Use of a Plackett–Burman Design with Multivariate Calibration for the Analysis of Polycyclic Aromatic Hydrocarbons in Micellar Media by Synchronous Fluorescence

ARTICLE *in* ANALYTICAL CHEMISTRY · MARCH 2002

Impact Factor: 5.64 · DOI: 10.1021/ac010756w · Source: PubMed

6 AUTHORS, INCLUDING:

Olivier Devos
Université des Sciences et Technologies de Li…
39 PUBLICATIONS **266** CITATIONS
SEE PROFILE

Abdel-Ilah Saber
Université Savoie Mont Blanc
11 PUBLICATIONS **102** CITATIONS
SEE PROFILE

Emmanuel Naffrechoux
Université Savoie Mont Blanc
71 PUBLICATIONS **1,021** CITATIONS
SEE PROFILE

Use of a Plackett−Burman Design with Multivariate Calibration for the Analysis of Polycyclic Aromatic Hydrocarbons in Micellar Media by Synchronous Fluorescence

Olivier Devos,* Bernard Fanget, Abdel-Ilah Saber, Laurent Paturel, Emmanuel Naffrechoux, and Jean Jarosz

Laboratoire de Chimie Moléculaire et Environnement, Ecole Supérieure d'Ingénieurs de Chambéry, Université de Savoie, 73376 Le Bourget du Lac Cedex, France

This work describes the use of a multilevel Plackett−Burman design (PB) for the creation of a calibration set for partial least square regression (PLS). The PB calibration set was compared to a collinear analogue by testing these two PLS models for the analysis of six polycyclic aromatic hydrocarbons (PAHs). These compounds were analyzed in micellar media by synchronous fluorescence after determination of the experimental conditions (choice of surfactant, analytical conditions such as $\Delta\lambda$, step, and scan range). The external validation shows that the collinear set is inappropriate to quantify PAH in real samples, but the PB calibration set affords optimal results.

The multivariate chemometric techniques are powerful tools of broad use in analytical chemistry. The classical methods, such as the principal component regression (PCR), multilinear regression (MLR), or partial least squares (PLS), have been applied to obtain selective information from nonselective data.

The more sound method,[1] PLS, is mainly used in spectroscopy in the case of overlapping standard spectral bands. This model has been applied to various analyses, such as the determination of nitrate in groundwater by UV−vis absorbance,[2] kerosene properties by infrared spectra,[3] or ash content in sugar samples by fluorescence spectroscopy.[4] PLS has also been used for the synchronous fluorescence determination in Brij-35 media of 10 PAHs[5] included in the list of the 16 PAHs classified as priority pollutants by the Environmental Protection Agency (EPA).

The prediction ability of the PLS model is relevant to the concentration standards. For the analysis of two components, the calibration set may be constructed with the help of a graphic with N1 and N2 levels. Each intersection in the graphic corresponds to a mixture of standards. Therefore, the total number of standards required is given by N1 × N2. In the case of a three component analysis, Eriksson et al. used a mixture design with only two independent concentrations.[6]

In the literature, most reports using multivariate methods with a number of compounds higher than three do not provide indications about the concentrations in the calibration set. Guiteras et al. have chosen 70 standards to quantify 10 PAHs without indication of concentrations;[5] thus, so far, no reliable calibration method for PLS models has been developed.

We propose the use of a Plackett−Burman (PB) design[7] to create such a calibration set for PLS, which was tested for the analysis of six PAHs in comparaison with a collinearity calibration set. The PAHs are referenced by the European Economic Community (EEC) as quality indicators for drinking water.[8] The analyses were effected by synchronous fluorescence in Genapol X-80 media.

The excitation and emission wavelengths (λ_{ex} and λ_{em}) for synchronous fluorescence measurements are scanned together, with a constant wavelength interval ($\Delta\lambda$) between λ_{ex} and λ_{em}. A simplified spectral profile is obtained, together with a reduction of spectral overlap, as compared with classical fluorescence methods.[9]

PAHs' fluorescence spectra in micellar media are characterized by an increased intensity, as compared to pure water.[10,11] This increase can be attributed to the higher quantum yield as a result of the lower rate of the radiationless relaxation processes. Indeed, different factors can affect the latter: (1) alteration of the micropolarity; (2) restricted motion; and (3) effective shielding of the excited singlet state against quenchers, such as O_2.

* Corresponding author. Tel: +33-479-75-88-33. Fax: +33-479-75-88-85. E-mail: odevo@univ-savoie.fr.

(1) Thomas, E. V.; Haaland, D. M. *Anal. Chem.* **1990**, *62*, 1091−1099.
(2) Dalhén, J.; Karlson, S.; Bäckström, M.; Hagber, J.; Pettersson, H. *Chemosphere* **2000**, *40*, 71−78.
(3) Guarrigues, S.; Andrade, J. M.; De La Guardia, M.; Prada, D. *Anal. Chim. Acta* **1995**, *317*, 95−105.
(4) Norgaard, L. *Talanta* **1995**, *42*, 1305−1324.
(5) Guiteras, J.; Beltràn, J. J.; Ferrer, R. *Anal. Chim. Acta* **1998**, *361*, 233−240.
(6) Eriksson, L.; Johansson, E.; Wikström, C. *Chemometr. Intell. Lab. Syst.* **1998**, *43*, 1−24.
(7) Plackett, R. L.; Burman, J. P. *Biometrika* **1946**, *33*, 305−325.
(8) Directive 80/778/EEC, Council of European Communities, *Off. J. Eur. Commun*; L229, 1980, 11.
(9) Vo-Dihn, T. *Modern Fluorescence Spectroscopy*; Plenum: New York, 1981; pp 167−192.
(10) Singh, H.; Hinze, W. L. *Anal. Lett.* **1982**, *15*, 221−243.
(11) Hinze, W. L.; Singh, H. N.; Baba, Y.; Harvey, N. G. *Trends Anal. Chem.* **1984**, *3*, 8, 193−199.

10.1021/ac010756w CCC: $22.00
Published on Web 00/00/0000

THEORY

PLS Calibration. PLS is a multivariate calibration procedure based on a factor analysis method,[12] by which spectra (X) and concentrations (Y) are transformed simultaneously into a structure part (TQ' and TP') and a noise part (E and F),

$$X = TQ' + EY = TP' + F$$

where Q' and P' are the factors, and T is the score (eigenvector loading).

The number of scores is linked to the complexity of the spectra and can be determined by using a cross validation (internal validation) method leaving out one sample each time.[13] The predicted concentrations are compared with the known concentrations and the prediction error sum of square (PRESS) is calculated as follows,

$$\mathrm{PRESS} = \sum_{i=1}^{n} (\hat{c}_i - c_i)^2$$

where n is the number of calibration samples, $\hat{c}_i$ is the estimated concentration, and c_i is the reference concentration for the ith sample.

The numbers of factors were chosen in order to minimize the PRESS. The smaller the PRESS value, the better the model predicts the concentrations of the constituents in the samples.

To quantify the prediction ability of the models, the relative root mean square difference was used,

$$\mathrm{RRMSD\%} = \frac{100}{\bar{c}} \sqrt{\frac{\mathrm{PRESS}}{n}}$$

where n is the number of sample, and $\bar{c}$, the mean value of the concentration of the component considered. RRMSD% can be obtained by internal validation (cross validation) or external validation.

Calibration Conditions. Classical calibration methods measure a physical value proportional to a component concentration.[14] The proportionality constant is empirically determined from known concentration standards. The standards must match the following criteria: (1) no interference, (2) multilevel concentrations in standard solutions, and (3) reproducibility and linearity between concentration and measured physical values in the considered range. Samples with concentrations different from those of the calibration set are used for validation. When the observed signal includes several components (x, number of components), the PLS model is applied. Last, the minimal number of experiments must be fixed to $x + 1$ mixtures with no collinearity due to the decomposition of spectra and concentrations by PLS.

To conclude with these points, only the use of a multilevel experiment design to create the calibration set can remove the collinearity constraint and respect classical calibration conditions.

Plackett−Burman Design-Calibration Set. The experimental design fixes the number of experiments (N) to be realized for a given factor level. Generally, the experimental design is constructed for two factor levels (L), which takes only two extreme values, −1 and + 1. To obtain information for more than two factor levels, a multilevel, that is, full factorial design is necessary. This implies that all the combinations of the levels for each factor must be present, requiring a large number of experiments. Therefore, to reduce this number, only a fraction of the full-factorial-design-like PB design[7] is used.

Since this design is orthogonal (e.g., no covariance between experimental factors) and centered,[7] it can be applied for PLS calibration where the factor levels (L) and the number of experiments (N) correspond, respectively, to the number of concentration levels and of mixtures.

The number of components studied (n_c) and concentration levels condition the choice of the PB table. Because n_c is fixed, the efficiency (E%)[15] calculation of the experimental design naturally optimizes L,

$$E\% = 100 \times \frac{\mathrm{dfn}}{N} = 100 \times \frac{1 + n_c(L-1)}{kL^2}$$

where N is a multiple of L^2 to respect the orthogonality, dfn (degree of freedom number) is the minimal number of independent experiments, and k is a positive integer determined by

$$k \geq \frac{1 + n_c(L-1)}{L^2}$$

A few tables have been given in the original paper,[7] but nowadays any type of tables can be computed.

They are generated by using the first column of the design. The complete design is obtained by cyclical permutation of $[(N-1)/(L-1)] - 1$ times and by the addition of a row of the lower level. The maximum number of experimental factors (components) to obtain an orthogonal design is equal to $(N-1)/(L-1)$.

The correspondence between concentrations and centered coded data results from the following transformation equation

$$C_{i,j} = C_{0,j} + X_{i,j}(C_{\max,j} - C_{\min,j})/(L-1)$$

where $C_{i,j}$ is the concentration of the constituent j for the mixture i; $C_{0,j}$ is the midrange concentration; $C_{\max,j}$ is the maximum concentration value of constituent j; $C_{\min,j}$ is the minimum concentration value of constituent j; L is the number of concentration levels, and $X_{i,j}$ is the centered coded data of constituent j for mixture i.

This rigorous construction of the calibration set minimizes the number of experiments with a maximum of efficiency. Because this design is orthogonal, the set number of each single concentration level of each component is, therefore, reproduced with an independent analysis order.

EXPERIMENTAL SECTION

Reagents. Polyaromatic compounds were purchased from

(12) Geladi, P.; Kowalski, B. R. *Anal. Chim. Acta* **1986**, *185*, 1−17.
(13) Eastment, H. T.; Krzanowski, W. J. *Technometrics* **1982**, *24*, 73−77.
(14) Skoog, D. A.; West, D. M.; Holler, F. J. *Chimie Analytique*; De Boeck & Larcier: Paris, 1997; 601−610.
(15) Pillet, M. *Les Plans d'Expérience par la Méthode TAGUCHI*; Les Editions d'Organisation: Paris, 1997.

Table 1. Spectrofluorometric Characteristics of BaP at 1 μg L^{-1} in Micellar Media[a]

medium	λ_{ex} (nm)	λ_{em}(nm)	I_m/I_a[b]
aqueous	299	408	-
Brij 35	300	405	3.7
Tergitol TMN-6	299	405	4.6
Tergitol 15-S-7	299	405	8.2
Genapol X-80	300	405	8.4

[a] All surfactant at 0.1% (w/w) in water. [b] Ratio of fluorescence intensities in micellar (I_m) and aqueous (I_a) media.

Table 2. Analytical Parameters of Synchronous Fluorescence Spectra of the Six PAHs with 0.1% of Genapol X-80

PAH	λ_{ex} (nm)[a]	λ_{em} (nm)[a]	$\Delta\lambda$ (nm)	emission acquisition range (nm)
BaP	288, 300, 369	405,430	105	380−480
BkF	300, 310, 360	409, 436	99	380−480
BbF	295, 304, 376	431	127	400−500
BghiP	292, 303, 362	408, 417, 421	105	380−480
IP	306, 317, 407	471, 503	165	440−540
F	290, 384	440	150	400−500

[a] The more intense bands are underlined for each PAH.

Table 3. Concentration Data for the Mixtures in the PB Calibration Set

	μg L^{-1}					
coded data[a]	BaP	BkF	BbF	BghiP	IP	F
−2	0.10	0.04	0.20	0.4	2.0	2.0
−2	0.10	0.31	0.65	3.1	15.5	11.0
2	1.00	0.04	1.55	1.3	15.5	15.5
−1	0.33	0.40	0.20	3.1	6.5	15.5
−1	0.33	0.13	2.00	0.4	15.5	6.5
0	0.55	0.13	0.65	4.0	2.0	15.5
−1	0.33	0.22	0.65	1.3	20.0	2.0
−2	0.10	0.13	1.10	1.3	6.5	20.0
1	0.78	0.04	0.65	2.2	6.5	6.5
0	0.55	0.31	0.20	1.3	11.0	6.5
0	0.55	0.22	1.55	0.4	6.5	11.0
2	1.00	0.22	1.10	3.1	2.0	6.5
0	0.55	0.40	1.10	2.2	15.5	2.0
−2	0.10	0.22	2.00	2.2	11.0	15.5
−1	0.33	0.04	1.10	4.0	11.0	11.0
2	1.00	0.13	0.20	2.2	20.0	11.0
2	1.00	0.40	0.65	0.4	11.0	20.0
1	0.78	0.40	2.00	1.3	2.0	11.0
2	1.00	0.31	2.00	4.0	6.5	2.0
−2	0.10	0.40	1.55	4.0	20.0	6.5
0	0.55	0.04	2.00	3.1	20.0	20.0
1	0.78	0.22	0.20	4.0	15.5	20.0
1	0.78	0.31	1.10	0.4	20.0	15.5
−1	0.33	0.31	1.55	2.2	2.0	20.0
1	0.78	0.13	1.55	3.1	11.0	2.0

[a] First column of the Plackett−Burman design.

Supelco, and methanol (analytical grade), from Merck. UHQ water was prepared from a Maxima system (Elga).

The surfactant Genapol X-80 was purchased from Fluka; Tergitol TMN-6 and Tergitol 15-S-7, from Sigma-Aldrich; and Brij 35 from Technicon.

Stock standard solutions of benzo(*a*)pyrene (BaP), benzo(*k*)-fluoranthene (BkF), benzo(*ghi*)perylene (BghiP), benzo(*b*)fluoranthene (BbF), fluoranthene (F) and indeno(1,2,3-*cd*)pyrene (IP) were prepared by dissolving the pure solid in methanol. Dilute solutions were prepared from stock solutions.

Safety Precautions. Several PAHs are suspected carcinogens and caution must be exercised with these compounds. All handling should be performed in a ventilated hood with latex gloves to avoid inhalation or skin contact.

Preparation of the PAH Mixture in Micellar Media. Each mixture was prepared in aqueous solution with UHQ water, 1% (v/v) of methanol, 0.1% (w/w) of surfactant, and an adequate concentration of PAH. All volumes were weighed to determine the actual concentrations. The surfactant concentration was ∼36 times the critical micellar concentration (CMC = 0.05 mmol L^{-1}) in order to increase the fluorescence intensities of PAHs.[10,11]

Instrumentation and Software. A homemade spectrofluorometer using a computerized system for spectral acquisition and data processing has been developed.[16] The spectrum of the sample was obtained using a 1000 W xenon lamp in a metal-coated cell, and the following instrumental parameters were used: 2-mm excitation slit widths and 1.5-mm emission slit widths for 4- and 0.6-nm band-pass, respectively. No correction factors for excitation intensity and detection response were applied to the signal. Furthermore, the PLS models are valid only on the same spectrofluorometer. The GRAMS-386 package with PLSplus/IQ application software (Galactic Industries Co., Salem, NH) was used for the processing of spectral data.

RESULTS AND DISCUSSION

Spectral Characteristics of PAH in Micellar Media. The fluorescence characteristics of BaP (λ_{ex}, λ_{em}, ratio of the fluorescence intensity in micellar-to-water media) were studied in the presence of four nonionic surfactants: Genapol X-80, Tergitol TMN-6, Tergitol 15-S-7, and Brij 35.

Table 1 shows that the excitation and emission maxima wavelengths of BaP at 1 μg L^{-1} are very similar in the different micellar media that produce an important increase in fluorescence intensity, ∼4−8 times that measured in aqueous solution. Genapol X-80, 0.1% (w/w) in water, producing the maximum enhancement, is used for calibration.

The optimum $\Delta\lambda$ observed for the best quantification of each PAH corresponds to the difference between the maximum intensity of the excitation and emission wavelengths. The results obtained for the six PAHs are summarized in Table 2. Last, scanning of the spectra is performed with five selected $\Delta\lambda$'s: 99, 105, 127, 150, and 165 nm, each being more adequate for a specific PAH. The emission data are recorded in a range of 100 nm with a step of 0.5 nm, and a calibration set using the five selected $\Delta\lambda$'s based on a PB design has been realized. The first results observed with $\Delta\lambda$ = 105 nm are quite similar to those obtained using the five optimal $\Delta\lambda$'s cited above (Table 4). Thus, $\Delta\lambda$ = 105 nm was chosen for the direct PAH quantification in order to simplify the analysis. Figure 1 shows the spectra of the six PAHs with $\Delta\lambda$ = 105 nm.

PLS Calibration. Two calibration sets have been carried out, the first one based on a PB design, and the second, on a

(16) Paturel, L.; Saber, A. I.; Fachinger, C.; Suptil, J.; Turnar, C. *Polycyclic Aromatic Compounds* **1999**, *13*, 151−164.

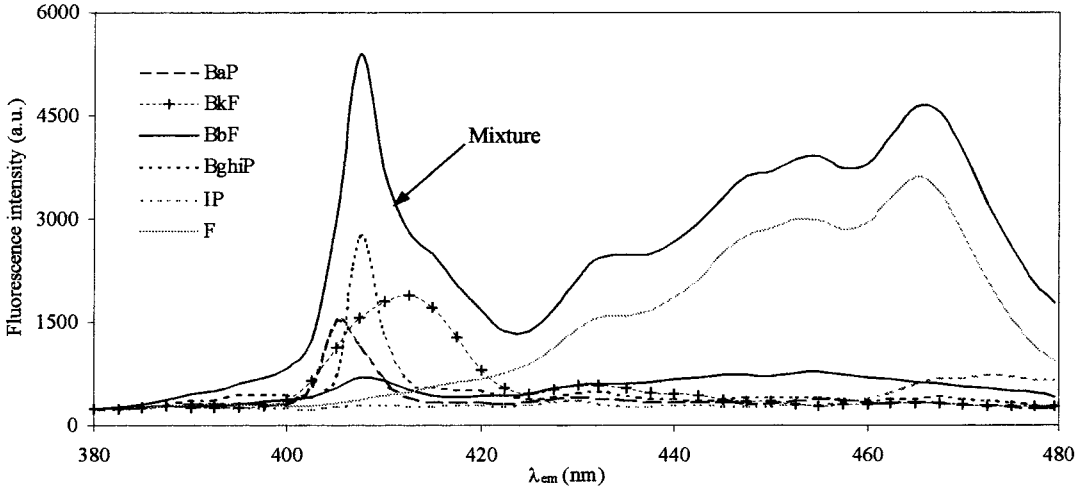

Figure 1. Synchronous spectra ($\Delta\lambda$ = 105 nm) of the six PAH in 0.1% Genapol X-80. Concentrations in μg L^{-1}: BaP, 0.83; BkF, 0.33; BbF, 1.66; BghiP, 1.66; IP, 16.29; and F, 16.75.

Table 4. Result of the PLS Calibration by Cross Validation with the Five $\Delta\lambda$ and $\Delta\lambda$ = 105 nm

	$\Delta\lambda$ (nm)	RRMSD%
BaP	105	6, 0
BkF	99	6
	105	6, 1
BbF	127	3, 4
	105	9, 8
BghIP	105	7
IP	165	6, 8
	105	8, 2
F	150	3, 8
	105	5, 2

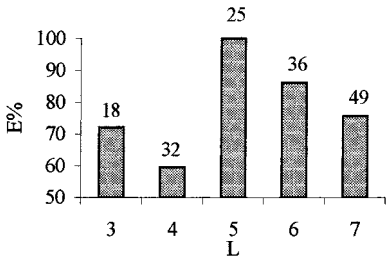

Figure 2. Efficiency (*E*%) of a PB design with six components (n_c = 6) and several levels (*L*).

collinearity set. The concentration ranges in μg L^{-1} used were for BaP, 0.1−1; for BkF, 0.04−4; for BbF, 0.2−2; for BghiP, 0.4−4; for IP, 2−20; and for F, 2−20.

Choice of the PB Design. To develop the most appropriate design for the analysis of the six PAHs (n_c = 6), the efficiency (*E*%) and the number of mixtures (*N*) have been calculated as a function of *L* (Figure 2). The five-concentration-level design with 25 mixtures led to a maximum efficiency. The PB design yielded the theoretical concentration values shown in Table 3.

The collinear set was realized with the same number of mixtures as the PB set. These 25 mixtures were obtained by dilution in UHQ water with Genapol X-80, from the same methanol stock solution as the six PAHs. These mixtures are uniformly distributed on the chosen concentration range.

For PLS calibration, cross validation is used to determine the optimum number of factors. On the other hand, testing mixtures different from those of the calibration set (external validation) is essential to validate the entire model. Figure 3 shows the average RRMSD% of the six PAHs obtained by cross (internal) and external validation. The collinear calibration matrix is totally inappropriate for quantification, despite a better internal validation than with the PB set. No significant difference between internal and external validation occurs with the latter, which proves its efficiency not only for the residual error but also for the quantification.

In comparison with the determination of 10 PAHs by synchronous fluorometry in Brij 35 media,[5] Guiteras et al. obtained nearly the same values of RRMSD% for each PAH between 6 and 11 for higher concentration range (0−20 μg L^{-1}) and a larger number of samples (70 calibration mixtures). Amador-Hernandez et al.[17] quantified six PAHs in polyoxyethylene 10 lauryl ether (POLE) media by linear variable angle fluorescence. They obtained similar values of RRMSD% (between 3 and 12%) from 34 calibration mixtures and a more selective analysis than by synchronous fluorescence.

The PB PLS models have also been tested using two samples of an artificial drinking water ($NaHCO_3$, 96 mg L^{-1}; $CaSO_4$, 60 mg L^{-1}; $MgSO_4$, 60 mg L^{-1}, and KCl, 4 mg L^{-1}). The concentrations of the spiked PAHs are different from the concentration mixtures of the calibration sets.

The results obtained for the two samples (Table 5) indicate a good agreement between the amounts of PAHs added and found.

Justification of PB Design. Standard mixtures cannot be perfectly obtained; therefore, the calibration set is not orthogonal. Thus, the multiple correlation criteria (R^2_m) between 0 and 1 define a collinearity rating of a matrix (R^2_m = 0 in the case of an orthogonal matrix). The progressive addition of collinear mixtures to an orthogonal calibration set increases R^2_m. The effect of R^2_m on RRMSD% is shown by the composite sets, made by gradual substitution of a defined number of mixtures from the PB set by the same number of mixtures from the collinear set, the mixtures being obviously randomly chosen and independent.

(17) Amador-Hernandez, J.; Cladera, A.; Estela, J. M.; Lopez-de-Alda, P. L.; Cerda, V. *The Analyst* **1998**, *123*, 2235−2241.

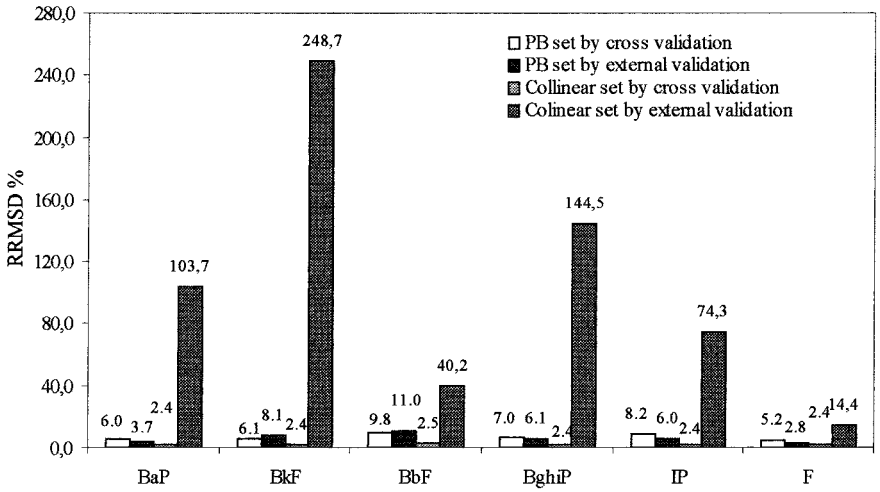

Figure 3. Comparison between PB calibration set and collinear calibration set with cross validation or external validation. The RRMSD% was determined using four independent samples for external validation and 25 samples for cross validation.

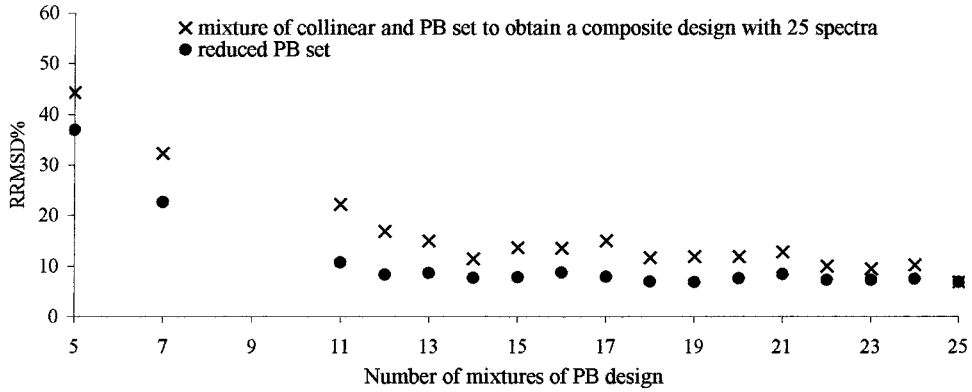

Figure 4. Evolution of RRMSD% for reduced PB and composite sets. The RRMSD% was determined by external validation (four independent samples).

Table 5. Concentrations of the PAHs in Artificial Natural Water Using the PB Calibration Set

	μg L^{-1}			
	sample 1		sample 2	
	found[a]	added	found[a]	added
BaP	0.52 ± 0.02	0.54	0.12 ± 0.02	0.13
BkF	0.16 ± 0.02	0.17	0.22 ± 0.02	0.24
BbF	0.71 ± 0.05	0.61	0.85 ± 0.05	0.79
BghiP	1.31 ± 0.10	1.24	0.28 ± 0.10	0.40
IP	1.80 ± 0.60	1.83	14.27 ± 0.60	13.67
F	4.30 ± 0.58	3.86	13.99 ± 0.58	14.11

[a] ± confidence interval for one measurement.

Figure 4 shows the evolution of average RRMSD% determined by external validation versus collinearity ratings of composite sets. To simplify the interpretation of the cited graph, the collinearity rating is replaced by the noncollinear mixtures number. Nevertheless, the replacement of noncollinear mixtures by collinear ones leads to a progressive inaccuracy of the concentration prediction. The impact of mixtures from PB design is shown by using reduced PB sets. The latter have been created by removing the collinear mixtures from the composite sets. They include a variable number of mixtures, and the reduced sets having less than five mixtures have been discarded, because the number of computing factors was too low (less than 3). R^2_m for composite sets is not only correlated but is also higher than those obtained for reduced PB sets (30% average). Although the number of mixtures is smaller in the case of reduced PB sets, we obtain a lower RRMSD% (35% average) than the one obtained for the corresponding composite set. This increase seems to be linked to that of R^2_m.

The lowest RRMSD% was obtained from the complete PB design using 25 mixtures. For the reduced design obtained with 12−25 mixtures, the difference between the predicted and the experimental values does not significantly increase (RRMSD% max of 8.5%). For a design of less than 12 mixtures, the maximum number of computing factors is too low to explain spectral variations, therefore leading to a certain increase of RRMSD%.

CONCLUSION

This paper shows that the adequation of PLS models is influenced by the construction of the calibration set. A conceptually simple procedure to calibrate a PLS multivariate analysis has been introduced. This procedure allows the standardization of a calibration procedure, regardless the number of compounds, with a reasonable number of mixtures. Once an efficient optimization on L, N, n_c is achieved, the complete PB design can be considered as totally defined on the chosen concentration ranges. In this study, we demonstrate by external validation that the PB design

offers results better than other types of set incorporating collinearity. The lowest prediction error is obtained through the complete PB design. Nevertheless, with an incomplete PB design, the prediction becomes less accurate but satisfactory. The number of computing factors is, moreover, sufficient. This incomplete design can be optimized before experimentation by choosing appropriate mixtures for the correlation criteria, which are linked to the accuracy of the result. These criteria can be applied to all types of calibration sets to estimate their predictive ability.

Received for review July 6, 2001. Accepted October 17, 2001.

AC010756W