Partial least squares model inversion in the chromatographic determination of triazines in water

M.C. Ortiz [a,*], L.A. Sarabia [b], A. Herrero [a], C. Reguera [a], S. Sanllorente [a], M.M. Arce [a], O. Valencia [b], S. Ruiz [b], M.S. Sánchez [b]

[a] *Dpt. of Chemistry (Analytical Chemistry), Facultad de Ciencias, Universidad de Burgos, Pza. Misael Bañuelos s/n, 09001 Burgos, Spain*
[b] *Dpt. of Mathematics and Computation, Facultad de Ciencias, Universidad de Burgos, Pza. Misael Bañuelos s/n, 09001 Burgos, Spain*

ARTICLE INFO

ABSTRACT

Inside the framework of Analytical Quality by Design, a model-based approach has been developed and used to identify operating conditions (control method parameters) related to the composition and flow rate of the mobile phase for a liquid chromatographic determination with preset quality characteristics.

The approach starts by defining these desired characteristics of the intended chromatogram (proper resolution for consecutive peaks and short time of analysis) and then looking for the needed control method parameters via inversion of a Partial Least Squares (PLS) prediction model.

The procedure has been applied to the determination of eight triazines (simazine, simetryn, atrazine, ametryn, propazine, terbuthylazine, prometryn and terbutryn) in surface waters by means of SPE-HPLC-DAD. These triazines either are forbidden or have a maximum allowable limit due to their potential toxicity.

The experimental verification of the selected parameters showed that the experimental results were significantly equal to those predicted. Besides, the validation of the developed method allowed concluding that accuracy was fulfilled for the eight triazines and there was not bias. With a probability of false positive equal to 0.05, CCβ was less than 3 µg L^{-1} for every triazine, except for simazine and terbutryn, which was less than 6 µg L^{-1} being the probability of false negative less than 10^{-6}.

No triazine was found, above their maximum allowable concentration, in any of the samples of surface water picked at fifteen different locations, mostly from streams and the Arlanzón river, near Burgos (Spain).

1. Introduction

The so-called Process Analytical Technology (PAT) [1] created by the US FDA (Food & Drug Administration) and harmonized [2] with the EMA (European Medicines Agency) in the field of regulation of pharmaceutical processes, establishes a flexible context for the maintenance of the quality of a product.

The Q8(R2) guideline [2] defines PAT as *a system for designing, analyzing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes with the goal of ensuring final product quality.*

An analytical method can be considered as a process that must have an output of acceptable quality [3], and the QbD concept for manufacturing processes could also be applied to analytical methods [4,5].

Consequently, the development of a chromatographic method to determine triazines is entirely included into the concept of Quality by Design (QbD), which is described in the annex of the Q8(R2) guideline 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*. At present, the implementation of the QbD approach in the development of analytical method is known as AQbD (Analytical Quality by Design), with parallel terminology, already acknowledged in document [6].

Under these premises, the paper addresses the development of a chromatographic method, designed as if it were a process but starting by defining the desired quality for the output. This is done by means of the so called Analytical Target Profile (ATP) and the goal is then to find the necessary experimental conditions to achieve this intended ATP. These

experimental conditions are the settings of the Control Method Parameters (CMP) that should be used when conducting the measurements to obtain chromatograms with the pre-defined characteristics.

Looking for the corresponding CMP requires building and validating a prediction model that relates the control method parameters with some ATP-related characteristics, and its inversion in a fully data driven product design context. As there are several CMP and several ATP-related characteristics to fulfill, the prediction model will be a latent variable model, precisely a PLS2 (Partial Least Squares) model. For that reason, the product design with these prediction models is also known as latent variable model inversion [7].

Another advantage of using a PLS2 model is that the Q and T^2 statistics provides a systematic procedure to decide whether the model can be applied to a new set of CMP. These valid configurations are in a region, the PLSbox [8], that is determined by imposing limits on the mentioned statistics with a confidence level.

In this paper, determination of eight triazines, namely simazine (SZ), simetryn (ST), atrazine (AZ), ametryn (AT), propazine (PZ), terbuthylazine (TZ), prometryn (PT) and terbutryn (TT), in surface waters was chosen to illustrate the methodology developed. Triazines are herbicides widely used all over the world to remove or inhibit the growth of unwanted plants. These compounds have some potential for movement (leaching) into soil, and eventually into surface waters. In the assessment of their environmental fate, atrazine and simazine are classified as moderately mobile in Commission Regulation (EU) 2016/266 [9], whereas terbuthylazine, prometryn and propazine are considered as slightly mobile. Low-level contamination of surface waters is a potential problem since many of the triazines have been suspected of being potential endocrine disruptors [10,11]. Atrazine was banned in the EU in 2004 [12] because of its persistence in the environment and toxicity for wildlife and possible effects on human health.

Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC, lays down environmental quality standards (EQS) for 45 priority substances in the field of water policy with the objective of achieving good surface water chemical status; atrazine, simazine and terbutryn are among these priority substances. In Annex II, maximum allowable concentrations (MAC-EQS) for inland surface waters of 2, 4 and 0.34 µg L^{-1} are established for atrazine, simazine and terbutryn, respectively. In addition, a Spanish national transposition, in Annex V of Real Decreto 817/2015 de 11 de septiembre [13], establishes an annual average (AA-EQS) of 1 µg L^{-1} for terbuthylazine, classified within the preferential substances that are substances presenting a significant risk to the Spanish surface waters on the basis of their particular toxicity, persistence and bioaccumulation or because the importance of their presence in the aquatic environment.

2. Material and methods

2.1. Reagents and standards

Simazine (CAS no. 122-34-9), simetryn (CAS no. 1014-70-6), atrazine (CAS no. 1912-24-9), ametryn (CAS no. 834-12-8), propazine (CAS no. 139-40-2), terbuthylazine (CAS no. 5915-41-3), prometryn (CAS no. 230-711-3) and terbutryn (CAS no. 212-950-5), PESTANAL grade, were purchased from Sigma–Aldrich (St. Louis, MO, USA). Acetonitrile, acetone and methanol, LiChrosolv® for liquid chromatography, were supplied by Merck (Darmstadt, Germany). Deionised water was obtained by using the Milli-Q gradient A10 water purification system from Millipore (Bedford, MA, USA).

Stock solutions containing 400 mg L^{-1} of each triazine were prepared in methanol. Intermediate solutions and final standard solutions were prepared daily in water to each contains the appropriate concentration of each triazine.

2.2. Instrumental and materials

The vacuum manifold used for the SPE step was purchased from Waters Corporation (Milford, MA, USA) and is coupled to a vacuum pump (Sartorius AG, Goettingen, Germany). SPE cartridges Oasis HLB 6 cc, 200 mg (Waters Corporation, Milford, MA, USA) were used. SPE cartridges were centrifuged using a Sigma 2–16 K refrigerated centrifuge (Osterode, Germany). The evaporation of the eluent was performed in a miVac Modular Concentrator (GeneVac Limited, Ipswich, UK), which consists of a miVac Duo concentrator, a SpeedTrap and a Quattro pump. A ZX3 vortex mixer (VELP Scientifica, Milan, Italy) was used for homogenizing reconstituted samples.

Analyses were carried out on an Agilent 1260 Infinity HPLC system (Agilent Technologies, Santa Clara, CA, USA) consisting of the quaternary pump VL (G1311C), a standard autosampler (G1329B), a thermostatted column compartment (G1316A) and a diode array detector (G7117C). A Kinetex 5.0 µm EVO C18 100 Å (150 mm length × 4.6 mm i. d.) analytical column (Phenomenex, Torrance, CA, USA) was used.

2.3. Standards and samples of water

To check the linear range of the method, for every triazine (SZ, ST, AZ, AT, PZ, TZ, PT and TT) a calibration model was fitted covering a large range of concentrations (specifically, 3, 6, 9, 12, 15, 18, 21, 24, 48, 75, 126, 150, 174, 201, 225, 249, 276 and 300 µg L^{-1}). Eight additional samples were determined to compute the prediction errors in the range from 3 to 249 µg L^{-1}, namely, samples with 3, 12 (in triplicate), 150 (in triplicate) and 249 µg L^{-1}. Finally, for the computation of the detection limits, seven standards were used with 3, 6, 9, 12, 18, 21, and 24 µg L^{-1}.

Surface water samples were collected in fifteen sites located in the province of Burgos, Spain. Exact locations are shown in Fig. S1(a) of the supplementary material, samples A to O. All samples come from rivers and streams, except for G, N and O, that correspond to springs and they were collected at the point of their emergence to the surface through fountains. Further, water from the fountains at the locations of samples N and O goes to the municipal network (sewerage) and, finally, at the location of point N there is a sign that specifies "water without sanitary guarantees".

The main criterion for the selection of the sampling sites was the closeness of the locations to orchards or crops, which could have been treated with pesticides containing triazines. Another factor is the mouth or confluence of some rivers or streams on others (described in the diagram in Fig. S1(b)) in order to interpret the results obtained when measuring the samples.

Sampling was carried out in a single day. Between a liter and a half and two liters were collected from each place in polyethylene bottles and stored at 4 °C in dark for further analysis.

2.4. Surface water sample preparation

Surface water samples were filtered through Whatman grade GF/C glass microfibre filters (Whatman, Maidstone, UK) and extracted with reversed phase SPE C18 cartridges. The cartridge was preconditioned with 10 mL of MeOH plus 10 mL of MilliQ grade water. 50 mL of the sample was loaded and percolated through the cartridge at a flow rate of 10 mL min^{-1}. The cartridge was washed with 20 mL of MilliQ grade water; residual water was removed effectively by centrifugation (2000 rpm, 5 min) and then using a stream of nitrogen (5 min). Triazines were eluted with 3 mL of acetone. The eluent was evaporated to dryness under vacuum in a centrifugal concentrator during 20 min at 40 °C. Finally, the residue was reconstituted with 1 mL of MilliQ grade water and transferred into a vial for the chromatographic analysis.

2.5. HPLC conditions

A volume of 20 µL of samples and standards was injected into the

chromatographic system. The separation was carried out under isocratic elution conditions. The column compartment was set at 25 °C. The mobile phase was a mixture of water–methanol–acetonitrile; the proportions of solvents differed according to the stage of the methodological approach, which includes their selection. In addition, different flow rates are set according to an experimental design. The final selected control method parameters were as follows: mobile phase was a mixture of acetonitrile–water (50:50, v/v) and a flow rate of 0.68 mL min^{-1}. The diode array detector was set at 225 nm.

2.6. Software

OpenLab CDS ChemStation software was used for acquiring data. The PLS2 models were fitted with the PLS_Toolbox [14] within MATLAB [15]. Inversion of the PLS2 model through the computation of the Pareto front was done with in-house programs written in MATLAB. The regression models were fitted and validated using STATGRAPHICS Centurion 18 [16]. The experimental design was selected with NEMRODW [17]. The capability of detection (CCβ) was calculated with DETARCHI [18]. The map that shows the exact coordinates of the location of the collected samples was arranged with Google Earth Pro [19].

3. Elements and description of the work procedure

In order to analyze the content of triazines in the collected water samples, an adequate chromatographic procedure should be stablished and validated. In the context of the AQbD (Analytical Quality by Design), a chromatographic determination is seen as a real process where the experimental conditions constitute the CMP (Control Method Parameters), which influence the obtained chromatogram. The CMP under consideration in this case are the ones that define the mobile phase, namely its composition and flow rate.

No all chromatograms show separation of the eight triazines under study or compliance with any other condition that can be of interest. The pursued conditions in terms of measurable characteristics of the chromatogram constitute the Analytical Target Profile (ATP). In this case, the ATP is defined imposing conditions for the resolution between consecutive peaks and for the initial and final time, nine values for every chromatogram.

As it has been explained in the introduction, once the ATP is defined, the question that remains is which CMP, if any, will produce a chromatogram with the characteristics in the ATP. The question is answered with the inversion of the corresponding prediction model, which is fitted to predict the nine ATP-related characteristics of the chromatogram from the CMP used to obtain it. That means that, in order to fit this model, some representative data need to be collected, from the own process.

The steps to accomplish the goals, which will be described for the case of the chromatographic determination that has been designed, are the following:

1. It is necessary to systematically cover the experimental domain, i. e., the region that contains viable values for the four CMP.

This was done by means of an appropriate experimental design, namely a combined mixture-process variable design to account for the mobile phase of the chromatographic determination, a ternary mixture (water, methanol and acetonitrile) and the flow rate.

2. The chromatograms obtained after conducting the experiments in the design should be quantitatively evaluated, in terms of the nine characteristics that define the ATP.

3. Using the nine values (responses) together, a PLS2 prediction model should be fitted to adequately predict the ATP-related characteristics of the expected chromatogram from given CMP in the experimental domain, and some squares or cross terms of CMP. This includes establishing the PLSbox [8] where the feasible conditions to apply the PLS2 model lie.

In this case, the boundary of the PLSbox was defined with the 95% confidence limits for the statistics Q and T^2.

4. The PLS2 model must be inverted to find the CMP needed to predict the defined ATP. The direct inversion [20] is unfeasible, so the computational alternative in [21] was used to look for the Pareto front of the conflicting characteristics demanded in the ATP.

A Pareto front in a multiobjective or multiresponse optimization context contains those situations where one of the responses reaches its best possible value to the point that it cannot be improved without worsening at least one other response. The method to seek these solutions uses a multiobjective evolutionary algorithm, summarized as:

4.1. Start with a population with a given number of settings for the CMP, i.e., different values of the composition of the mobile phase and its flow rate. These values are chosen with uniform distribution in the experimental domain, provided they satisfy the constraints imposed in the composition of the mobile phase and on the PLSbox.

4.2. Compute prediction for each member of the population, with the PLS2 model fitted, to obtain the nine-dimensional vector of ATP-related characteristics, which is compared to the defined ATP, in terms of the fitness function.

4.3. Apply selection, crossover and mutation operators to build new valid experimental conditions, which are also evaluated in terms of the fitness function.

4.4. Merge the old and newly generated population.

4.5. Arrange the members of the extended populations according to the Pareto order for multidimensional vectors.

4.6. Keep to survive for the next generation the non-dominated solutions and select, when needed, the most dispersal according to the crowding distance.

4.7. Repeat for a given number of generations.

4.8. Extract the Pareto front from the final population.

5. Exploration of the Pareto front provides insight about the effect of CMP on the ATP and allows selecting the needed CMP, selection usually based on expert knowledge.

6. The selected CMP are experimentally validated by conducting experiments, ten replicates in this case, with the chosen CMP. Then, the experimental results are compared with the predicted ones via confidence intervals for the mean, computed for both the experimental and the predicted results.

7. The developed chromatographic procedure, with these conditions, is validated in terms of accuracy and capability of detection.

8. The validated procedure is applied to quantify every triazine in the water samples.

4. Results and discussion
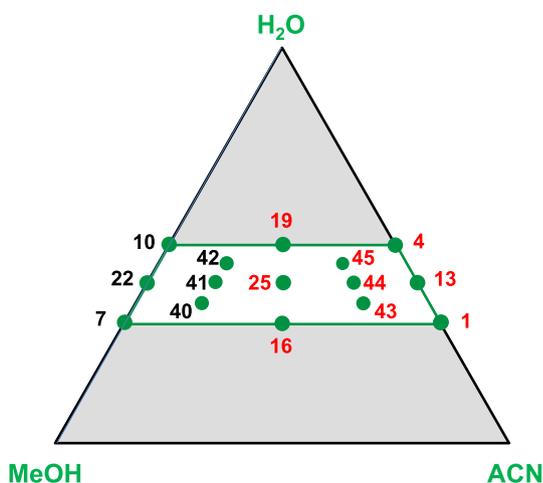
4.1. Experimental design

The Control Method Parameters (CMP) in the present case define the ternary mobile phase to be used, along with its flow rate. Therefore, there are four factors, three of them are components of a mixture, with constraints, and the fourth is a process variable, the flow rate.

With the goal of obtaining representative chromatograms from the whole experimental domain to be explored, a 'mixtures-process design' was used, which is adequate for linear models with interactions and quadratic terms [22].

The ternary mixture was made of water, methanol and acetonitrile with these constraints: the content of water, Z_1, should be between 30 and 50%, and the proportion of any of the organic solvents, Z_2 and Z_3, should not be greater than 70%. The process variable on its part was explored in the interval [0.40, 0.80] mL min^{-1} with 3 levels, 0.4, 0.6 and 0.8 mL min^{-1}.

The application of the constraints on the components of the mixture produced a reduced domain, a constraint simplex per level of the flow rate, which is depicted in Fig. 1 for one of the values of flow rate.

It also shows the distribution of the planned experiments (the

Fig. 1. Distribution of the experiments in the restricted simplex for the mixtures, at one of the levels of flow rate. The green dots constitute the experimental design, numbers in red highlight the experiments kept after further narrowing the experimental domain when fitting the final PLS2 model. The numbers are those of Table 1 to identify the composition.

design), the green filled circles, numbered according to the rows in Table 1. The mixture design is the so-called Scheffé design for a polynomial of degree two with four test points (which are experiments 40, 42, 43 and 45 of Fig. 1 and Table 1). The four test points are added because, otherwise, it would be only three ternary mixtures. The same reason explains the two new ternary mixtures (experiments number 41 and 44 in Fig. 1 and Table 1). The inclusion of these additional six experiments is a result of the need of handling ternary mixtures for the composition of the mobile phase. The same distribution is repeated for the three levels of the flow rate. In total, the experimental design consisted of 48 experiments, the 15 different experiments per level of flow rate in Table 1 plus one additional replicate at the centroid of the reduced simplex (experiment number 25 in Fig. 1). The same design is repeated for flow rate at 0.6 and 0.8 mL min^{-1}.

4.2. Prediction model

With the 48 resulting chromatograms, the nine ATP-related characteristics were computed, namely the resolution between consecutive peaks, R_{12}, R_{23}, R_{34}, R_{45}, R_{56}, R_{67}, and R_{78}, the initial time, t_i, and the total (final) time, t_f, both in minutes. The resolution $R_{i,i+1}$ between the consecutive i-*th* and (i + 1)-*th* chromatographic peaks is computed by

means of the following Eq. (1).

$$R_{i,i+1} = \frac{2.35\left(t_{R,i+1} - t_{R,i}\right)}{2\left(w_{0.5,i+1} + w_{0.5,i}\right)} \quad (1)$$

where $t_{R,i}$ is the retention time and $w_{0.5,i}$ is the width at half height of the i-*th* chromatographic peak. Matrix **Y** of responses was thus 48×9.

The supposed form of the PLS2 model to be fitted to the experimental data should take into account the characteristics of the procedure, so that it contains:

1. The terms that constitute the Control Method Parameters, that is, the composition of the mobile phase and its flow rate (Z_1, Z_2, Z_3, X_4), which are the main factors, the only ones that can be modified inside the domain.
2. Some crossed terms to take into account the possible expected interaction among the composition of the mixture and between them and the flow rate, and possible quadratic effects among all the variables.

Therefore, the predictor variables for fitting the PLS2 model have to account for these possible interactions and quadratic effects as shown in Eq. (2), which has 19 coefficients.

$$\begin{aligned}
\mathbf{Y} =\ & \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 X_4 + \\
& \beta_{12} Z_1 Z_2 + \beta_{13} Z_1 Z_3 + \beta_{23} Z_2 Z_3 + \beta_{41} X_4 Z_1 + \beta_{42} X_4 Z_2 + \beta_{43} X_4 Z_3 + \\
& \beta_{412} X_4 Z_1 Z_2 + \beta_{413} X_4 Z_1 Z_3 + \beta_{423} X_4 Z_2 Z_3 + \\
& \beta_{441} X_4^2 Z_1 + \beta_{442} X_4^2 Z_2 + \beta_{443} X_4^2 Z_3 + \\
& \beta_{4412} X_4^2 Z_1 Z_2 + \beta_{4413} X_4^2 Z_1 Z_3 + \beta_{4423} X_4^2 Z_2 Z_3
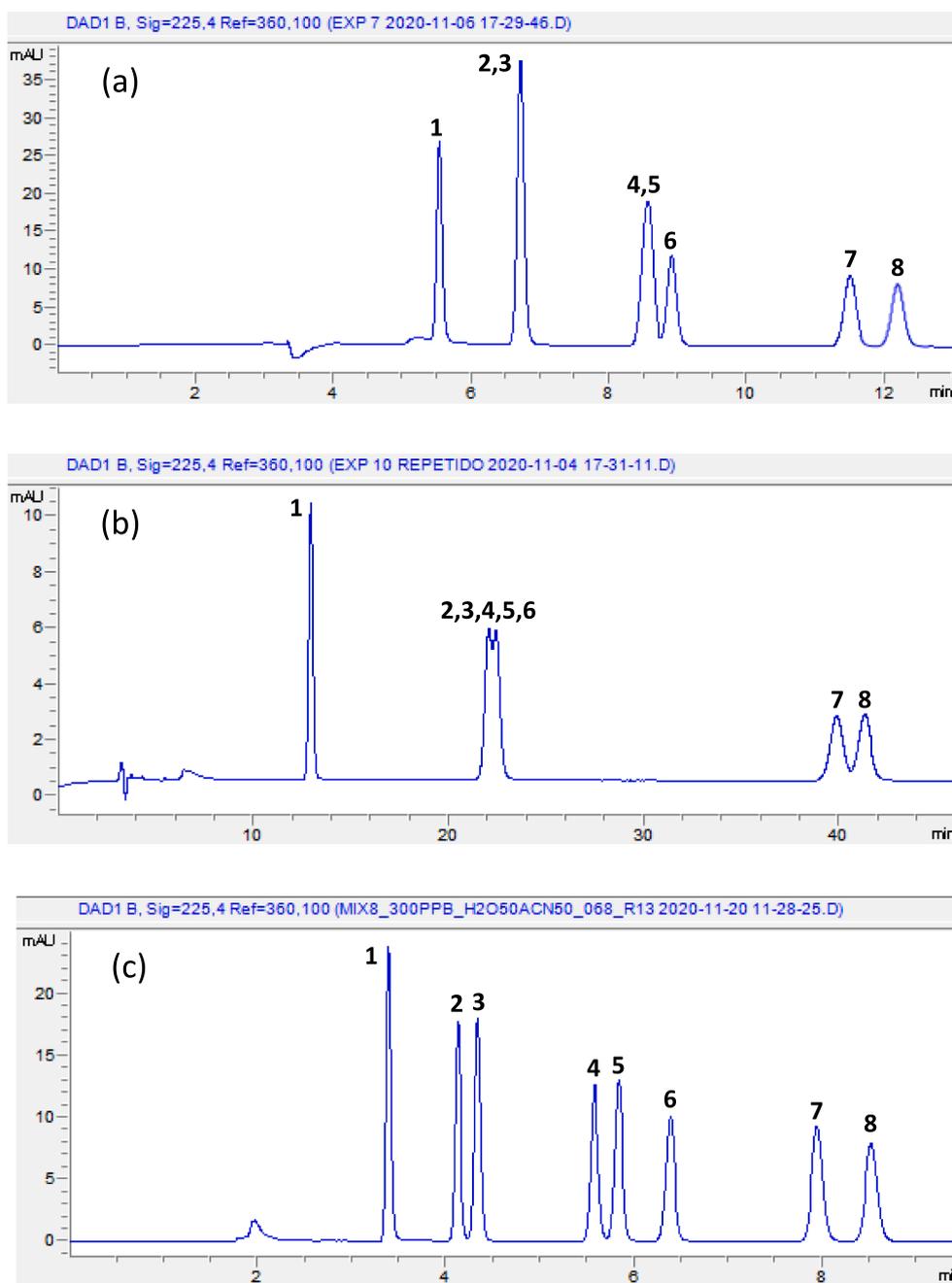\end{aligned} \quad (2)$$

It was observed that, in all the experiments that are not in bold in Table 1, only 6 or fewer chromatographic peaks appeared instead of one peak for each of the 8 triazines that had been initially injected into the chromatograph (i.e., two or more peaks were totally overlapped). Those experiments corresponded to a composition of the mobile phase without acetonitrile or with the percentage of methanol greater than that of acetonitrile, regardless of the flow rate of the mobile phase. Fig. 2 shows three chromatograms corresponding, specifically, to experiments number 7 (Fig. 2a) and 10 (Fig. 2b) in Table 1, both with several overlapping peaks, and both with binary methanol–water mixtures and the slowest flow rate, 0.4 mL min^{-1}.

It seems obvious that the mathematical model cannot model peaks that do not appear in the chromatogram. Even so, a PLS2 model was initially fit with the 48 analyses (chromatograms) trying to predict the nine responses (resolutions between contiguous peaks and initial and final time of the chromatogram), but it had a poor predictive capacity.

This led to narrowing the experimental domain, imposing an additional constraint, that the composition of acetonitrile be greater than or equal to that of methanol ($Z_3 \geq Z_2$) to avoid the total overlap observed in some peaks. Consequently, instead of 15 experiments per level of the flow rate, only the nine marked in bold in Table 1 were used. In the chromatograms corresponding to these nine experimental conditions, mostly eight chromatographic peaks can be observed even if they were partially overlapped, so that non-null resolutions were obtained when applying Eq. (1).

These nine experiments, in red in Fig. 1 and in bold in Table 1, repeated in each of the three level of flow rate, constituted the 27 chromatograms finally taken into account. Therefore, the PLS2 model was fitted with the predictor matrix **X** (27×19) and the response matrix **Y** (27×9).

Like in the previous case, in the present model and in all the following models, the corresponding raw variables were auto-scaled and the selection of the number of latent variables was made by using leave-one-out cross-validation. Besides, as it has been already stated, to define the region of the experimental domain where the model can be applied, the PLSbox [8], the threshold values at 95% confidence level for *Q* and

Table 1
Experimental design graphically depicted in Fig. 1 for a flow rate of 0.4 mL min^{-1}. In bold, the nine experiments (per level of flow rate) kept to fit the final PLS2 model.

Number in Fig. 1	Z_1(water)	Z_2(methanol)	Z_3(acetonitrile)	X_4 (flow rate, mL min^{-1})
1	**0.30**	**0**	**0.70**	**0.4**
4	**0.50**	**0**	**0.50**	**0.4**
7	0.30	0.70	0	0.4
10	0.50	0.50	0	0.4
13	**0.40**	**0**	**0.60**	**0.4**
16	**0.30**	**0.35**	**0.35**	**0.4**
19	**0.50**	**0.25**	**0.25**	**0.4**
22	0.40	0.60	0	0.4
25	**0.40**	**0.30**	**0.30**	**0.4**
40	0.35	0.50	0.15	0.4
41	0.40	0.45	0.15	0.4
42	0.45	0.40	0.15	0.4
43	**0.35**	**0.15**	**0.50**	**0.4**
44	**0.40**	**0.15**	**0.45**	**0.4**
45	**0.45**	**0.15**	**0.40**	**0.4**

Fig. 2. Chromatograms recorded from a sample containing 300 μg L^{-1} of each triazine: (a) for run 7, and (b) for run 10 of the experimental design in Table 1; (c) chromatogram obtained in the selected conditions found for the composition and flow rate of the mobile phase, ACN-H$_2$O (50:50, v/v) and 0.68 mL min^{-1}. Peak labels: 1, simazine; 2, simetryn; 3, atrazine; 4, ametryn; 5, propazine; 6, terbuthylazine; 7, prometryn; and 8, terbutryn.

T^2 statistics were always imposed. Finally, the final time had to be transformed as $Y_9 = t_f^{1/1.5}$ to obtain a better fit.

After fitting the model, a selection of variables was performed, keeping those variables whose VIP (Variable Importance in the Projection) score [23] was greater than one in at least one of the nine responses. The resulting reduced PLS2 model with the nine variables thus selected is in Eq. (3).

$$\begin{aligned}
\mathbf{Y} = \ &\beta_1 Z_1 + \beta_3 Z_3 + \\
&\beta_{12} Z_1 Z_2 + \beta_{13} Z_1 Z_3 + \beta_{41} X_4 Z_1 + \beta_{43} X_4 Z_3 + \\
&\beta_{423} X_4 Z_2 Z_3 + \\
&\beta_{443} X_4^2 Z_3 + \beta_{4423} X_4^2 Z_2 Z_3
\end{aligned} \tag{3}$$

This PLS2 model was fitted with the predictor variables in the reduced matrix $\mathbf{X_{red}}$ (27 × 9) because with these nine variables, the model had better global prediction ability and explained variance for each individual response. This is precisely the model inverted in the next section to obtain the CMP that predict the characteristics of the chromatograms closest to the ATP.

The characteristics of the final PLS2 model are in Table S1, in the supplementary material. The seven latent variables selected explain 99.92% of the variance in $\mathbf{X_{red}}$ and 97.70% of the total variance in $\mathbf{Y}$.

Furthermore, to evaluate a possible overfitting of the seven latent variables model, a permutation test was also done. Fifty random permutations of each response were performed and the RMSEC (Root Mean Squared Error in Calibration) and the RMSECV (Root Mean Squared Error in Cross Validation) were calculated for each permutation, and used to compute the probability that the prediction with the original

undisturbed PLS2 model could have been obtained at random. In this case, the probability of model insignificance versus permuted samples for the model with seven latent variables was less than $5\ 10^{-3}$ in all nine responses and with the three hypothesis tests conducted (sign test, rank test and rand t-test). At 5% significance level, probabilities less than 0.05 indicate that the chosen model is significant.

Table 2 contains the individual determination coefficients in both fitting and prediction. They are high for all the responses (greater than 0.95 and 0.87 in fitting and prediction, respectively) meaning that a large variance was explained in all responses (the resolutions between contiguous peaks, as well as for the initial and final time of the chromatogram). Furthermore, differences less than 0.08 between R^2 and R^2_{cv} for every response indicate that the model was highly predictive.

4.3. Inversion of the PLS2 model and selection of CMP

Once the characteristics of the desired (or target) chromatogram were set, in the form of the Analytical Target Profile, ATP, it was necessary to invert the fitted PLS2 model to find the composition of the mixture and the flow rate of the mobile phase to conduct the experiments, so that the target chromatogram is obtained.

The ATP specifies that each of the seven resolutions between consecutive peaks, $R_{i,i+1}$ ($i = 1, …, 7$) should be 2 with the shortest possible time of analysis. That means that the optimization problem is posed with different type of conditions for the objectives, a kind of 'mixed' optimization problem. In this case, seven of the responses have a target value, there is not condition for the initial time, and the remaining response (the final time) has to be just minimized in the experimental domain with no target value.

Besides, the experimental domain is the restricted domain used to fit the prediction model, where the variables Z_i are linearly dependent $\left(\sum_{i=1}^{3} Z_i = 1\right)$, and further Z_2 must be less than or equal to Z_3.

There is still another consideration to take into account: the experimental domain is in four dimensions, but there are nine predictor variables for the PLS2 model. Therefore, for any four-dimensional vector in this domain, the extended nine-dimensional vector should be computed to apply the model, as long as it belongs to the established PLSbox.

In short, a population of four-dimensional points belonging to both the experimental domain and the PLSbox evolved by maintaining the best possible value in at least one of the pursued objectives, while also preserving the diversity among solutions. In other words, the population evolved towards the Pareto optimal front of all the responses at hand. After a predefined number of generations, the Pareto front was extracted from the final population.

Several runs with different population size (from 100 to 300) and number of generations (from 500 to 1500) were performed. In none of them, R_{45} achieved more than resolution 1.75. All final populations were merged together, extracting the Pareto front, which is depicted in Fig. 3 in the form of a parallel coordinates plot.

In the graph, a vertical line represents each coordinate and the height in this line is the value of the corresponding coordinate. Broken lines join the values of each vector to follow the corresponding solution. To improve visualization, all the values have been range-scaled. The original ranges appear at the bottom and top of the corresponding coordinate (vertical line).

Furthermore, to be able to interpret the influence of the CMP on the predicted characteristics (ATP-related) of the chromatogram, they are
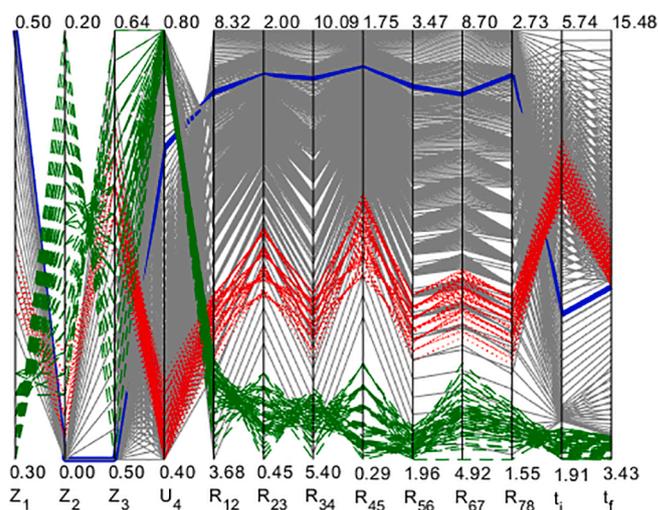

Fig. 3. Parallel coordinates plot of the Pareto front obtained when inverting the PLS2 model. Gray and colored lines separate the binary ($Z_2 = 0$) and ternary mobile phases, respectively. The blue line highlights the selected CMP along with the expected characteristics of the chromatogram.

depicted together in Fig. 3. The first four lines (coordinates) correspond to the proportion of Z_1 (water), Z_2 (methanol), Z_3 (acetonitrile), and X_4 (the mobile phase flow rate in mL min^{-1}), joint to the nine predicted characteristics, first the resolution between consecutive peaks, then the initial time and the last coordinate is the final time (i.e., undoing the transformation when defining Y_9), both in minutes.

Looking at the ranges in the solutions found in the Pareto front, the process variable (X_4, the flow rate) can take any value in its defining interval, just like the proportion of water, which also covers its whole allowable range (from 0.3 to 0.5), whereas the proportion of methanol scarcely achieves 0.2. The Pareto-optimal values attainable for the responses, last nine coordinates, show the difficulty of obtaining large values for R_{23} and, above all, R_{45}. The initial time varies between 1.91 and 5.74 min while the final (total) time goes from 3.43 up to just over 15 min.

More interesting is the effect of the CMP on the predicted responses. The grey lines identify the solutions with no methanol ($Z_2 = 0$) in binary mobile phases. The ternary mobile phases are marked with different colors.

In green discontinuous lines, a high flow rate (near 0.8 mL min^{-1}) for ternary mobile phases with little more than 30% of water, are clearly linked to the smallest values on all nine responses, which means the shortest runs to obtain the chromatogram (up to 4 min) but also the smallest values for resolution. In particular, R_{23} and R_{45} are less than one, with R_{78} and R_{56} barely reaching the target value 2, representing thus unacceptable solutions.

In red, still ternary mobile phases, now with up to 5% of methanol and more water (near 40%), linked to low levels of flow rate (between 0.4 and 0.48 mL min^{-1}), need longer time to finish the chromatogram (between 8 and 9.7 min), but all the resolutions increased. Although none of them is now less than one, still R_{23} and R_{45} are far from two.

With the binary mobile phases in grey, the greatest resolutions are expected, mostly with the longest runs. The most identifiable behavior is that increasing the flow rate, when it is at low levels (near 0.4 mL min^{-1}), increases all the responses up to their maximum achievable

Table 2

Coefficient of determination in fitting and prediction (estimated by cross-validation, CV) for the nine responses fitted with the PLS2 model.

	R_{12}	R_{23}	R_{34}	R_{45}	R_{56}	R_{67}	R_{78}	t_i	$t_f^{1/1.5}$
R^2	0.95	0.98	0.96	0.99	0.99	0.99	0.99	0.97	0.97
R^2_{cv}	0.87	0.95	0.91	0.98	0.98	0.99	0.98	0.92	0.90

values, so the worst solutions concerning the final time t_f. Like with ternary mobile phases, flow rate near 0.8 mL min^{-1} shortens the final time worsening the resolution, with unacceptable values for R_{23} and R_{45}. With some intermediate values of flow rate (between 0.6 and 0.7 mL min^{-1}) and more water, the resolutions are near their best values but the final time can be reduced.

Finally, the blue line in Fig. 3 highlights the CMP chosen to conduct the determination of the triazines: a binary mixture with equal proportion of water and acetonitrile and 0.68 mL min^{-1} of flow rate. With these conditions, the worst expected resolution is 1.6 for R_{45} and the final time is less than 9 min. (Fig. 2c) depicts one chromatogram obtained with the selected CMP.

4.4. Experimental verification of the selected CMP

Ten determinations of a mixture of 300 µg L^{-1} of each triazine were performed with the control method parameters found in the inversion of the PLS2 prediction model. These ten chromatograms were conducted in a single day.

Table 3 shows the mean values and their 95% confidence intervals for each of the nine characteristics in the ATP. The first block of rows corresponds to the experimental results obtained with the 10 replicates, whereas the second block contains the values computed with the predicted responses. For all the characteristics, it is seen that the intervals with the chromatographic data are included in the confidence intervals computed with the predictions of the PLS2 model.

A bibliographic revision about analytical determination of triazines by means of HPLC-DAD, limited to years 2010 to 2020 reveals 17 papers in which at least two of the eight triazines of the present work were determined. Those works are summarized in Table S2 of the supplementary material together with the retention time for each triazine, in the last 8 columns, though not all the triazines considered here were determined in all the referenced works. For comparative purposes, the characteristics in row 18 are the ones obtained in the present work with the validated procedure.

In most of the cases, 9 out of 17 published works, the mobile phase was programmed in gradient mode, mostly binary mixtures of acetonitrile and water. The eight remaining works usually used binary mixtures of methanol and water.

In relation to the ATP-related values, the method proposed here shows a clear advantage compared to the other papers. The elution time of the last appearing triazine in our case, less than 9 min, is much less than any of those found in the references shown in Table S2, which vary from 20.7 to 44.0 min.

4.5. Validation of the experimental procedure

4.5.1. Calibration lines, linear range and relative errors

Weighted least squares linear regression is the usual procedure to restore homoscedasticity when the variance of the response lacks it [24]. In analytical calibration, to counteract the greater influence of larger concentrations, different weights have been suggested ($1/y^2, 1/y, 1/y^{0.5}, 1/x^2, 1/x, 1/x^{0.5}$ being x the concentration and y the

signal) aiming at selecting the one that best fit the data [25–28]. Given the large range of the calibration standards used in the present work, a least squares linear regression was used with the response weighted by the inverse of the square of the experimental signal, $1/y^2$. Standardized residuals greater than three, in absolute value, were removed when computing the corresponding calibration line.

Rows 1 to 5 in Table 4 contain slope, intercept and standard error of estimation of the fitted models, for every triazine, further to the correlation coefficient and p-values related to the significance test with null hypothesis H_0: The model does not explain the variance of the response. All models were significant at 5% significance level because all p-values were less than 0.05. The next row in Table 4 contains the concentration of the calibration standards identified as outliers when calibrating. Therefore, the linear range in row 7 was not the same for all triazines since it depended on the outliers identified during fitting. It went from 3 to 300 µg L^{-1} for SZ, 6 to 300 µg L^{-1} for ST and AZ, 3 to 276 µg L^{-1} for AT, PZ, TZ and PT, and 6 to 276 µg L^{-1} for TT.

When comparing the results with the weighted regression as against the usual (unweighted) regression, the mean of the absolute value of the relative errors (MARE) for the $n = 8$ test samples were always less with the former than with the latter, except for AZ and TZ, as can be seen in rows 8 and 9 in Table 4. In any case, all of them are less than 2%.

4.5.2. Accuracy lines

Slope, intercept and standard error of estimation of the accuracy regression lines (predicted versus true concentration [29]) are in the rows 10, 11 and 12 in Table 4 for every triazine. Now, the goal of the regression line is not predicting but checking accuracy by comparing the predicted and true concentrations. Therefore, the usual least squares (unweighted) criterion was used. Then, 95% joint confidence regions were computed for (intercept, slope) and depicted in Fig. 4, where it is seen how all the ellipses contain point (0, 1), that is, null intercept and slope unity. The conclusion is that the method is unbiased for every triazine with neither constant nor proportional errors. Fig. 4 also shows that the two least precise determinations (ellipses with largest area) are for TT and SZ, which are the ones with longest and shortest elution time.

4.5.3. Decision limit and detection capability

Decision limit, CCα, and detection capability, CCβ, were computed to guarantee the probabilities of false negative (β) and false positive (α) which were set at 0.05 as stated in ISO-11843 [30]. For null concentration, CCα and CCβ were computed with the calibration lines fitted with weighted least squares, in a reduced range. The results (expressed in concentration in vial) are in the last two rows in Table 4, all of them are less than the least concentration in the calibration standards. Hence, to avoid extrapolation, the capability of detection is 3 µg L^{-1} for every triazine, except for ST and TZ, which is 6 µg L^{-1} since the first standard (3 µg L^{-1}) was an outlier in both cases. This means that the actual probability of false negative is negligible, less than 10^{-6} in the eight analyzed triazines.

These values, expressed in concentration in the surface waters taking into account the factors of preconcentration and recovery, are around 0.29 µg L^{-1} for ST and TT, and 0.12 µg L^{-1} for the rest of triazines; values

Table 3

95% confidence intervals for the mean of the nine ATP-related characteristics. Experimental data from HPLC-DAD analysis and theoretical data from the PLS2 predicted values.

	R_{12}	R_{23}	R_{34}	R_{45}	R_{56}	R_{67}	R_{78}	t_i	t_f
Experimental									
Lower limit	7.43	1.87	9.47	1.68	3.31	7.99	2.56	3.28	9.00
Mean	7.45	1.88	9.49	1.69	3.32	8.01	2.57	3.28	9.06
Upper limit	7.46	1.88	9.51	1.69	3.32	8.03	2.57	3.29	9.12
Predicted									
Lower limit	6.60	1.70	8.72	1.53	3.14	7.88	2.50	2.83	6.39
PLS2	7.68	1.85	9.59	1.63	3.28	8.16	2.62	3.29	8.49
Upper limit	8.76	1.99	10.46	1.72	3.42	8.44	2.75	3.74	10.78

Table 4

Performance criteria of the analytical method. Parameters of calibration lines (in the overall range and around the detection limit). Accuracy lines (s_{yx} is the standard error of estimation of the regression; MARE, is the mean of absolute values of relative errors). Decision limit, CCα and detection capability, CCβ for α and β set at 0.05.

	SZ	ST	AZ	AT	PZ	TZ	PT	TT
Calibration line (overall range)								
Intercept	0.162	−0.045	−0.078	0.022	0.009	0.024	0.054	−0.055
Slope	0.276	0.250	0.262	0.241	0.240	0.225	0.250	0.241
r	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
s_{yx}	0.006	0.092	0.010	0.006	0.006	0.005	0.009	0.009
p-value (sig. of regression)	<10⁻⁴	<10⁻⁴	<10⁻⁴	<10⁻⁴	<10⁻⁴	<10⁻⁴	<10⁻⁴	<10⁻⁴
Conc. of outliers (µg L⁻¹)	9, 21, 24	3	3	300	6, 300	12, 300	6, 300	3, 300
Linear range (µg L⁻¹)	[3, 300]	[6, 300]	[6, 300]	[3, 276]	[3, 276]	[3, 276]	[3, 276]	[6, 276]
MARE unweighted (%)	1.88	1.78	1.45	1.12	0.79	0.68	1.40	2.01
MARE weighted (%)	0.92	1.27	2.00	1.09	0.27	1.26	0.98	2.00
Accuracy line (overall range)								
Intercept	−0.084	−0.120	−0.002	−0.027	−0.079	0.055	−0.108	0.037
Slope	1.001	1.001	0.998	1.000	0.998	0.998	1.001	1.000
s_{yx}	0.583	0.471	0.525	0.305	0.295	0.341	0.379	0.697
Calibration line (detection limit)								
Intercept	0.188	−0.004	0.020	0.024	0.007	0.024	0.043	−0.034
Slope	0.270	0.246	0.256	0.240	0.240	0.225	0.251	0.238
s_{yx}	0.012	0.008	0.030	0.007	0.009	0.005	0.020	0.012
Conc. of outliers (µg L⁻¹)	3				6	12		3
CCα (µg L⁻¹)	0.112	0.090	0.291	0.072	0.098	0.063	0.192	0.144
CCβ (µg L⁻¹)	0.217	0.171	0.562	0.140	0.187	0.120	0.371	0.276
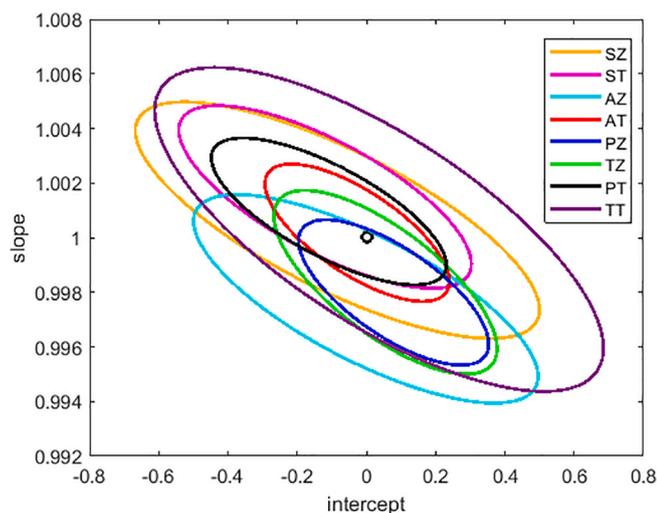
Fig. 4. 95% joint confidence region for intercept and slope of the accuracy lines. SZ in brown, ST in pink, AZ in cyan, AT in red, PZ in blue, TZ in green, PT in black and TT in purple. The circle 'o' is point (0, 1).

that are below the limits established in the regulations in force.

4.5.4. Recovery

Recovery was calculated from three MilliQ grade water samples fortified with 4 µg L⁻¹ of the eight triazines, which were pretreated as described in Section 2.4. The average recovery rate is expressed as the percentage of the amount of each triazine initially added in each sample found with the analytical procedure. The found recovery rates were 94, 50, 54, 46, 51, 44, 40, and 34% for SZ, ST, AZ, AT, PZ, TZ, PT, and TT respectively.

4.6. Determination of triazines in the samples of surface water

A new calibration line was fitted with 13 calibration standards covering the range from 0.34 to 24 µg L⁻¹. Four, out of the 15 samples analyzed, presented residues of SZ, ST, AT and TZ, but only the concentration of simazine in sample D was significantly different from zero, 21.7 µg L⁻¹ in vial (after applying the corresponding preconcentration and recovery factors, it is equal to 0.46 ± 0.01 µg L⁻¹ in the surface

water analyzed).

The corresponding concentration, once the correction factors are applied, falls well below the MAC-EQS of 4 µg L⁻¹ established for simazine in Directive 2013/39/EU for inland surface waters. The concentration of the remaining triazines in this sample, and of the eight triazines in the rest of water samples, it either was null or below the detection limit of the analytical procedure.

5. Conclusions

The work represents a novel application of AQbD to design a chromatographic procedure.

With preset characteristics for the desired chromatogram, the inversion of the PLS2 model fitted to the chromatographic process allows the selection of CMP to conduct the determination of eight triazines in very competitive conditions, compared to similar works published in the last ten years.

These selected conditions, namely a mobile phase made of a binary mixture with 50% water and 50% acetonitrile and a flow rate of 0.68 mL min⁻¹ give a procedure with not overlapping peaks for the eight triazines in less than 9 min. The method is also sensitive and more sustainable.

The found conditions were experimentally verified and the method validated before being used to measure samples of surface water from the basin of Arlanzón river, near Burgos in Spain. No triazines were found in the surface water samples analyzed, except for sample D, with a concentration of simazine well below its maximum allowable limit.

CRediT authorship contribution statement

M.C. Ortiz: Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing - original draft, Writing - review & editing. **L.A. Sarabia:** Conceptualization, Formal analysis, Methodology, Writing - original draft, Writing - review & editing. **A. Herrero:** Investigation, Methodology, Writing - review & editing. **C. Reguera:** Investigation, Methodology, Writing - review & editing. **S. Sanllorente:** Writing - review & editing. **M.M. Arce:** Writing - original draft, Writing - review & editing. **O. Valencia:** Formal analysis, Methodology, Writing - review & editing. **S. Ruiz:** Formal analysis, Methodology, Writing - review & editing. **M.S. Sánchez:** Formal analysis, Methodology, Software, 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.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.microc.2021.105971.

REFERENCES

[1] FDA-2003-D-0032: Guidance for Industry PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance, September 2004. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pat-framework-innovative-pharmaceutical-development-manufacturing-and-quality-assurance. Accessed on 19/7/2020.

[2] ICH guideline Q8 (R2) on pharmaceutical development (harmonised tripartite guideline), August 2009. https://www.ema.europa.eu/en/ich-q8-r2-pharmaceutical-development#current-effective-version-section. Accessed on 12/12/2020.

[3] A. Dispas, P. Lebrun, B. Andri, E. Rozet, P. Hubert, Robust method optimization strategy—A useful tool for method transfer: The case of SFC, J. Pharmaceut. Biomed. 88 (2014) 519–524, https://doi.org/10.1016/j.jpba.2013.09.030.

[4] P. Borman, M. Chatfield, P. Nethercote, D. Thompson, K. Truman, The application of quality by design to analytical methods, Pharm. Technol. 31 (2007) 142–152.

[5] S. Orlandini, S. Pinzauti, S. Furlanetto, Application of quality by design to the development of analytical separation methods, Anal. Bioanal. Chem. 405 (2013) 443–450, https://doi.org/10.1007/s00216-012-6302-2.

[6] EMA-FDA pilot program for parallel assessment of Quality-by-Design applications: lessons learnt and Q&A resulting from the first parallel assessment, August 2013. https://www.ema.europa.eu/en/human-regulatory/research-development/quality-design#parallel-assessment-with-the-united-states-section. Accessed on 16/08/2020.

[7] C.M. Jaeckle, J.F. MacGregor, Industrial applications of product design through the inversion of latent variable models, Chemom. Intell. Lab. Syst. 50 (2000) 199–210, https://doi.org/10.1016/S0169-7439(99)00058-1.

[8] S. Ruiz, L.A. Sarabia, M.C. Ortiz, M.S. Sánchez, Residual spaces in latent variables model inversion and their impact in the design space for given quality characteristics, Chemom. Intell. Lab. Syst. 203 (2020) 104040, https://doi.org/10.1016/j.chemolab.2020.104040.

[9] Commission Regulation (EU) 2016/266 of 7 December 2015 amending, for the purpose of its adaptation to technical progress, Regulation (EC) No 440/2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), Off. J. Eur. Union L 54 (2016) 1–446.

[10] A.M. Fan, Triazines, in Wexler P, Encyclopedia of Toxicology (Third Edition) Academic Press, 2014, pp. 810-815. https://doi.org/10.1016/B978-0-12-386454-3.01066-6.

[11] D. Želježić, S. Žunec, M. Bjeliš, V. Benković, M. Mladinić, B.L. Tariba, I. Pavičić, A.M.M. Čermak, V. Kašuba, M. Milić, A. Pizent, A.L. Vrdoljak, N. Kopjar, Effects of the chloro-s-triazine herbicide terbuthylazine on DNA integrity in human and mouse cells, Environ. Sci. Pollut. R. 25 (2018) 19065–19081, https://doi.org/10.1007/s11356-018-2046-7.

[12] 2004/248/EC: Commission Decision of 10 March 2004 concerning the non-inclusion of atrazine in Annex I to Council Directive 91/414/EEC and the withdrawal of authorisations for plant protection products containing this active substance. Off. J. Eur. Union L 78 (2004) 53–55.

[13] Real Decreto 817/2015, de 11 de septiembre, por el que se establecen los criterios de seguimiento y evaluación del estado de las aguas superficiales y las normas de calidad ambiental, Spanish BOE n° 219, 12 October 2015.

[14] 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.

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

[16] STATGRAPHICS Centurion 18 Version 18.1.12, Statpoint Technologies, Inc., Herndon, VA, USA, 2020.

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

[18] 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.

[19] Google Earth Pro version 7.3.3.7786 (64 bit), Google LLC, Mountain View, CA, USA, 2020.

[20] M. Ottavian, E. Tomba, M. Barolo, Advanced Process Decision Making Using Multivariate Latent Variable Methods, in: M. Ierapetritou, R. Ramachandran (Eds.), Process Simulation and Data Modeling in Solid Oral Drug Development and Manufacture. Methods in Pharmacology and Toxicology, Humana Press, New York, NY, 2016, pp. 159–189, https://doi.org/10.1007/978-1-4939-2996-2_6.

[21] S. Ruiz, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A computational approach to partial least squares model inversion in the framework of the process analytical technology and quality by design initiatives, Chemom. Intell. Lab. Syst. 182 (2018) 70–78, https://doi.org/10.1016/j.chemolab.2018.08.014.

[22] Mixtures Designs, Chapter 25 in Handbook of Chemometrics and Qualimetrics: Part A, D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke (Eds.), Volume 20 of Data Handling in Science and Technology, Elsevier, 1998, pp. 739-769. https://doi.org/10.1016/S0922-3487(97)80055-X.

[23] I.G. Chong, C.H. Jun, Performance of some variable selection methods when multicollinearity is present, Chemom. Intell. Lab. Syst. 78 (2005) 103–112, https://doi.org/10.1016/j.chemolab.2004.12.011.

[24] M.C. Ortiz, M.S. Sánchez, L.A. Sarabia, Quality of Analytical Measurements: Univariate Regression. In Brown SD, Tauler R, Walczak B, (Eds.) Comprehensive Chemometrics. Chemical and Biochemical Data Analysis, 2nd ed. Amsterdam, Elsevier, 2020, pp. 71-105. https://doi.org/10.1016/B978-0-12-409547-2.14869-3.

[25] A.M. Almeida, M.M. Castel-Branco, A.C. Falcão, Linear regression for calibration lines revisited: weighting schemes for bioanalytical methods, J. Chromatogr. B 774 (2002) 215–222, https://doi.org/10.1016/S1570-0232(02)00244-1.

[26] E. Alladio, E. Amante, C. Bozzolino, F. Seganti, A. Salomone, M. Vicenti, B. Desharnais, Effective validation of chromatographic analytical methods: The illustrative case of androgenic steroids, Talanta 215 (2020) 120867, https://doi.org/10.1016/j.talanta.2020.120867.

[27] J.M. Sánchez, Linear calibrations in chromatography: The incorrect use of ordinary least squares for determinations at low levels, and the need to redefine the limit of quantification with this regression model, J. Sep. Sci. 43 (2020) 2708–2717, https://doi.org/10.1002/jssc.202000094.

[28] H.O. Gomes, R.S. Cardoso, J.G.M. da Costa, V.P.A. da Silva, C.A. Nobre, R.N.P. Texeira, R.F. do Nascimento, Statistical evaluation of analytical curves for quantification of pesticides in bananas, Food Chem. 345 (2021), 128768, https://doi.org/10.1016/j.foodchem.2020.128768.

[29] J. Mandel, F.J. Linning, Study of accuracy in chemical analysis using linear calibration curves, Anal. Chem. 29 (5) (1957) 743–749, https://doi.org/10.1021/ac60125a002.

[30] ISO 11843, Capability of detection, Part 1: Terms and definitions and Part 2: methodology in the linear calibration case, International Organization for Standardization, 2000.