

Three Dimensional Image Analysis for Prediction of Vinyl Triethoxysilane Concentration and Porosity Percent of Methyl Methacrylate/Ethylene Glycol Dimethacrylate Copolymers using Scanning Electron Microscopy

Mahnaz Esteki,^a Taghi Khayamian^{*,a} and Hamid J. Naghash^b

^aDepartment of Chemistry, Isfahan University of Technology, Isfahan 84154, Iran

^bDepartment of Chemistry, Shahreza Islamic Azad University, Shahreza, Isfahan, Iran

Neste trabalho, micrografias eletrônicas de varredura de esferas de copolímero baseadas em metacrilato de metila (MMA), etilenoglicoldimetacrilato (EGDM) e vinil-trietoxissilano (VTES) foram usadas para construir modelos de regressão multivariada de imagem. Microscopia eletrônica de varredura (SEM) e ensaios de intumescimento indicaram que aumentando a concentração de silicone (VTES) dos copolímeros ocorre um aumento da quantidade de poros e porcentagem de porosidade. Micrografias eletrônicas de varredura dos copolímeros sintetizados com diferentes concentrações de VTES foram usadas para construir dois modelos de calibração, que puderam relacionar a concentração de VTES e a porcentagem de porosidade dos copolímeros com suas imagens de SEM. Para construção dos modelos foram utilizados o método dos mínimos quadrados parciais multi-modo (N-PLS) e componentes principais desdobrados (u-PCR). O poder de previsão dos modelos construídos foi avaliado utilizando um conjunto teste. Os resultados indicaram que os modelos conseguiram prever de forma muito efetiva a concentração de VTES e a porcentagem de porosidade dos copolímeros a partir de suas imagens de SEM.

In this work, scanning electron micrographs of copolymer beads based on methyl methacrylate (MMA), ethylene glycol dimethacrylate (EGDM) and vinyl triethoxysilane (VTES) were used to construct multivariate image regression models. Scanning electron microscopy (SEM) and swelling measurements indicated that increasing the silicone concentration (VTES) of the copolymers enhances the amount of pores and porosity percent. Scanning electron micrographs of the synthesized copolymers with different VTES concentrations were used to construct two soft models, which were able to relate the VTES concentration and the porosity percent of the copolymers with their SEM images. The constructed models were the three-way partial least squares analysis (N-PLS) and unfolded principal component regression (unfold-PCR). The predictive ability of the constructed models were evaluated by a test set. The results showed the models were able to effectively predict the VTES concentration and the porosity percent of the copolymers using their corresponding SEM images.

Keywords: copolymers, scanning electron micrographs, N-PLS, unfold-PCR

Introduction

Digital images are representative of two dimensional data with a finite set of digital values, which any numerical, mathematical or statistical technique can be applied on them.¹ Images have been used in various scientific fields for a long time. Large amounts of data originate from complex systems might only be represented by visualization as images. Image analysis is the extraction of meaningful information from images by means of digital image

processing techniques. The role of image analysis in our modern society is becoming increasingly important and it is applied in different fields of science and technology. When images are collected in multiple variables such as different wavelengths or other parameters,² multivariate images data are produced. Multivariate image analysis³⁻⁸ (MIA) can be defined as image analysis using diverse chemometric tools.¹ The multivariate image analysis is an important technique due to its ability to extract quantitative and qualitative information. In this work, one of the applications of multivariate image analysis of the copolymer beads is examined.

*e-mail: taghi@cc.iut.ac.ir

Cross-linked copolymer beads (typically $> 100 \mu\text{m}$ diameter) are commonly used as basis for ion exchange resins, catalyst, reagent supports and enzyme immobilization.^{9,10} Such applications frequently require large particle surface areas, which requires the formation of pores in the bead structure. The copolymer beads size ranging from $1 \mu\text{m}$ to 1mm are produced by suspension polymerization.¹¹ Porosity in the copolymer beads may be created by the inclusion of an inert organic liquid (diluent) to monomer phase, during copolymerization process.¹²⁻¹⁴ Characteristics of the copolymer beads, like porosity and sorbent ability depend on the reaction conditions of polymerization such as diluent type and reagents' concentration. Therefore, it would be useful to construct a reliable model to predict these characteristics using minimum, but enough information about the copolymers.

In this regards, scanning electron micrographs of the copolymers are one of the most interesting kinds of information that can be used. Scanning electron micrograph is an image obtained by scanning electron microscopy (SEM) technique. In this work, the micrographs of copolymer beads based on methyl methacrylate (MMA), ethylene glycol dimethacrylate (EGDM) and vinyl triethoxysilane (VTES) were used to construct N-PLS and unfold-PCR models for relating the porosity percent and VTES concentration of the copolymer beads to their corresponding SEM images.

Multivariate image regression

Multivariate image analysis appears in all experimental fields of science and technology and its applications are growing at a fast rate. One of the most interesting features of MIA is the multivariate image regression (MIR).^{2,15} Multivariate image regression can be used as a tool for predicting quantitative features in multivariate image data. In this method, the main goal is to establish the relationship (regression) between images and physical or chemical parameters.¹

There are different methods of MIR including parallel factor analysis (PARAFAC), N-PLS, unfold-PLS and unfold-PCR. These methods can be categorized into weak and strong N-way methods.¹⁶

Unfold-PCR/PLS are considered weak multi-way methods because they actually unfold three-way data array

into two-way, followed by ordinary two way analysis, instead of utilizing the intact multi-way structure during the modeling.¹⁷

Unfold-PCR

Unfold-PCR is one of the N-way analysis methods which can be used in multivariate image regression. In this method, the three-way array \underline{X} of dimensions I, J and K becomes the two-way array X of dimensions $(I \times JK)$. In other words, each row of the two-dimensional array X corresponds to one image. Figure 1 shows how to unfold three-way data for two-way analysis in multivariate image regression. Since in PCR method, the regression model is built between scores (of the independent variables) and the dependent variables, therefore scores of the constructed matrix should be calculated.

Singular value decomposition (SVD) is used for this calculation. A singular value decomposition of a matrix X ($I \times JK$) is as follows:

$$X = UDV' = \sum_{a=1}^A u_a d_a v'_a$$

Here, U is a $(I \times I)$ matrix, which its columns are eigenvectors of XX' ; V is a $(IJ \times IJ)$ matrix, which its columns are eigenvectors of $X'X$; D is a diagonal matrix of size $(I \times JK)$, The diagonal elements d_a are the singular values, the square roots of the eigenvalues λ_a of XX' or $X'X$.¹⁸ In this work, because the unfold matrix is too large (5×344608), the classical singular value decomposition (SVD) of the matrix is a very time consuming task. The score vectors may be derived from the SVD of the correlation matrix (XX'). As a matter of fact, the eigenvectors of XX' are exactly the score vectors of a matrix X . Therefore the scores may be calculated by SVD on XX' matrix, and then the loading matrix can be obtained from the score matrix. This fact gives a simple method for calculating the scores and the loadings of a large matrix. This is shown schematically in Figure 1.

N-PLS

In this method, three dimensional images data are decomposed and processed simultaneously. Suppose \underline{X} is the $I \times L \times M$ array of independent data and X the $I \times JK$

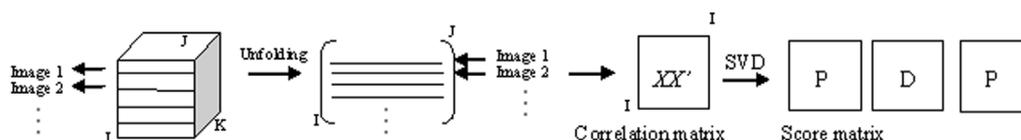


Figure 1. The unfold matrix X obtained from multivariate image \underline{X} , correlation matrix creates (XX'), score vectors calculated from SVD on the correlation matrix.

unfolded array; yet \underline{Y} is the $I \times L \times M$ array of dependent data and Y the $I \times LM$ unfolded array. The N-PLS model decomposes \underline{X} as:¹⁶

$$X = T(W^K | \otimes | W^J)^T + E_x$$

And \underline{Y} as:

$$Y = U(Q^M | \otimes | Q^L)^T + E_y$$

Where E is the residual and \otimes denotes the kronecker product. The score vectors of \underline{X} and \underline{Y} are called t and u , respectively, and the weight vectors are called w and q . A superscript J or K , L or M is used to specify which mode the vectors refer to. The above decomposition models are estimated component-wise; a set of vectors w^J , w^K , q^M , and q^L should be found in such a way that the least squares score vectors t and u have maximal covariance. Through these models of the data sets the prediction model between \underline{Y} and \underline{X} is found using a regression model for the so-called inner relation:

$$U = TB + E_u$$

by determining T , the scores in the \underline{Y} -space can be predicted and through the model of \underline{Y} the prediction of \underline{X} is obtained. Note that the decomposition of the model of \underline{X} , the decomposition model of \underline{Y} and the regression model relating these two models, constitute the N-PLS regression model.

Experimental

The procedure of preparation of copolymer beads has been reported in our previous published paper.¹⁹ In brief, the copolymers have been prepared using toluene as diluent, using the suspension polymerization technique. In the synthesized copolymers, concentration and chosen diluent (toluene) were kept constant, but the concentration of VTES was changed. The compositions of the synthesized copolymers are listed in Table 1. The apparent density of the copolymers, d_o , was determined by the mercury pycnometric method. Porosity percent, $P\%$, and pore volume, V_p , were calculated from d_o as :

$$P\% = (1 - d_o / d_2) \times 100\%$$

$$V_p = 1/d_o - 1/d_2$$

where d_2 is the density of homogeneous MMA-EGDM copolymers.

Table 1. Composition of the reaction mixtures of synthetic copolymers

sample number	MMA (mL)	EGDM (mL)	VTES (mL)	VTES conc. (mol L ⁻¹)	Toluene (mL)
calibration set:					
1 (a)	2.5	2.0	0.5	0.24	5.0
2 (b)	2.0	2.0	1.0	0.48	5.0
3 (d)	1.5	2.0	1.5	0.72	5.0
4 (e)	0.5	2.0	2.5	0.96	5.0
5 (h)	0.0	2.0	3.0	1.44	5.0
prediction set:					
6 (c)	1.8	2.0	1.2	0.58	5.0
7 (e)	1.0	2.0	2.0	0.96	5.0
8 (g)	0.3	2.0	2.7	1.29	5.0

The SEM experiments were performed on a Philips, XL 30 scanning electron microscope. Magnification for all samples was 500 and a clear image could be obtained. All images were achieved in a constant contrast and brightness. SEM imaging of the various copolymer beads also reveals some other information such as bead diameter and porosity patterns. All images are displayed in the gray scale. The N-PLS model was applied on a three-way array $5 \times 484 \times 712$, and unfold-PCR on 5×344608 two-way array. The N-PLS modeling was carried out with the N-way toolbox for MATLAB,²⁰ using a Pentium 4 computer for processing and analysis.

Results and Discussion

The porosity percent of synthesized copolymers with different VTES concentrations is shown in Figure 2. According to this figure, VTES concentration efficiently influences the porosity percent of the copolymers. A set of images consist of five SEM images from synthetic copolymer beads (five samples) with different VTES concentrations (Table 1, calibration set) were obtained without any pretreatment. SEM images are shown in Figure 3 (a, b, d, f, h). Figure 3 shows that different concentrations of VTES in the copolymer beads produce copolymers with different porosity patterns.

Image pre-treatment

In this work, logarithm transformation, a kind of nonlinear transformation, was used as a preprocessing method, since without performing this transformation the modeling results were poor. In some situations, non-linear scaling may be useful for image analysis.¹ This may be originated from the Beer-Lambert law. This law simply

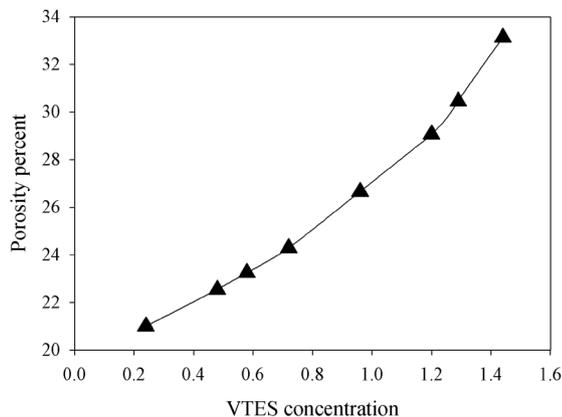


Figure 2. Porosity percent of the synthesized copolymers versus VTES concentration used in the reaction mixture.

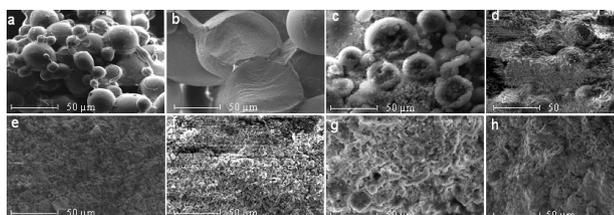


Figure 3. SEM images of synthetic copolymers with different concentration of VTES used in the reaction mixture. (a): 0.24, (b): 0.48, (c) 0.58, (d): 0.72, (e): 0.96, (f): 1.20, (g): 1.29 and (h): 1.44 mol L⁻¹.

states that absorbance is linearly related to concentration of absorbing species and since most of the imaging equipment measures transmission, its transformation to absorbance may be a useful tool. In order to obtain reasonable results in unfold-PCR modeling, mean-centering and auto-scaling were performed as a data pretreatment method on SEM images.

N-PLS modeling

In N-PLS model the number of latent variables is selected by plotting predicted residual error sum of squares (PRESS) in terms of number of latent factors, using full cross validation method,²¹ which is shown in Figure 4a. Usually the number of latent variables, which gives a minimum PRESS value is chosen as the optimum number. As it can be seen in Figure 4a, the minimum error of the calibration and cross-validation is obtained using two factors. Therefore, two latent variables were used in this model. Figure 4b shows the calculated *versus* experimental values of the VTES concentration. This figure shows that the predictive ability of the model is suitable.

Porosity percent of the copolymer beads was another property, predicted by means of the modeling SEM images.

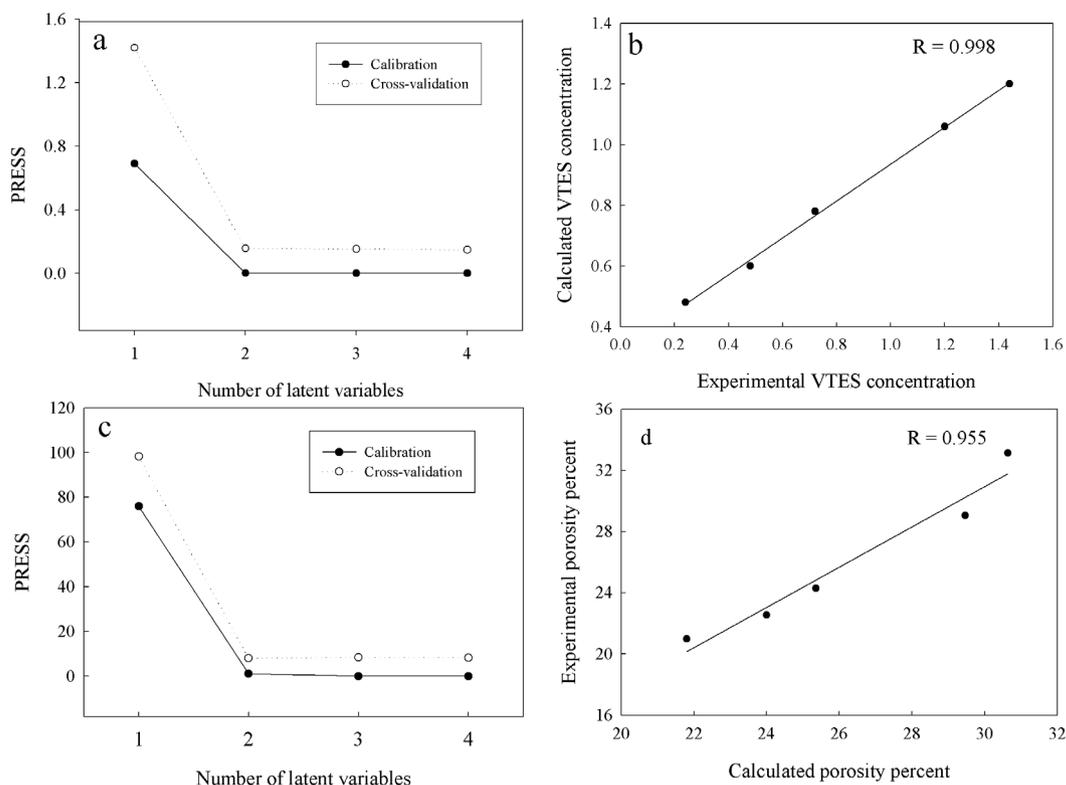


Figure 4. a) PRESS vs. number of latent variables (constructing model for predicting VTES concentration), using the full cross-validation method in the N-PLS modeling. b) Calculated vs. experimental value of VTES volumes using the constructed N-PLS model. c) PRESS vs. number of latent variables (constructing model for predicting the porosity percent), using full cross-validation method in the N-PLS modeling. d) Experimental vs. calculated values of porosity percent using the constructed N-PLS model.

Selection of the optimum number of latent variables, in this case, was accomplished based on the previous procedure. Figure 4c shows the PRESS for calibration and cross-validation *versus* number of latent variables. In this case, two latent variables are also suitable for the modeling. Figure 4d shows the experimental *versus* calculated values of the porosity percent of the copolymers.

Two-way PCR modeling

In order to investigate the capability of the unfold-PCR modeling, this type of the multivariate image regression was also used for predicting the mentioned parameters. In this case, the images were all unfolded into vectors of length 484×712 pixels row-wised, which subsequently results a matrix with dimensions of 5×344608 .

The PRESS values *versus* number of latent variables (Figure 5a), indicates that three latent variables are required for the modeling. The calculated *versus* experimental values of VTES concentrations is shown in Figure 5b. The porosity percent of the copolymers was also investigated using unfold-PCR model. The PRESS values *versus* number of latent variables for calibration and the cross-validation (Figure 5c) also represent that three latent variables are

the best choice for the modeling. Experimental *versus* calculated values of the porosity percent (Figure 5d) shows that the ability of the unfold-PCR model for predicting porosity percent is inferior than that in the N-PLS model. The results of the two models are listed in Table 2. According to this table, both models are able to relate VTES concentration and porosity percent of the copolymer beads to their images. However, the N-PLS models are superior to unfold-PCR for predicting VTES concentration and especially for porosity percent. Furthermore, the N-PLS model needs two latent variables for the modeling. This is essentially because strong multi-way methods are more parsimonious, with specific use of the three-way structure in the modeling.¹⁷ The tri-linear models are more restricted

Table 2. The comparison of the predictive ability of the N-PLS and unfold-PCR models for VTES concentration and the porosity percent of the porous copolymers

Calibration methods	Mentioned parameters	R	PRESS	No. of factors
N-PLS	VTES concentration	0.998	0.6679	2
	Porosity percent	0.955	7.9747	2
Unfold- PCR	VTES concentration	0.997	1.7312	3
	Porosity percent	0.760	22.880	3

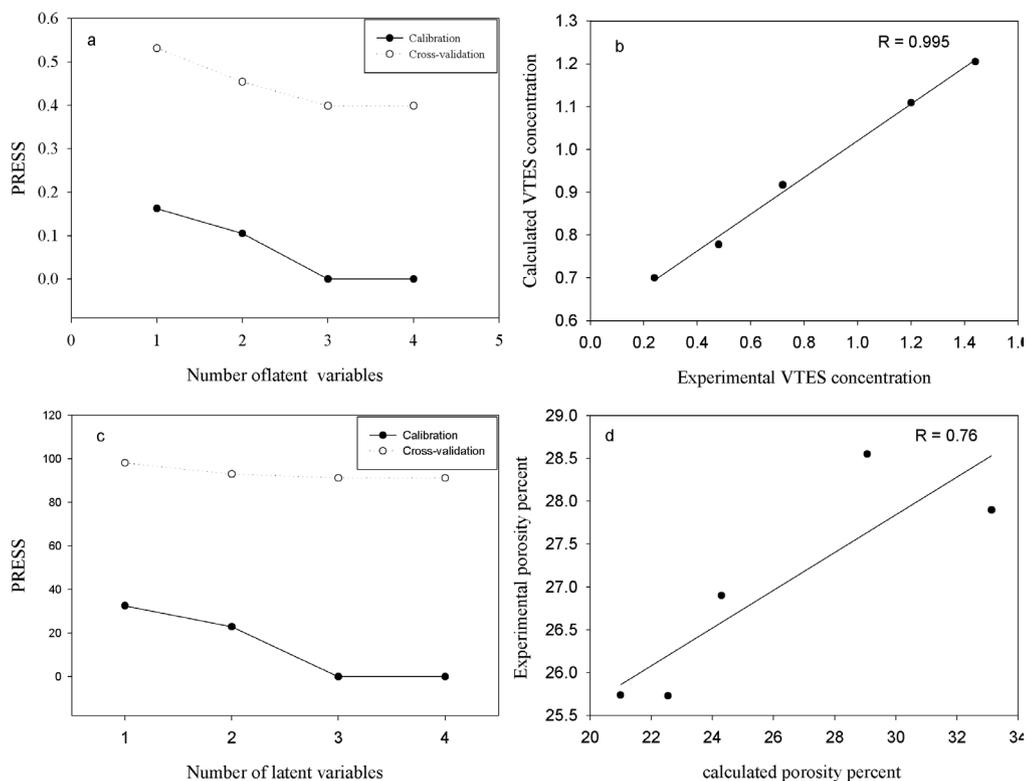


Figure 5. a) PRESS *vs.* number of latent variables (for predicting VTES concentration), using full cross-validation method in the unfold-PCR modeling. b) Calculated *vs.* experimental values of VTES concentrations using unfold-PCR model. c) PRESS *vs.* number of latent variables (for predicting the porosity percent), using full cross-validation method in the unfold-PCR modeling. d) Experimental *vs.* calculated values of porosity percent using the constructed unfold-PCR model.

Table 3. Experimental and predicted values of VTES concentration and porosity percent of the copolymers using constructed N-PLS and unfold-PCR models

copolymer #	calibration method	experimental VTES conc.	calculated VTES conc.	Relative error	Experimental Porosity percent	Calculated Porosity percent	Relative error
1 (c)	N-PLS	0.576	0.581	0.87	23.25	22.42	-3.57
	unfold-PCR	0.576	0.662	14.93	23.25	25.00	7.53
2 (e)	N-PLS	0.96	0.945	-1.5	26.65	26.57	-0.30
	unfold-PCR	0.96	0.878	-8.5	26.65	26.33	-1.20
3 (g)	N-PLS	1.29	1.162	-9.92	31.24	30.16	-3.46
	unfold-PCR	1.29	1.166	-9.61	31.24	28.82	-7.75

than unfolding methods and unfolding may risk taking away some three-way information and fitting more noise to the structural model.^{1,17}

Finally, in order to confirm the predictive ability of the models, another three SEM images of synthetic copolymers (Figures 3c, e and g and Table 1, prediction set) were used for predicting VTES concentration and porosity percent. The predicted results are tabulated in Table 3 showing that both models are able to predict the parameters with the satisfactory results. However, as expected, the N-PLS model is superior to unfold-PCR.

Conclusion

The N-PLS and unfold-PCR models were constructed between SEM images of the copolymer beads and their corresponding VTES concentrations and porosity. Comparison of the results demonstrates that the predictive ability of N-PLS is superior to unfold-PCR model. The results indicate that SEMs of the copolymers could be used as an alternative technique for predicting some of their characteristics. This relationship between some copolymer components properties with their SEM images using multi-way models opens a new research area for further investigation.

Acknowledgments

The authors express their appreciation to the Isfahan University of Technology Research Councils for financial support of this work.

References

- Geladi, P.; Grahn, H.; *Multivariate Image Analysis*, John Wiley & Sons Press: New York, 1996, ch. 1.
- Lied, T.; Esbensen, K. H.; *Chemom. Intell. Lab. Syst.* **2001**, *58*, 213.

- Geladi, P.; Burger, J.; Lestander, T.; *Chemom. Intell. Lab. Syst.* **2004**, *72*, 209.
- Geladi, P.; Sethson, B.; Nystrom, J.; Lilhonga, T.; Lestander, T.; Burger, J.; *Spectrochim. Acta B* **2004**, *59*, 1347.
- Geladi, P.; *Microchim. Acta.* **1995**, *120*, 211.
- Geladi, P.; Esbensen, K. In *Multivariate Image Analysis in Chemistry: An Overview*; Devillers, J.; Karcher W., eds., Kluwer: Dordrecht, 1991, p. 415.
- Geladi, P.; Grahn, H.; Esbensen, K.; Bengtsson, E.; *Trends Anal. Chem.* **1992**, *11*, 121.
- Esbensen, K.; Geladi, P.; *Chemom. Intell. Lab. Syst.* **1989**, *7*, 67.
- Dorfner, K.; *Ion Exchangers*, Walter de Gruyter Press: Berlin, 1991.
- Dowding, P. J.; Goodwin, J. W.; Vincent, B.; *Colloids Surf. A* **1998**, *145*, 263.
- Arshady, R.; *J. Chromatogr. A* **1991**, *586*, 181.
- Kun, K. A.; Kunin, R.; *J. Polym. Sci. A* **1968**, *6*, 2689.
- Sederel, W. L.; De Jong, G. J.; *J. Appl. Polym. Sci.* **1973**, *17*, 2835.
- Jacobelli, H.; Bartholin, M.; Guyot, A.; *J. Appl. Polym. Sci.* **1979**, *23*, 927.
- Geladi, P.; Esbensen, K.; *J. Chemom.* **1991**, *5*, 97.
- Bro, R.; *PhD Thesis*, University of Amsterdam (NL) & Royal Veterinary and Agricultural University (DK); available online at: <http://www.models.kvl.dk/users/rasmus/brothesis.pdf>, Multi-way Analysis in the Food Industry, 1998, pp 50-60.
- Huang, J.; Wium, H.; Qvist, K. B.; Esbensen, K. H.; *Chemom. Intell. Lab. Syst.* **2003**, *66*, 141.
- Trefethan, L.; Bau, D. I.; *Numerical Linear Algebra*, SIAM Press: Philadelphia, 1997.
- Javaherian Naghash, H.; Dehghan, F.; Massah, A. R.; *e-polymers* **2008**, *53*, 1.
- Andersson, C. A.; Bro, R.; *Chemom. Intell. Lab. Syst.* **2000**, *52*, 1; <http://www.models.kvl.dk/source/nwaytoolbox/> (*The N-way Toolbox for Matlab*).
- Wold, S.; *Technometrics* **1978**, *20*, 397.

Received: May 21, 2009

Web Release Date: February 1, 2010