Spectrophotometric Analysis of α- and β-Naphthol Mixtures Applying Partial Least-Squares Method

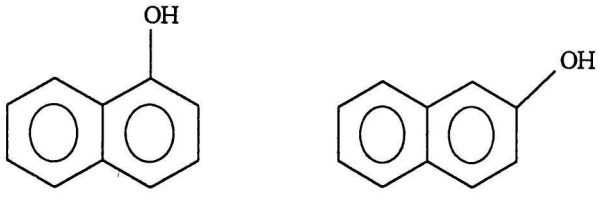
A. L. REVILLA, G. VARGAS, and J. HAVEL*

Department of Analytical Chemistry, Faculty of Science, Masaryk University, CZ-611 37 Brno, Czech Republic
e-mail: havel@chemi.muni.cz

Received 29 January 1998

Pharmaceutical and other chemical industries often require high purity of α-naphthol (αN) because it is used like a precursor for further synthesis. The determination of β-naphthol (βN) in mixture with αN by spectrophotometry applying partial least-squares procedure was developed. Using the partial least-squares calibration, a model based on calibration measurement was developed and, afterwards, used to predict the test set (samples). Normal or derivative spectra in λ-ranges *a*) 190—250 nm or *b*) 260—345 nm were used. The method is suitable for fast determination of βN in mixtures with αN with relative standard deviation around 2 % for βN (if the content is higher than 1.5—3.0 %). Both studied λ-ranges can be used for the determination, but derivative spectra are advantageous for the 190—250 nm region.

α- or β-naphthol are weak acids (pK_a 9.34 and 9.51) [1]. These compounds are structural isomers differing only in the hydroxy group position.



αN βN

They are used as precursors for azo dyes and the other mostly pharmaceutically relevant compound syntheses. In many of the processes involved, a high purity of the naphthols is required. For example, only 5 % content of βN in αN causes a colour instability of resulting azo dyes [2]. For the synthesis of medicaments, the purity of αN required is less than 0.5 % βN content. At present, the mixtures of αN and βN are analyzed mostly by separation techniques like chromatographic methods as HPLC [3], picochromatography [4], and more recently by capillary zone electrophoresis [5] with UV detector.

Both compounds have a high molar absorptivity in the UV range ($\varepsilon \approx 40\,000$ dm^3 mol^{-1} cm^{-1}) at 210 nm for βN and at 226 nm for αN, respectively. It offers a good possibility for sensitive spectrophotometric analysis as an alternative method without the necessity to use organic solvents. However, the absorption spectra are quite similar (Fig. 1) and the direct determination of traces of βN in αN reagent is very difficult by classical multicomponent spectrophotometric analysis.

Multivariate calibration methods have an increased importance in multicomponent analysis in analytical chemistry, especially those using the partial least-squares (PLS) method with decomposition into latent variables [6—8]. It has already been shown that the PLS method gives favourable results in comparison with multiple regression methods for spectrophotometric data [9]. Thus the use of this kind of approach opens the possibility to determine the traces of βN in the presence of αN excess directly by spectrophotometry. In this work, we have examined the application of multivariate calibration using the PLS method with the aim to find the possibility to evaluate the spectrophotometric data (from normal or derivative spectra) and thus to determine traces of βN in mixtures with αN as the main component.

EXPERIMENTAL

Naphthol samples were reagent grade, Merck (Darmstadt, Germany). Additionally, β-naphthol was recrystallized prior to use to assure the highest possible purity. The other reagents were anal. grade supplied by Lachema (Brno, Czech Republic). Thrice distilled water used was produced in a commercial apparatus from Heraeus (Hanau, Germany). Standard pH buffers were from the Institute of Serum and Vaccines (Prague, Czech Republic).

*The author to whom the correspondence should be addressed.

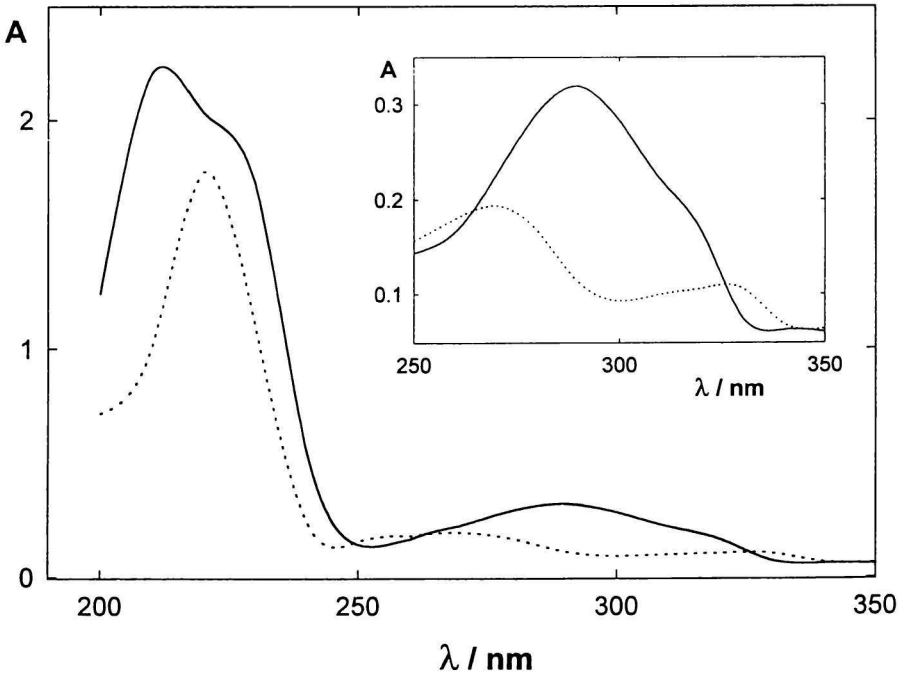


Fig. 1. Absorption spectra of the pure compounds: βN, $c = 0.190$ mmol dm^{-3} (——); αN, $c = 0.253$ mmol dm^{-3} (– – –).

UV2 Quartz Series spectrophotometer (ATI-UNICAM, Cambridge, UK) was applied while a cell with 1-cm path-length quartz was used for the absorbance measurements. The derivative spectra smoothing was applied (*high mode*) using a Vision 3.0 software of the spectrophotometer. Due to the property of the software used some data were lost (around 5 wavelengths). A pH-meter PHM 84 (Radiometer, Copenhagen, Denmark) with saturated calomel electrode and glass electrode was used for pH measurements. The absorption spectra were measured using 5 mM borate (pH 8.5) as buffer.

Solutions of βN and αN were prepared separately. The weighted amount of the substance was dissolved in 1 cm^3 of ethanol and then the solution was filled up to 50 cm^3 volume with water. Posterior dilutions were done with water.

Computational Approach and Programs

The PLS program written according to published algorithms [6, 7] was used for the calculations. The theory of PLS has already been described in many papers, and an overview can be found elsewhere [6—9], for this reason only a brief explanation about the application of PLS for spectrophotometric data is given here. The absorbance matrix $\mathbf{A}$ for N_s solutions with known concentrations (calibration matrix) of N_c components which absorb at the chosen spectral range, is given as

$$\mathbf{A} = \mathbf{CE} \tag{1}$$

where $\mathbf{C}$ is $N_s \times N_c$ concentration matrix and $\mathbf{E}$ is $N_c \times N_\lambda$ matrix of molar absorptivities, N_λ denotes the number of wavelengths. By using the PLS method the matrixes $\mathbf{A}$ and $\mathbf{C}$ are decomposed into latent variables

$$\mathbf{A} = \mathbf{F}_A\mathbf{L}_A^T + \mathbf{E}_A \tag{2}$$

$$\mathbf{C} = \mathbf{F}_C\mathbf{L}_C^T + \mathbf{E}_C \tag{3}$$

where $\mathbf{F}_A$ is the latent absorbance matrix ($N_c \times$ LV), LV is the number of abstract component, $\mathbf{L}_A$ is the loading A-matrix (LV $\times$ N_s), and $\mathbf{F}_C$ is the latent concentration matrix, while $\mathbf{E}_A$, $\mathbf{E}_C$, and $\mathbf{E}_D$ are error matrixes (of unexplained variability). Relating both equations (*2* and *3*) for a diagonal regression matrix $\mathbf{V}$, the equation $\mathbf{F}_C = \mathbf{F}_A\mathbf{V} + \mathbf{E}_D$ is obtained. The matrix $\mathbf{V}$ is then used in the prediction step to determine unknown sample concentration using the relation

$$c_0 = A_0(\mathbf{F}_C^T\mathbf{A})^T\mathbf{V}\mathbf{L}_C \tag{4}$$

where A_0 represents the experimentally obtained absorbance.

Experimental Design

A three-level factorial design was used for the multivariate calibration matrix. Only two components are present in the matrix, αN and βN, in this way a 3^2 design was employed. Table 1 shows the concentration levels for each wavelength range studied.

Table 1. Concentration Levels for the Calibration Set

Level	Code	Wavelength range 190—240 nm		245—365 nm	
		αN $c \cdot 10^6/(\text{mol dm}^{-3})$	βN $c \cdot 10^6/(\text{mol dm}^{-3})$	αN $c \cdot 10^5/(\text{mol dm}^{-3})$	βN $c \cdot 10^5/(\text{mol dm}^{-3})$
Low	−1	4.505	0.069	2.486	0.026
Average	0	13.514	0.206	12.43	0.255
High	1	22.523	0.344	24.86	2.549

c = concentration.

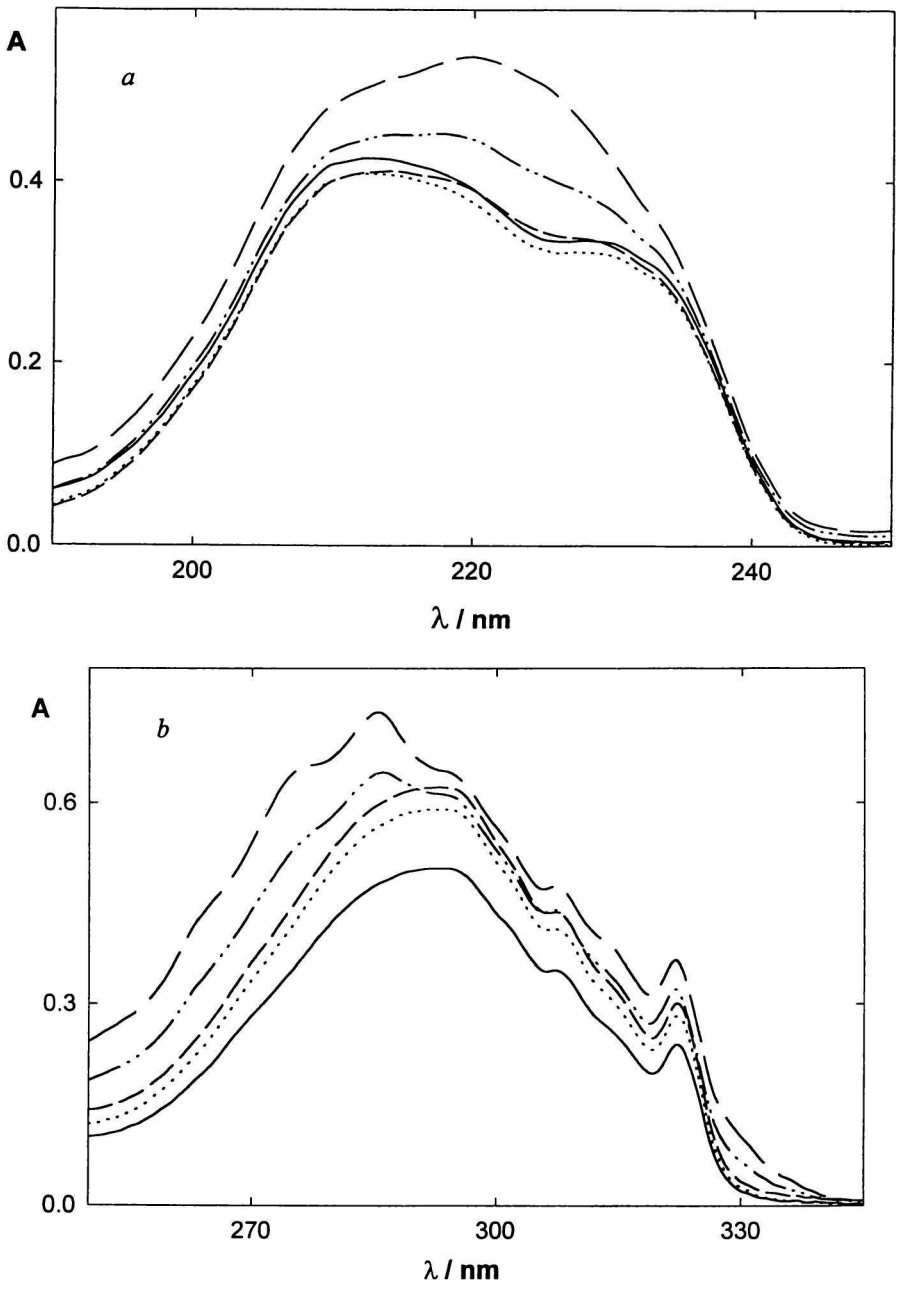


Fig. 2. *a*) Absorption spectra for different contents of βN in the wavelength range 190—250 nm at 0.010 mmol dm^{-3} αN concentration; βN content/%: 0 (——); 0.5 (·····); 2.4 (– – –); 9.1 (– ·–); 20 (— —). *b*) Absorption spectra for different contents of βN in the wavelength range 260—345 nm at 0.125 mmol dm^{-3} αN concentration; βN content/%: 0 (——); 0.6 (·····); 2 (– – –); 16.7 (– · · –); 28.6 (— —).

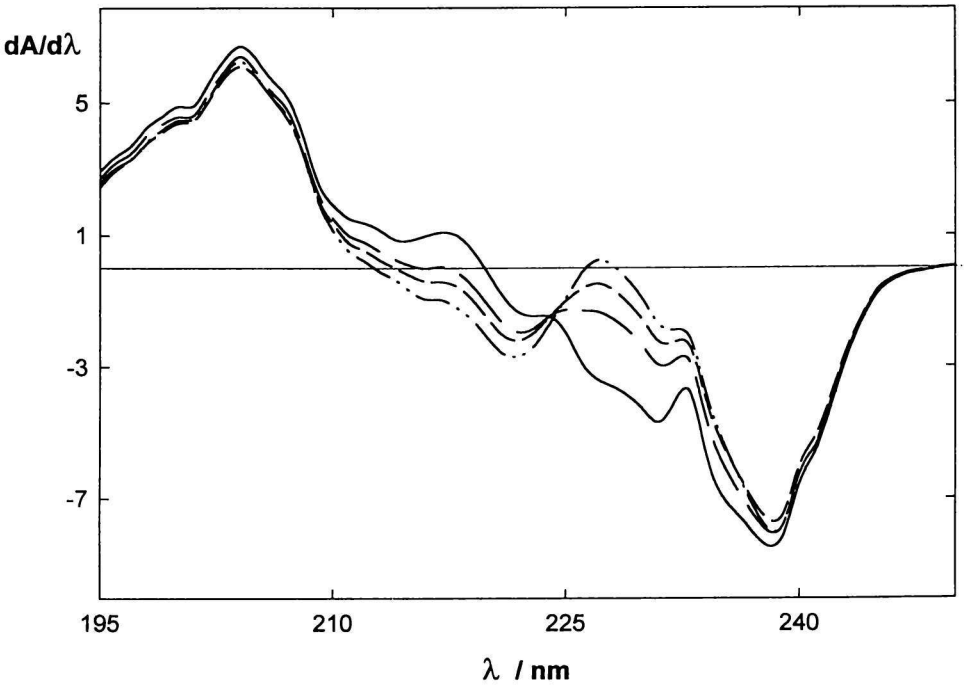


Fig. 3. First derivative absorption spectra for different contents of βN in the wavelength range 190—250 nm at 0.010 mmol dm^{-3} αN concentration; βN content/%: 0 (– · · –); 0.5 (— —); 4.7 (– – –); 20 (——).

RESULTS AND DISCUSSION

All aromatic compounds are strongly absorbing in the 200—260 nm range. We can suggest that also impurities will be absorbing. However, the aim of this work is to control αN and βN content in highly pure preparatives where we can suggest that the content of impurities will be low. The absorption spectra of the pure compounds measured are shown in Fig. 1. In these spectra there are two wavelength zones, 190—250 nm and 260—345 nm, where there are sufficient differences between the spectra of αN and βN. In the range 260—345 nm, the differences in the absorption spectra are greater than in the 190—250 nm region, but the sensitivity is lower. Because of the differences in sensitivity in the two wavelength ranges, higher αN and βN concentrations were used in the range 260—345 nm than in the former one.

The spectra obtained for mixtures of βN and αN, for the two wavelength ranges, keeping αN concentration constant, are shown in Figs. 2*a* and *b*. Remarkable differences were found if more than 5 % of βN content was present but, if βN concentration is lower than 5 % the spectrum of such a mixture is almost the same as that of the pure compound. We should notice that the concentrations of the components are different for the two ranges, due to the sensitivity, as was mentioned above (details are given in the text to the figures).

Derivative spectra from the first to the fourth derivative were calculated. However, we have to be careful because using derivative spectra, the statistical distribution is changing and also the errors can be magnified. In spite of this, there are many papers dealing with the use of derivative spectra for multicomponent analysis; mixtures with quite similar absorbance spectra of the components can be well resolved using derivative spectra [10]. For the first derivative spectra, greater differences among the spectra were observed in the range of 190—250 nm (Fig. 3).

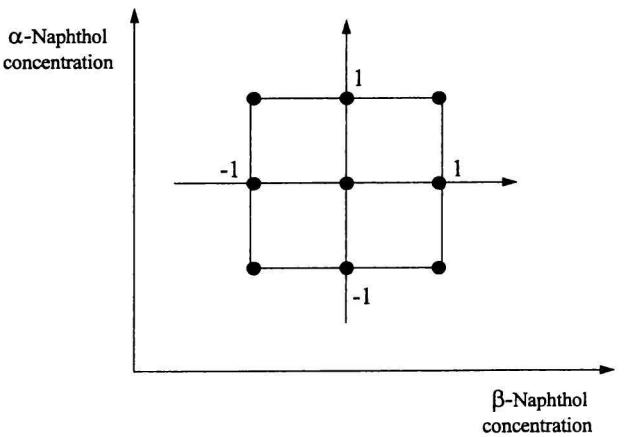


Fig. 4. Experimental design of the calibration set.

The calibration matrix was designed applying a multivariate factorial design (Fig. 4) and then used for the prediction of the test set (samples). Due to the different sensitivity in the wavelength ranges, different concentration levels were used in each one (see Table 1). The data of original and derivative spectra were processed with the PLS program using scaling

Table 2. Results of PLS Method Evaluation of the Data of the First Derivative Spectra for Binary Mixtures of βN and αN (Test Set) and for the Wavelength Range 190—240 nm

αN				βN			
c(given) · 10^6	c(found) · 10^6	SD	Content	c(given) · 10^6	c(found) · 10^6	SD	Content
mol dm^{-3}	mol dm^{-3}	%	%	mol dm^{-3}	mol dm^{-3}	%	%
13.8305	13.862	0.23	99.51	0.0677	0.0941	39.03	0.49
18.0619	18.029	0.19	99.24	0.1376	0.1738	26.29	0.76
23.0508	22.594	1.98	99.12	0.2039	0.1674	17.88	0.88
13.8305	12.914	6.62	98.54	0.2039	0.1891	7.25	1.45
18.0619	18.056	0.04	98.50	0.2752	0.2824	2.62	1.50
13.5474	13.169	2.79	98.50	0.2064	0.2097	1.61	1.50
13.5474	13.462	0.63	98.50	0.2064	0.2122	2.80	1.50
13.5474	13.713	1.22	98.50	0.2064	0.2021	2.08	1.50
13.8305	12.782	7.58	97.60	0.3398	0.3400	0.07	2.40
4.6102	4.463	3.19	95.77	0.2039	0.2135	4.71	4.23

SD = Standard deviation.

Table 3. Results of PLS Method Evaluation of the Data of the Normal Spectra for Binary Mixtures of βN and αN (Test Set) and for the Wavelength Range 240—365 nm

αN				βN			
c(given) · 10^5	c(found) · 10^5	SD	Content	c(given) · 10^5	c(found) · 10^5	SD	Content
mol dm^{-3}	mol dm^{-3}	%	%	mol dm^{-3}	mol dm^{-3}	%	%
19.0309	19.078	0.247	99.26	0.14265	0.18639	30.661	0.74
12.6873	12.639	−0.383	98.89	0.14265	0.18257	27.984	1.11
19.0309	19.078	0.247	98.66	0.25936	0.18639	−28.136	1.34
12.6873	12.728	0.318	98.00	0.25936	0.28965	11.679	2.00
6.3436	6.237	−1.682	97.80	0.14265	0.13475	−5.537	2.20
6.3436	6.4436	2.375	97.07	0.25936	0.27199	4.869	3.93
19.0309	18.750	−1.477	93.03	1.4265	1.4152	−0.788	6.97
6.3436	6.4368	1.468	81.64	1.4265	1.3992	−1.912	18.36

SD = Standard deviation.

factor equal to one.

After processing all data (original spectra and up to fourth derivative spectra) by PLS, it was noticed that better results were found for the 190—250 nm range using the first derivative spectra data and for the range 260—345 nm, using the normal spectra data. During the processing of the data, the number of wavelengths used in order to optimize the results, was studied. Finally, 23 wavelengths (196—240 nm, step 2 nm) and 16 wavelengths (262—294 nm, step 2 nm) for the 190—250 nm and 260—345 nm wavelength ranges were selected, respectively.

PLS program was run applying several abstract components (2 to 6). The lowest value of the residual square sum in both wavelength ranges was obtained for 4 abstract components (0.008 and 0.01 for 190—250 nm and 260—345 nm, respectively) by the cross validation method. For higher value of the abstract components, overfitting was found.

For 190—250 nm range of the wavelengths, the results of PLS method obtained for the prediction of the test set (first derivative spectra) are summarized in Table 2. From these results it follows that the content of βN equal to 1.5 % could be quantified with a relative standard error in the range 1.61—2.80 %. A satisfactory value of the relative error equal to 1.2 % (average) for the calibration set was obtained.

The results achieved for the test set (absorption spectra) for the range 260—345 nm are summarized in Table 3. The content of βN above 2.0 % (rel.) can be quantified with a relative error lower than 5.6 % for βN, while relative error of 2.1 % (average) for the whole calibration set was obtained. Comparison of the results from both wavelength ranges is quite difficult, because different concentrations and different types of spectral data were used in both regions.

The spectrophotometric analysis gave quite satisfactory results for βN contents higher than 2 % and also acceptable ones for the lower contents. However, we suggest that for the content below the mentioned content (2 %) the determination *e.g.* by capillary zone electrophoresis [5] or HPLC techniques [3, 4] could be done preferably. Nevertheless, the spectrophotometric method developed could be applied to analyze rapidly and simply the mixtures with higher βN content.

CONCLUSION

PLS calibration method may be successfully applied to evaluation of spectrophotometric data for βN and αN mixtures. Spectrophotometric determination applying the PLS calibration method permits a suitable, simple, and fast analysis of α- and β-naphthols. Derivative spectra for the 190—250 nm λ-region while normal spectra for the 260—345 nm region should be used to obtain the best results. The determination of βN content higher than 2 % (in the presence of an αN excess) by this procedure gives acceptable results. For βN content higher than 1.5—3.0 %, the relative standard deviation around 2 % was obtained. If βN content is lower than 2 % the concentration value can also be predicted, but the relative standard deviation is then much higher (about 30—40 %).

Acknowledgements. Revilla, A. L. and Vargas, G. would like to thank the U.N.A.M. (National Autonomous University of Mexico) for its invaluable support.

REFERENCES

1. *Handbook of Chemistry and Physics.* (David R. Lide, Editor.) 65th Edition. P. 380. CRC Press LLC, U.S.A., 1976.
2. Borodkin, V. F., *Chemistry of Organic Dyes.* Nakladatelství technické literatury (Publishers of Technical Literature), Prague, 1987.
3. Vrchlabský, M., Pollaková, N., Navrátilová, St., and Kuthan, P., *Proc. Colorchem '94, Vth International Conference*, Research Institute for Organic Synthesis (VÚOS), Czech Republic, 1994, p. 115.
4. Pollaková, N., Vrchlabský, M., Kahle, V., Haruda, J. and Kuthan, P., *Scr. Fac. Sci. Nat. Masaryk Univ. Brun. 25*, 3 (1994).
5. Revilla, A. L., Havel, J., and Jandik, P., *J. Chromatogr. A 745,* 225 (1996).
6. Wold, H., *System under Indirect Observation.* (Jöreskog, K. G. and Wold, H., Editors.) Vol. *2.* North-Holland, Amsterdam, 1982.
7. Lorber, A., Wangen, L. E., and Kowalski, B. R., *J. Chemom. 1,* 19 (1987).
8. Lindberg, W., Öhman, J., Wold, S., and Martens, H., *Anal. Chem. Acta 41,* 174 (1985).
9. Wold, S., Kettaneh-Wold, N., and Skagerberg, B., *Chemom. Intell. Lab. Syst. 7,* 53 (1990).
10. Jiménez, A., Jiménez, F., and Arias, J. *Analyst 114,* 93 (1989).