

Hyperplane Intercept Vector Calibration Method in Grey Analytical Systems

Huajie LIU, Kuiyi YOU, Pingle LIU,[†] and He'an LUO[†]

College of Chemical Engineering, Xiangtan University, Xiangtan 411105, P. R. China

A novel vector calibration method, hyperplane intercept, is proposed for grey analytical systems to resolve the concentrations of the chemical components from the multi-component data gained from attenuated total reflectance-Fourier transform infrared (ATR-FTIR) spectroscopy. This method is capable of quantifying a particular component known to be present in the mixture without having to know the identity of the rest of the components. Through the target factor analysis (TFA), a section of the IR spectrum of the interested component is chosen to calculate the concentration by the proposed hyperplane intercept method. Results calculated from on-line ATR-FTIR spectroscopy data of cyclohexanone ammoximation process are used to illustrate the simplicity and efficiency of this proposed method.

(Received January 17, 2010; Accepted March 2, 2010; Published May 10, 2010)

Introduction

In situ mid-IR monitoring can provide a large amount of information concerning a chemical reaction. This information can be used for fundamental mechanism study,^{1,2} reaction kinetic investigation³⁻⁵ and process control.⁶⁻⁹ As for mid-IR monitoring of some less complicated reactions, calibration using a single frequency or the ratio of two frequencies has been proved adequate to obtain concentration profiles.^{10,11} But for reactions that are more complex, it is still a challenge to obtain concentration profiles because of spectra overlap and collinearity of spectra matrix.¹² In order to keep track of the concentration of components in a reaction, a variety of chemometric methods are used. Partial least squares (PLS) regression has been used to analyze *in situ* mid-IR monitoring data of an acid-catalyzed esterification reaction in toluene.¹³ PLS calibration is unfeasible if the samples cannot be analyzed with a reference method. This is the case when a reaction is largely unknown or external analysis is not possible, *e.g.*, due to sampling problems. In addition, the development of reference analysis methods is time-consuming. Later, more research is focused on multivariate curve resolution (MCR)¹⁴⁻¹⁹ which is based on self-modeling curve resolution. However, the main problem with MCR is the rotational ambiguities associated with the solutions.¹⁴

The analytical systems can be divided into white, grey and black systems.²⁰ Generally, PLS is used in white analytical systems and MCR is used in black analytical systems. A chemical reaction monitored by *in situ* mid-IR is a grey analytical system and the obtained mid-IR spectra are vector data. It is a difficult chemometric task to quantify the desired analyte concentrations from the vector data in grey analytical systems. Rutan *et al.*^{21,22} proposed an adaptive kalman filtering method, which can be used to resolve overlapping spectral variables if the calibration model is incomplete or inaccurate.

However, the calibration model information must be selective for some wavelength region in the spectral response,²³ and the result is not good if more than one interferent exists. Later, a method called the additional iterative target transformation factor²⁴ was developed under the condition that concentration of interferents remained unchanged. By adding a series of standards which were to be tested, the grey analytical system would be transformed into white analytical system, and then it could be analyzed by calibration methods for white analytical systems. However, this method is complicated and not suitable for the analytical system in which concentration of interferents are changeable. Karstang *et al.*²⁵ proposed a local curve fitting technique (LCF). In this method, the interferents are modelled and then differentiation is used to remove the interferents at their local maxima and minima. If the modeling of the interferents is good, LCF performs well and the results can be verified for internal consistency by using models obtained at several maxima and minima. However, to obtain the maxima or minima of the background by using models is not easy, especially when the numbers of unknown interferents are large and/or the spectra overlap very seriously. Recent theories of multivariate calibration methods for grey analytical systems summarized by Liang²⁶ considered that it was impossible to obtain the unique physically meaningful solutions because the spectra of the interferents were unknown.

In this work, we would like to report a novel vector calibration method, hyperplane intercept (HI), for grey analytical systems to resolve the concentrations of the chemical components. Compared with the above-mentioned three vector calibration methods, it has the advantage of being able to obtain the unique solution, operation is easy and the method is suited to the analytical systems in which the concentration of the interferents are changeable. This method is applied to deal with the data of on-line attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) monitoring cyclohexanone ammoximation reaction to calculate the concentration of cyclohexanone and cyclohexanone-oxime, and the obtained results are better than those from separate peak method based on MCR.

[†] To whom correspondence should be addressed.
E-mail: liupingle@xtu.edu.cn; heanluo@126.com

Theory

Lambert-Beer's law for a grey analytical system can be expressed as

$$y_t = c_A x_A + \sum_{i=1}^q c_i x_i + e \quad (1)$$

Here y_t denotes the test sample spectrum vector, x_A denotes the standard spectrum vector of the desired component A (A stands for the measured component), c_A denotes the concentration of component A, x_i ($i = 1, 2 \dots q$) denotes the standard spectrum of the i th interferent, c_i stands for the concentration of the i th interferent and e denotes the error vector.

Suppose a grey analytical system with three components A, I_1 and I_2 . Herein, A is the component of interest, whose spectrum vector is x_A ; I_1 and I_2 are the interferents, whose spectrum vectors are x_1 and x_2 , respectively, these are unknown. Suppose y_t is the spectrum vector of the mixture sample to be tested. If we know the vectors x_1 and x_2 , the concentration of A can be calculated by Eq. (2).

$$y_t = [x_A, x_1, x_2]c \quad (2)$$

Where $c = [c_A, c_1, c_2]^t$. The superscript t indicates a transposed matrix or vector. The process to obtain c_A from Eq. (2) can be shown in Fig. 1. A directed line segment in three-dimensional vector space Ψ denotes a vector. $\overrightarrow{OX} = x_A$, $\overrightarrow{OX_1} = x_1$, $\overrightarrow{OX_2} = x_2$, $\overrightarrow{OY} = y_t$. The Greek letter ω denotes the OX_1X_2 -plane. Through the terminal point Y of vector y_t and parallel to the plane ω , a plane ω_1 is obtained. Plane ω_1 crosses vector x_A at point C. The quotient obtained by dividing the length of vector \overrightarrow{OC} by the length of vector \overrightarrow{OX} is equal to c_A . $\overrightarrow{OC} = c_A x_A$, $\overrightarrow{CY} = y_t - c_A x_A$. However, as the vectors x_1 and x_2 are unknown, the plane ω is unknown. So plane ω_1 can not be obtained. But we thought of a clever approach to obtain ω_1 .

Suppose y_1 and y_2 are two known mixtures. In Fig. 1, $\overrightarrow{OY_1} = y_1$, $\overrightarrow{OY_2} = y_2$. The concentrations of A in these two mixtures are c_{A1} and c_{A2} , respectively. As y_1, y_2 and x_A are linearly independent, then $y_1 - c_{A1}x_A, y_2 - c_{A2}x_A$ are linearly independent. In Fig. 1, $\overrightarrow{OC_1} = c_{A1}x_A$, $\overrightarrow{OC_2} = c_{A2}x_A$, $\overrightarrow{C_1Y_1} = y_1 - c_{A1}x_A$, $\overrightarrow{C_2Y_2} = y_2 - c_{A2}x_A$. With geometric vectors $\overrightarrow{C_1Y_1}$ and $\overrightarrow{C_2Y_2}$, a plane ω_2 can be obtained. Plane ω_2 parallels plane ω . So plane ω_1 can be obtained from plane ω_2 . As vector \overrightarrow{CY} parallels vectors $\overrightarrow{C_1Y_1}$ and $\overrightarrow{C_2Y_2}$, so the vector $y_t - c_A x_A$ is a linear combination of vectors $y_1 - c_{A1}x_A, y_2 - c_{A2}x_A$, then there are scalars k_1, k_2 , not all of which are zero, such that

$$y_t - c_A x_A = [(y_1 - c_{A1}x_A), (y_2 - c_{A2}x_A)]k \quad (3)$$

where $k = [k_1, k_2]^t$. Concentration of A can be calculated through Eq. (3).

If there are $q + 1$ components in a grey analytical system, ω, ω_1 and ω_2 is a q -dimensional vector subspace of the $q + 1$ -dimensional vector space Ψ . Subspaces ω, ω_1 and ω_2 are defined as hyperplanes of Ψ . Where plane ω_1 crosses vector x_A is called the hyperplane intercept.

Vector $y_t - c_A x_A$ is a linear combination of $y_1 - c_1 x_a, y_2 - c_2 x_a \dots y_q - c_q x_a$, then there are scalars $k_1, k_2 \dots, k_q$, not all of which are zero, such that

$$y_t - c_A x_A = [(y_1 - c_{A1}x_A), (y_2 - c_{A2}x_A) \dots (y_q - c_{Aq}x_A)]k \quad (4)$$

where $k = [k_1, k_2 \dots k_q]^t$.

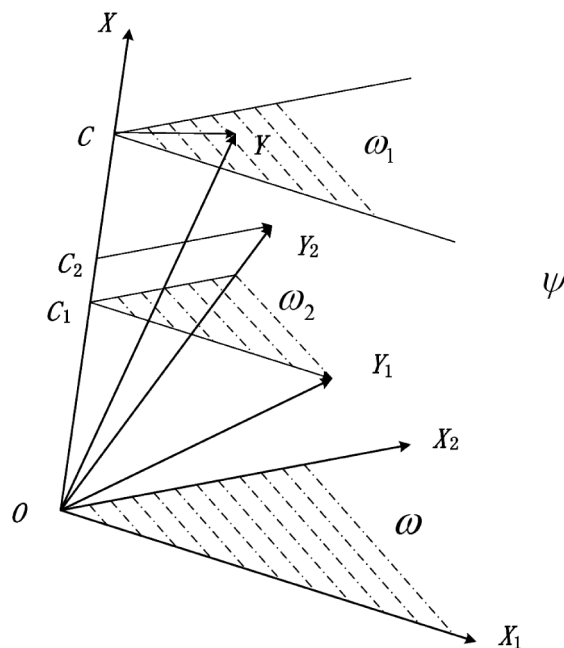


Fig. 1 Geometric representation of the spectra vectors in three-dimensional vector space in three component systems.

It is an over-determined system of linear equations. Let $M = [(y_1 - c_{A1}x_A), (y_2 - c_{A2}x_A) \dots (y_q - c_{Aq}x_A)]$. Then

$$y_t - c_A x_A = Mk \quad (5)$$

In order to obtain the concentration of A from Eq. (5), the variable step size search algorithm was developed (see Supporting Information).

Experimental

Materials and instrumentation

The titanium silicate catalysts were supplied by Hunan Jian Chang Chemical Co. Ltd. The materials cyclohexanone (>99%), hydrogen peroxide (28.84 wt% hydrogen peroxide in water), ammonia (22.67%), cyclohexanone-oxime (>99%), and *tert*-butyl alcohol (>99%) were obtained from commercial suppliers (China). Water is twice-distilled.

A ReactIR 4000 reaction analysis system from Mettler was used to collect the mid-FTIR spectra.

Experimental procedure

Each reaction was performed in a slurry reactor consisting of a 250-ml three-necked flask fitted with a condenser. The titanium silicate was used as catalyst, H_2O_2 as oxidant and *t*-butyl alcohol as the solvent. Materials in accordance with a certain order were added to the reactor. The mixture was stirred and heated up to 347.15 K by circulating hot water through the jacket of the reactor. The ReactIR 4000 probe was inserted into the reactor to collect mid-FTIR spectra of the reactions on-line (see Fig. 2). The spectra region is between 650 and 4000 cm^{-1} . The experimental conditions consisted of resolution, 8 cm^{-1} , and number of scans, 64. Spectra were collected every 1 min.

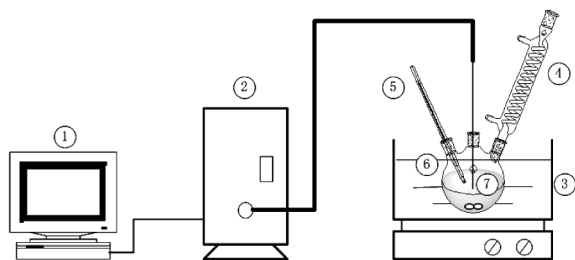


Fig. 2 Reactor system utilized in the present study. 1, Computer; 2, ReactIR 4000; 3, constant water bath and magnetic stirring apparatus; 4, condenser; 5, thermometer; 6, three-necked reaction flask; 7, diamond-composite insertion probe.

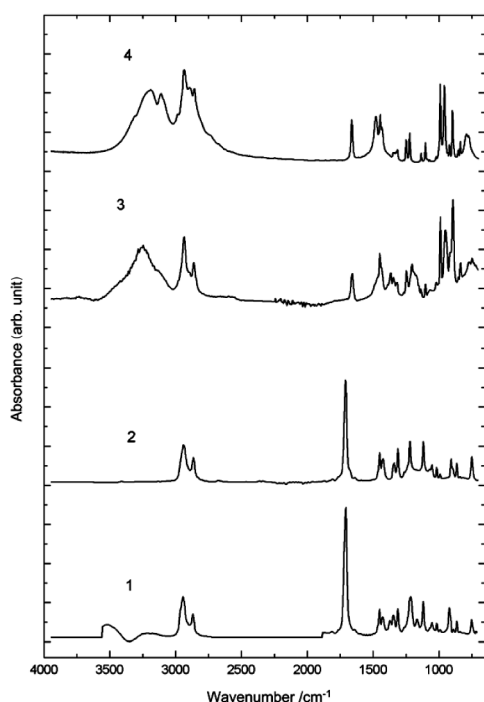


Fig. 3 FTIR spectra of cyclohexanone under different conditions. 1, Cyclohexanone in solvent; 2, pure cyclohexanone; 3, cyclohexanone-oxime in solvent; 4, pure cyclohexanone-oxime.

Results and Discussion

Spectra vectors of cyclohexanone and cyclohexanone-oxime

As mid-IR absorptions are changed in frequency and band shape due to intermolecular forces, *e.g.*, hydrogen bonding and temperature, the spectrum vector of component in the solvent must be obtained instead of the pure component. The concentration vector of a series of the mixture of cyclohexanone, cyclohexanone-oxime and solvent is used to obtain the spectra vectors of cyclohexanone and cyclohexanone-oxime in solvent by deconvolution (Fig. 3). The spectra of pure cyclohexanone and cyclohexanone-oxime are also plotted in Fig. 3. It can be seen that the spectra in solvent is different from that of pure component due to the influence of the solvent. It demonstrates that spectrum of the pure component could not be used as a real factor in this study.

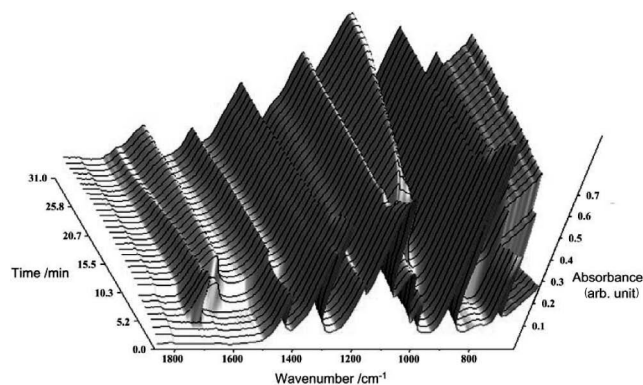


Fig. 4 The three-dimensional plot of the FTIR spectra of the cyclohexanone ammoxidation at 347.15 K.

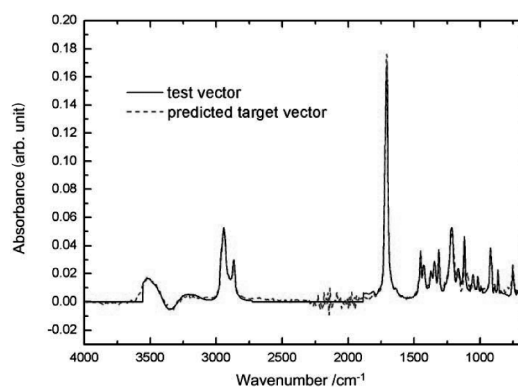


Fig. 5 Test vector and predicted vector of cyclohexanone.

Concentration profile obtained from hyperplane intercept

Figure 4 depicts a three-dimensional plot of the spectra that recorded the cyclohexanone ammoxidation reaction for the first 31 min. In order for us to obtain the region of the IR spectrum which is more close to the real situation, the target factor analysis (TFA)²⁷ was adopted to examine the target spectrum vector. In this step, the spectrum vector of cyclohexanone in solvent was selected as the test vector. Through the three-dimensional spectra data of cyclohexanone ammoxidation reaction, the predicted target vector of cyclohexanone could be obtained, this is shown in Fig. 5. It can be seen that the predicted target vector agrees well with the test vector, especially between wavenumbers 1683 and 1740. This demonstrates that the spectrum of cyclohexanone in this area is very close to the real situation. From wavenumber 1683 to 1740, the principal component analysis (PCA)²⁸ was used, and we found that there are two components, that is to say, there is one interferent in this region. The test vector and the predicted target vector of cyclohexanone-oxime are shown in Fig. 6. Wavenumbers 1660 and 1675 were chosen to calculate the concentration of cyclohexanone-oxime.

The concentration of cyclohexanone at the point of initial reaction is 0.867 mol L⁻¹, and at the end point of the reaction is 0.303 mol L⁻¹. Through the HI, the concentration profile of cyclohexanone can be calculated. The value of one of the $|\delta|$ (δ is the error vector in the variable step size search algorithm for HI, see Supporting Information) *versus* iteration is shown in Fig. 7. It is clearly that the $|\delta|$ decreases with the number of iterations. A comparison between the concentration profiles of

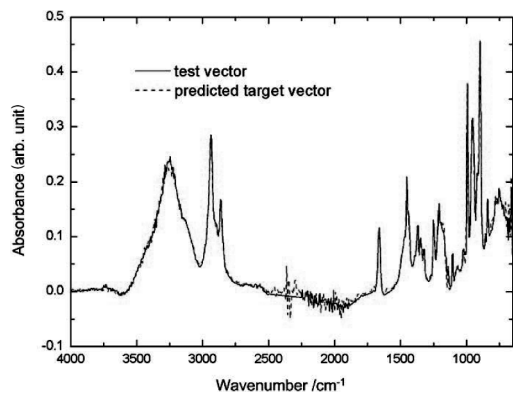


Fig. 6 Test vector and predicted vector of cyclohexanone-oxime.

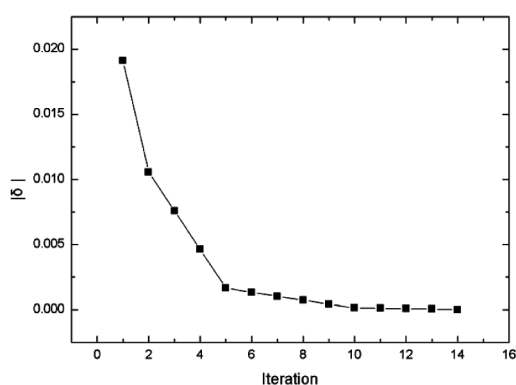


Fig. 7 Plot of the $|\delta|$ versus iteration number.

cyclohexanone obtained from HI and the method of separate peak method based on MCR (analytical software supplied by Mettler) is shown in Fig. 8. It can be seen that there is little difference of the concentration profiles from these two methods.

Through the same process, the concentration profile of cyclohexanone-oxime can be obtained. The concentration profiles of cyclohexanone-oxime obtained from HI and separate peak method are plotted in Fig. 9, and the concentration profile of cyclohexanone-oxime calculated from the cyclohexanone ammoximation reaction kinetics²⁹ is also plotted in Fig. 9. It can be clearly shown that the concentration profile of cyclohexanone-oxime from HI is close to the result calculated from reaction kinetics. In fact, the concentration of cyclohexanone-oxime during the cyclohexanone ammoximation process is impossible to reach the concentration profile obtained from the separate peak method, even if all of the converted cyclohexanone was supposed to produce cyclohexanone-oxime. This demonstrates that the result obtained by HI is better than that from the separate peak method.

In order to demonstrate the accuracy and reliability of HI, we sampled the reaction mixture at 8 min intervals and the datum was quantitatively analyzed by GC internal method. The concentrations of cyclohexanone and cyclohexanone-oxime from GC are 0.754 and 0.104 mol L⁻¹, respectively, while the concentrations of cyclohexanone and cyclohexanone-oxime obtained from HI are 0.745 and 0.091 mol L⁻¹, respectively, and the results from the separate peak method are 0.742 and 0.312 mol L⁻¹, respectively. As can be seen, although the concentration of cyclohexanone-oxime from HI is a little

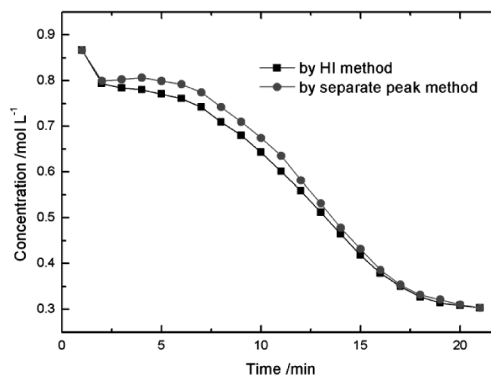


Fig. 8 Concentration profile of cyclohexanone calculated by different methods.

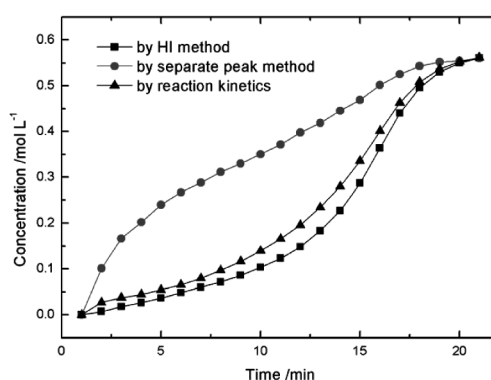


Fig. 9 Concentration profile of cyclohexanone-oxime calculated by different methods.

different from the result of GC, it is much better than that from the separate peak method. The reason may be that the spectrum of the desired component is used in HI, while the separate peak method based on MCR does not use this information.

Conclusions

Based on the theory of the vector space, a new hyperplane intercept (HI) method was proposed, and it is a simple and effective vector calibration method. It was applied to deal with the data of on-line ATR-FTIR spectra of cyclohexanone ammoximation. The result demonstrates that this method could extract useful and reliable information from the data of spectrum vectors. Compared with the separate peak method, the result from HI is more close to reality. The consistency of the results analyzed by GC and HI further shows the reliability of the proposed HI method. Since there is no restriction on the type of vector data, this method is general and should thus be applicable to deal with other types of vector data under the condition that unknown interfering constituents are a major problem for quantitative analysis.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (20976148; 20736009), Scientific

Research Fund of Hunan Provincial Education Department (09A090), Hunan Provincial Natural Science Foundation of China (07JJ6017) and Master Innovation Foundation of Hunan Province (s2008yjsx15).

Supporting Information

The variable step size search algorithm is available free of charge on the web at <http://www.jsac.or.jp/analsci/>.

References

1. X. C. Luo, H. Zhang, H. Duan, Q. Liu, L. Z. Zhu, T. Zhang, and A. W. Lei, *Org. Lett.*, **2007**, 9, 4571.
2. F. Visentin, G. Puxty, O. M. Kut, and K. Hungerbühler, *Ind. Eng. Chem. Res.*, **2006**, 45, 4544.
3. R. V. Wisman, J. G. Vries, B. J. Deelman, and H. J. Heeres, *Org. Process Res. Dev.*, **2006**, 10, 423.
4. D. G. Blackmond, M. Ropic, and M. Stefinovic, *Org. Process Res. Dev.*, **2006**, 10, 457.
5. A. Matsumoto, T. Otaka, and H. Aota, *Macromol. Rapid Commun.*, **2001**, 22, 607.
6. A. Friebe and H. W. Siesler, *Vib. Spectrosc.*, **2007**, 43, 217.
7. N. A. Magnus, P. B. Anzeveno, D. S. Coffey, D. A. Hay, M. E. Laurila, J. M. Schkeryantz, B. W. Shaw, and M. A. Staszak, *Org. Process Res. Dev.*, **2007**, 11, 560.
8. B. Kappler, A. Tuchbreiter, D. Faller, P. Liebetraut, W. Horbelt, and J. Timmer, *J. Honerkamp, Polymer*, **2003**, 44, 6179.
9. A. Pintar, R. Malacea, C. Pinel, G. Fogassy, and M. Besson, *Appl. Catal., A*, **2004**, 264, 1.
10. W. M. Doyle and N. A. Jennings, *Spectrosc. Int.*, **1991**, 2, 48.
11. W. R. Moser, J. R. Berard, and P. J. Melling, *Appl. Spectrosc.*, **1992**, 46, 1105.
12. M. Amrbein, B. Srinivasan, D. Bonwin, and M. M. Schumacher, *Chemom. Intell. Lab. Syst.*, **1996**, 33, 17.
13. P. MacLaurin, N. C. Crabb, I. Wells, P. J. Worsfold, and D. Coombs, *Anal. Chem.*, **1996**, 68, 1116.
14. E. Furusjö, L. G. Danielsson, E. Könberg, M. R. Jonas, and B. Skagerberg, *Anal. Chem.*, **1998**, 70, 1726.
15. J. R. Schoomover, R. Marx, and W. R. Nichols, *Vib. Spectrosc.*, **2004**, 35, 239.
16. J. Jaumot, V. Marchn, R. Gargallo, A. Grandas, and R. Tauler, *Anal. Chem.*, **2004**, 76, 7094.
17. S. Richards, M. Ropic, D. Blackmond, and A. Walnsley, *Anal. Chim. Acta*, **2004**, 519, 1.
18. M. Blanco, M. Castillo, and R. Beneyto, *Talanta*, **2007**, 72, 519.
19. S. C. Cruz, G. Rothenberg, and J. A. Westerhuis, *Chemom. Intell. Lab. Syst.*, **2008**, 91, 101.
20. Y. Z. Liang, O. M. Kvalheim, and R. Manne, *Chemom. Intell. Lab. Syst.*, **1993**, 18, 235.
21. S. C. Rutan and S. D. Brown, *Anal. Chim. Acta*, **1984**, 160, 99.
22. S. C. Rutan and S. D. Brown, *Anal. Chim. Acta*, **1985**, 167, 23.
23. D. D. Gerow and S. C. Rutan, *Anal. Chim. Acta*, **1986**, 184, 53.
24. Y. Z. Liang, Y. L. Xie, and R. Q. Yu, *Chin. Sci. Bull.*, **1989**, 34, 1533.
25. T. V. Karstang and O. M. Kvalheim, *Anal. Chem.*, **1991**, 63, 767.
26. Y. Z. Liang and R. Q. Yu, "Chemometrics", **2003**, Higher Education Press, Beijing, 139.
27. E. R. Malinowshi, "Factor Analysis in Chemistry", 3rd ed., **2002**, Wiley, New York.
28. I. T. Jolliffe, "Principal Component Analysis", 2nd ed., **2002**, Springer, New York.
29. Y. X. Li, W. Wu, and E. Z. Min, *Chin. J. Chem. Eng.*, **2005**, 13, 32.