 87–97, https://doi.org/10.1016/B978-0-12-820332-3.00010-8.

[22] M.M. Arce, M.C. Ortiz, S. Sanllorente, Univariate data analysis *versus* multivariate approach in liquid chromatography. An application for melamine migration from food contact materials, Microchem. J. 181 (2022), 107648, https://doi.org/10.1016/j.microc.2022.107648.

[23] M.C. Ortiz, S. Sanllorente, A. Herrero, C. Reguera, L. Rubio, M.L. Oca, L. Valverde-Som, M.M. Arce, M.S. Sánchez, L.A. Sarabia, Three-way PARAFAC decomposition of chromatographic data for the unequivocal identification and quantification of compounds in a regulatory framework, Chemometr. Intell. Lab. 200 (2020), 104003, https://doi.org/10.1016/j.chemolab.2020.104003.

[24] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Winding, R.S. Koch, PLS Toolbox 8.8.1, Eigenvector Research Inc., Wenatchee, WA, USA, 2020.

[25] MATLAB, The Mathworks, Inc., Natick, MA, USA, 2019 version 9.7.0.1190202 (R2019b).

[26] D. Mathieu, J. Nony, R. Phan-Than-Lu, NEMRODW (Version 2015), L.P.R.A.I, Marseille, France, 2015.

[27] M.M. Arce, S. Sanllorente, M.C. Ortiz, L.A. Sarabia, Easy-to-use procedure to optimise a chromatographic method. Application in the determination of bisphenol-A and phenol in toys by means of liquid chromatography with fluorescence detection, J. Chromatogr. A 1534 (2018) 93–100, https://doi.org/10.1016/j.chroma.2017.12.049.

[28] STATGRAPHICS Centurion 19, Statpoint Technologies, Inc., Herndon, VA, USA, 2023. Version 19.4.01.

[29] L.A. Sarabia, M.C. Ortiz, DETARCHI: a program for detection limits with specified assurance probabilities and characteristic curves of detection, TrAC-Trend. Anal. Chem. 13 (1994) 1–6, https://doi.org/10.1016/0165-9936(94)85052-6.

[30] G.A. Lewis, D. Mathieu, R. Phan-Tan-Luu, Pharmaceutical Experimental Design, Marcel Dekker, New York, 1999.

[31] R. Cela, R. Phan-Tan-Luu, Screening strategies, in: S. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics, second ed., Elsevier, 2020, pp. 209–250, https://doi.org/10.1016/B978-0-12-409547-2.14755-9.

[32] A. Herrero, S. Sanllorente, C. Reguera, M.C. Ortiz, L.A. Sarabia, A new multiresponse optimization approach in combination with a D-Optimal experimental design for the determination of biogenic amines in fish by HPLC-FLD, Anal. Chim. Acta 945 (2016) 31–38, https://doi.org/10.1016/j.aca.2016.10.001.

[33] M.M. Arce, S. Ruiz, S. Sanllorente, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A new approach based on inversion of a partial least squares model searching for a preset analytical target profile. Application to the determination of five bisphenols by liquid chromatography with diode array detector, Anal. Chim. Acta 1149 (2021), 338217, https://doi.org/10.1016/j.aca.2021.338217.

[34] A. Inselberg, Discovering and visualizing relations in high dimensional data, in: J. E. Gentle, W.K. Härdle, Y. Moriet (Eds.), Handbook of Computational Statistics, Springer-Verlag Berlin Heidelberg, 2012, pp. 299–333, https://doi.org/10.1007/978-3-642-21551-3_11.

[35] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, D. Arroyo, Improving the visualization of the Pareto-optimal front for the multi-response optimization of chromatographic determinations, Anal. Chim. Acta 687 (2011) 129–136, https://doi.org/10.1016/j.aca.2010.12.023.

[36] R. Güzel, Z.C. Ertekin, E. Dinç, A new application of PARAFAC model to UPLC dataset for the quantitative resolution of a tri-component drug mixture, J. Chromatogr. Sci. 59 (2021) 361–370, https://doi.org/10.1093/chromsci/bmaa119.

[37] M.B. Anzardi, J.A. Arancibia, A.C. Olivieri, Processing multi-way chromatographic data for analytical calibration, classification and discrimination: a successful marriage between separation science and chemometrics, TrAC-Trend. Anal. Chem. 134 (2021), 116128, https://doi.org/10.1016/j.trac.2020.116128.

[38] K.S. Booksh, B.R. Kowalski, Theory of analytical chemistry, Anal. Chem. 66 (1994) 782A–791A, https://doi.org/10.1021/ac00087a718.

[39] M.C. Barreto, R.G. Braga, S.G. Lemos, W.D. Fragoso, Determination of melamine in milk by fluorescence spectroscopy and second-order calibration, Food Chem. 364 (2021), 130407, https://doi.org/10.1016/j.foodchem.2021.130407.

[40] H. Abdollahi, R. Tauler, Uniqueness and rotation ambiguities in multivariate curve resolution methods, Chemometr. Intell. Lab. 108 (2011) 100–111, https://doi.org/10.1016/j.chemolab.2011.05.009.

[41] M.D. Carabajal, R.P. Vidal, J.A. Arancibia, A.C. Olivieri, A new constraint to model background signals when processing chromatographic-spectral second-order data with multivariate curve resolution, Anal. Chim. Acta 1266 (2023), 341354, https://doi.org/10.1016/j.aca.2023.341354.

[42] A.C. Olivieri, Evaluation of the ambiguity in second-order analytical calibration based on multivariate curve resolution. A tutorial, Microchem. J. 179 (2022), 107455, https://doi.org/10.1016/j.microc.2022.107455.

[43] A.C. Olivieri, R. Tauler, N-BANDS: a new algorithm for estimating the extension of feasible bands in multivariate curve resolution of multicomponent systems in the presence of noise and rotational ambiguity, J. Chemometr. 35 (2021) e3317, https://doi.org/10.1002/cem.3317.

[44] C.A. Andersson, R. Bro, The N-way toolbox for MATLAB, Chemometr. Intell. Lab. 52 (2000) 1–4, https://doi.org/10.1016/S0169-7439(00)00071-X.

[45] R. Bro, Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis, Chemometr. Intell. Lab. 46 (1999) 133–147, https://doi.org/10.1016/S0169-7439(98)00181-6.

[46] R. Bro, H.A.L. Kiers, A new efficient method for determining the number of components in PARAFAC models, J. Chemometr. 17 (2003) 274–286, https://doi.org/10.1002/cem.801.

[47] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A. Herrero, S. Sanllorente, C. Reguera, Usefulness of PARAFAC for the quantification, identification, and description of analytical data, in: A. Muñoz de la Peña, H.C. Goicoechea, G.M. Escandar, A. C. Olivieri (Eds.), Data Handling in Science and Technology: Fundamentals and Analytical Applications of Multiway Calibration, Elsevier, Amsterdam, 2015, pp. 37–81, https://doi.org/10.1016/B978-0-444-63527-3.00002-3.

[48] Commission implementing Regulation (EU), 2021/808 of 22 March 2021 on the performance of analytical methods for residues of pharmacologically active

substances used in food-producing animals and on the interpretation of results as well as on the methods to be used for sampling and repealing Decisions 2002/657/EC and 98/179/E, Off. J. Eur. Union L 180 (2021) 84–109.

[49] SANTE/11312/2021, Guidance Document on Analytical Quality Control and Method Validation Procedures for Pesticide Residues and Analysis in Food and Feed, European Commission, 2021 (implemented by 01 January 2022).

[50] L.A. Sarabia, M.C. Ortiz, M.S. Sánchez, Response surface methodology, in: S. D. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics, second ed., Elsevier, 2020, pp. 287–326, https://doi.org/10.1016/B978-0-12-409547-2.14756-0.

[51] International Organization for Standardization, ISO 11843, Capability of Detection, Part 1: Terms and Definitions and Part 2: Methodology in the Linear Calibration Case, Genève, Switzerland, 2000.

[52] Commission Decision (EC) No 2002/657/EC of 12 August 2002 Implementing council directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Off. J. Eur. Commun L 221 (2002) 8–36.

[53] Commission regulation (EU) No 836/2011 of 19 August 2011 amending Regulation (EC) No 333/2007 laying down the methods of sampling and analysis for the official control of the levels of lead, cadmium, mercury, inorganic tin, 3-MCPD and benzo(a)pyrene in foodstuffs, Off. J. Eur. Union L 215 (2011) 9–16.