Identification of chemical components in XPS spectra and images using multivariate statistical analysis methods

Kateryna Artyushkova, Julia E. Fulghum*

Chemistry Department, Kent State University, Kent, OH 44242, USA

Received 20 December 2000; received in revised form 11 May 2001; accepted 29 May 2001

Abstract

A variety of data analysis methods can be used to enhance the information obtained from a measurement, or to simplify extraction of significant components from large data sets. Much work is needed to improve the quantification and interpretation of XPS spectra and images from complex organics. Multivariate analysis (MVA) is increasingly used for applications in electron spectroscopy to aid the analyst in interpreting the vast amount of information yielded by spectroscopic techniques. In general, the goals of MVA are to determine the number of components present, identify the chemical components, and quantify component concentrations in the mixture. Principal component analysis (PCA) is frequently used to determine the number of mathematical components which describe the data set. These mathematical components must then be related to chemically meaningful components. Various approaches to solve rotational ambiguities of spectral resolution, including local rank method (EFA), pure variables method (Simplisma) and multivariate curve resolution (MCR), are tested in the determination of chemical components from XPS data. Limitations associated with the resolution of a single matrix are shown to be partially or completely overcome when several related matrices are treated together. The test data sets contain XPS images or spectra acquired from blends of poly(vinyl chloride), PVC, and poly(methyl methacrylate), PMMA. The PVC degrades rapidly upon exposure to the X-ray beam. Spectra and images from the blend, acquired as a function of time, provide the multi-dimensional data sets for algorithm evaluation. In addition to spectral resolution, multivariate image analysis methods, such as principal component analysis, are used to extract maps of the pure components from an images-to-spectra data set.

Keywords: XPS; PVC/PMMA polymer blend; Multivariate data analysis; Multivariate curve resolution; Multivariate image analysis; Algorithm evaluation

1. Introduction

X-ray photoelectron spectroscopy (XPS) is widely used for the quantitative analysis of heterogeneous, multicomponent polymer samples. Although sample composition can be obtained from quantification of the survey spectra, or curve fits of the core level, high energy resolution photopeaks, quantification is not always straightforward. For polymer samples containing components with no differences in elemental composition, overlapping photopeak components frequently result in ambiguities in the number of photopeaks, the photopeak positions and other parameters used in the curve fit [1,2].
The recent improvements in imaging photoelectron spectroscopy enhance lateral and vertical characterization of heterogeneous samples, at the cost of increasing complexity in the XPS datasets which are acquired. In an images-to-spectra experiment, for example, images are acquired as a function of binding energy. The intensities from single pixels or groups of pixels can then be plotted as a function of energy, providing spectra from smaller areas than is possible using direct spectroscopic acquisitions. The resulting multi-spectral imaging data set is a complex data structure which requires more sophisticated analysis methods than visual inspection if the data are to be interpreted effectively [3]. Multivariate analysis (MVA) methods are increasingly being utilized in surface spectroscopies to aid the analyst in interpreting the vast amount of information resulting from multi-dimensional data set acquisitions. It has been suggested that this area of study be referred to as surface chemometrics [4].

MVA can be used to elucidate additional information from both spectra and images. Factor analysis based methods have been widely used in quantification of depth profiles from both destructive XPS depth profiling and angle resolved XPS [5–7], in resolving the distribution of oxidation states of metals and catalysts [1,8,9], and in studies of surface modifications, such as oxidation [10], film growth [10] and other chemical changes [11]. A self-modeling curve resolution (SMCR) technique was successfully applied to a sequence of XPS spectra acquired over the time during which a photoinduced chemical reaction was observed [12], and to resolve XPS spectra into surface and bulk component spectra based on an analysis of a family of spectra acquired at different polar angles from flat specimens [13].

Appropriate methods for multispectral image analysis have been developed for Earth satellite image processing and are directly applicable to surface microanalysis [14,15]. The application of factor and principal component analysis has been illustrated for the surface analysis of semiconducting, catalytic and magnetic structures [16,17].

XPS data sets are generally less complex than the hyperspectral data sets which can be obtained using TOF-SIMS or FTIR, making them useful for an evaluation of MVA methods. In this work, we use XPS data from a heterogeneous polymer blend in the evaluation of results from a variety of MVA approaches. Multivariate methods are tested by comparing the extracted information (number and identification of components) to the chemical information obtained directly from the XPS spectra or images, which provide a critical reference point. Spectra and images acquired from 50/50 PVC/PMMA blends as a function of X-ray exposure time provide the multidimensional data sets for algorithm evaluation. The PVC degrades upon prolonged exposure to the X-ray beam through a dehydrohalogenation process, in which carbon–chlorine bonds are replaced by double bonds, with a release of hydrogen chloride. The PVC degradation results in changes in peak intensities and positions in both survey and high resolution spectra [18–20].

This article will also demonstrate that additional information can result from the application of MVA methods, even when direct spectral or image interpretation is possible.

2. PVC/PMMA system

In order to be able to compare information extracted by MVA methods to the knowledge about the PVC/PMMA system, the degradation of PVC under X-ray exposure and the morphology of PVC/PMMA blend should be well understood. Halogenated polymers, especially chlorinated ones, are especially sensitive to X-ray exposure, resulting in a decrease in the halogen peak intensity and an increase in the C 1s peak intensity in photoelectron spectra [18–20]. A decrease in the rate of poly(vinyl chloride) (PVC) dehydrochlorination was observed at lower temperatures, suggesting that degradation of PVC is a combination of thermal and X-ray-induced processes [21]. The degradation process probably involves a series of free-radical reactions initiated by X-ray impact-induced and thermal bond cleavage accompanied by HCl release. The labile chlorine, when cleaved, attacks the methylene protons leading to dehydrochlorination by a free radical mechanism, followed by the formation of conjugated double bonds [22]. It has been reported that PMMA does not degrade via rearrangements, as PVC does. The PVC degradation leads to volatile species and/or changes in stoichiometry. PMMA is reported to
degrade via main-chain scission without changes in stoichiometry and the associated changes in quantification. Additionally, PMMA has been shown to degrade at a rate 2.5 times slower than PVC [23]. It has also been shown that the dehydrochlorination rates of PVC in complex multicomponent polymer samples depend on the method of blending, the solvent used and the affinity of the polymers [24,25]. The PVC/PMMA polymer blend phase separates with the formation of two phases: one enriched in PVC and the second enriched in PMMA, although both PMMA and PVC are present in each phase. The rate of degradation of PVC, thus, will potentially be affected by the presence of PMMA.

Polymer blends containing PVC have been extensively studied by XPS [26–33], as a result of both theoretical and practical interests. A variety of other methods have been used to characterize blends containing PVC and PMMA, leading to inconsistencies and conflicting conclusions regarding surface structure [26,28–30]. These inconsistencies are the consequence of different sampling areas utilized in the techniques, as the level of heterogeneity which is detected depends upon the analysis area of the characterization technique, and the number of areas/films which are analyzed. XPS analysis utilizing small sampling areas, and TOF-SIMS imaging of such films, demonstrated an enrichment of PMMA at the surface with extensive phase segregation [28,31,32]. SEM analysis revealed bulk phase separation extending to the outermost one micron. A comparison of XPS linescans with SEM micrographs of the film suggested either a thin film of PMMA on top of the PVC or smaller PVC domains contained within large PMMA domains [33]. Results obtained by Briggs et al. from TOF-SIMS studies also suggested the presence of a thin overlayer of PMMA [29]. A comparison of XPS and FTIR images and spectra has also demonstrated surface enrichment of PMMA [34]. This work focuses on identification of chemical components in the blend using MVA.

3. Theory behind multivariate data analysis

Multivariate analysis has been discussed in detail by Malinowski [35]. The discussion below summarizes the methods which can be used to obtain chemical components after mathematical component resolution by PCA or FA. The notation and terminology used throughout the article are explained in Table 1.

3.1. Spectroscopic resolution

Multivariate analysis (MVA) methods comprise a group of statistical, mathematical or graphical techniques that analyze multiple variables simultaneously [36]. MVA soft modeling methods aim at a fully automatic deconvolution of the original data matrix based on very few physical restrictions, such as non-negativity of concentrations or molar absorptivities, the validity of Beer–Lambert’s law, closure, etc. Soft modeling presumes that the chemical system is complex, and that it is not possible to adequately describe the behavior of the system by a hard model. Hard models are based on theories or laws that describe the behavior of the system, and development of such models is not always possible or practical. In soft modeling the system is described by variations and correlations in the data. In general, the goals of MVA analysis are to find correlations and variations in the data, to determine the number of components that describe this correlation, identify the chemical components and quantify their concentrations in the mixture.

Principal component analysis (PCA) and factor analysis (FA) are utilized to determine the number of significant mathematical components present in a mixture and to extract these components from large collections of spectra [35,36]. PCA transforms a number of (possibly) correlated original variables into a (smaller) number of uncorrelated variables called principal components. PCA uses a singular value decomposition (SVD) to decompose the original data matrix $D_{m \times n}$ formed from $m$ spectra and $n$ data points into a set of new variables, called latent variables or loadings, $V$, which are a linear combination of the original measurable variables:

$$D_{m \times n} = U_{m \times s} S_{s \times n} V_{n \times n}^T$$

where $U$ is the score which represents coordinates of the data matrix $D$ in a new coordination system. In spectroscopic terms, the loading, $V$, is a spectra-like vector, i.e. a plot of intensity vs. binding energy,
<table>
<thead>
<tr>
<th>Term</th>
<th>Name of term</th>
<th>Chemometrics definition</th>
<th>Analogy in spectroscopic terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General notation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td></td>
<td>The square of the standard deviation, ( \sigma ), or measure of the dispersion of a series of results around their mean, ( \mu )</td>
<td>In soft modeling the system is described by variations and correlations in the data</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>A measure of the relation between two or more variables; the degree of linear dependence between two vectors</td>
<td></td>
</tr>
<tr>
<td><strong>m</strong></td>
<td>No. of samples</td>
<td>No. of spectra or samples or positions on a sample</td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>No. of variables/channels</td>
<td>No. of data points in a spectrum</td>
<td></td>
</tr>
<tr>
<td>( D_{m \times n} )</td>
<td>Data matrix</td>
<td>Data is arranged in matrices where rows correspond to samples (m spectra) and the columns correspond to variables (n data points)</td>
<td></td>
</tr>
<tr>
<td><strong>k</strong></td>
<td>No. of correlated matrices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D_{m \times n \times k} )</td>
<td>Multi-way array</td>
<td>Consisting of k data sets of size ( m \times n )</td>
<td></td>
</tr>
<tr>
<td>( D_{m \times n \times k} )</td>
<td>Augmented two-way data set</td>
<td>Obtained by stacking a set of matrices column-wise</td>
<td></td>
</tr>
<tr>
<td><strong>PCA terminology,</strong> ( D_{m \times n} \leftarrow U_{m \times k} \Sigma_{k \times k} V_{k \times n}^T )</td>
<td>Principal component</td>
<td>New uncorrelated variable obtained by mathematical transformation of correlated original variables; orthogonal, maximum variance estimators of the data; each PC has ( V ) and ( U ) associated with it</td>
<td></td>
</tr>
<tr>
<td><strong>V</strong></td>
<td>Loading</td>
<td>Vector comprising a new coordinate system and consisting of a linear combination of original measurable variables</td>
<td>Spectral profile obtained by linear combination of original spectra; shows relationship between variables</td>
</tr>
<tr>
<td><strong>U</strong></td>
<td>Score</td>
<td>Coordinates of data matrix ( D ) in a new coordination system</td>
<td>Intensity profile — a plot of contribution of loading ( V ) to the original data ( D ) vs. No. of samples (spectra), ( m ); shows relationship between samples</td>
</tr>
<tr>
<td><strong>( \lambda )</strong></td>
<td>Eigenvalue</td>
<td>The measure of the magnitude or importance of a component; amount of variance captured by each loading/score pair is proportional to eigenvalue</td>
<td>Importance of each resolved spectral profile in a new coordinate system</td>
</tr>
</tbody>
</table>
Table 1. Continued

<table>
<thead>
<tr>
<th>Term</th>
<th>Name of term</th>
<th>Chemometrics definition</th>
<th>Analogy in spectroscopic terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Diagonal matrix whose elements are $\sqrt{\lambda}$ Factor</td>
<td>An element of a data reduction in which many measurements are described by a few independent variables; the number of factors will always be less than the number of original variables $n$ and will be equal to $r$</td>
<td>Number of new spectral/intensity profiles (loadings/score pairs) that can describe the most significant information in the original data set $D$ (in terms of variation and correlation) within experimental error $E$</td>
</tr>
<tr>
<td>$r$</td>
<td>Rank</td>
<td>Number of principal components (PCs) or factors that are required to reproduce the original data matrix $D$ within experimental error $E$</td>
<td></td>
</tr>
<tr>
<td>$E$</td>
<td>Model error</td>
<td>Difference between original data $D$ and data reproduced using $r$ number of factors</td>
<td>The goal is to explain all the signal but noise ($E$) by $r$ factors</td>
</tr>
<tr>
<td>$D = CS^T + E$</td>
<td>Real decomposition</td>
<td>The goal of real decomposition is to obtain chemically meaningful real decomposition of matrix $D$ into spectral and concentration profiles</td>
<td>Spectral profile — plot of intensity vs. variable (BE)</td>
</tr>
<tr>
<td>$S$</td>
<td>Resolved spectral profile (rotated $V$)</td>
<td>The goal of real decomposition is to obtain chemically meaningful real decomposition of matrix $D$ into spectral and concentration profiles</td>
<td></td>
</tr>
<tr>
<td>$C$</td>
<td>Resolved concentration profile (rotated $U$)</td>
<td></td>
<td>Intensity or concentration profile — plot of intensity vs time, $t$, or sample number</td>
</tr>
<tr>
<td>$T$</td>
<td>Rotation matrix</td>
<td>Mathematical transformation of loading/score pairs ($V/U$) to chemically meaningful decomposition ($S/C$)</td>
<td></td>
</tr>
<tr>
<td>Simplisma, $p_{ij} = \sigma_j / (\mu_j + \alpha)$</td>
<td>Pure variable</td>
<td>Variable at which only one component is present in the system</td>
<td>Binding energy at which only one component in the multicomponent mixture is present</td>
</tr>
<tr>
<td>$p_{ij}$</td>
<td>Purity</td>
<td>Ratio between standard deviation $\sigma_j$ and mean of each variable $\mu_j$</td>
<td>The purer the variable the less overlap between peaks in spectra; a larger standard deviation for a particular variable indicates a higher purity value</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Offset</td>
<td>Coefficient used to avoid attributing a high purity value to noise</td>
<td></td>
</tr>
</tbody>
</table>
while the score, $U$, is an intensity-like vector, i.e. a plot of intensity vs. sample number or time. Principal component analysis provides a unique solution, allowing for the original data to be reconstructed from the results. It analyzes the total variance among the variables, so the solution generated will include as many components as there are variables, although not all of them will meet the criteria for retention. $S$ is a diagonal matrix which contains the information related to the criteria for component retention. The diagonal elements are the square root of the eigenvalues $\lambda$ of the correlation matrix $Z = D^T D$. The first principal component accounts for as much of the variability in the data as possible and has a largest eigenvalue $\lambda$, and each succeeding component accounts for as much of the remaining variability as possible.

Common factor analysis, usually referred to as factor analysis (FA), uses an estimate of common variance among the original variables to generate the factor solution. Because of this, the number of factors will always be less than the number of original variables and will be equal to $r$:

$$D_{m \times n} = U_{m \times r} V_{r \times r}^T + E_{m \times n} \quad (2)$$

The abstract factors $U_{m \times r}$ in Eq. (2) are computed from $US$ in Eq. (1). The primary factors are those corresponding to the largest $r$ eigenvalues and represent the set of $r$ abstract factors that are required to reproduce the original data matrix $D$ within experimental error $E$. The remaining factors, each describing a low variance, represent the noise in the data set.

The determination of the rank of the data $r$ is a crucial step in the resolution of chemical components. The rank can be determined by various statistical criteria, as discussed below for each particular mathematical algorithm.

Loadings vectors, $V$, describe the variation in measurements rather than chemical phenomena, while the goal of the mixture analysis is the real decomposition of data to concentration, $C$, and spectral, $S$, profiles, i.e. $D = CS^T + E$, within experimental error, $E$. Mathematically resolved vectors $V$ and $U$ must, thus, be related to chemically meaningful components through a rotation matrix $T$:

$$D = UTT^{-1}V^T + E = CS^T + E, \quad C = UT, \quad S^T$$

$$= T^{-1}V^T \quad (3)$$

The art in self-modeling curve resolution (SMCR) methods is to find the correct or best rotation matrix $T$. There are numerous approaches for finding the rotation matrix $T$. Several of the most widely used methods will be tested in the resolution of the chemical components present in the XPS spectra from PVC/PMMA polymer blends.

1) Methods based on local rank. Evolving factor analysis (EFA) takes advantage of information in the time domain [37]. As a new component appears in the data during an evolutionary process, or over the data set from sample to sample, the presence of an additional eigenvalue is required to explain the variability. Thus the magnitude of the eigenvalues changes with time. Tracking the number of significant PCs in the soft model over the time axis permits estimation of the number of species that vary over the data set and whether different species enter into the system. Spectra acquired as a function of time vary from the blend, in which PVC degrades, are used for EFA.

2) Once a FA has been performed, another mathematical bridge between the abstract and real solutions is Target Testing. In target factor analysis (TFA), the test vectors representing the concentration or spectral profiles of the pure components can be evaluated by oblique rotation to assess whether the corresponding vector lies in the subspace spanned by the chosen primary abstract factors [38]. Statistical tests applied to the predicted and test vectors determine whether these two vectors are one and the same. These tests serve as a method for accepting or rejecting possible fundamental components of the sample. External knowledge of concentration and spectral profiles as targets is necessary. In this case, spectra from the pure polymers are included among the targets used to evaluate the performance of TFA.

3) In self-modeling methods the contribution of each chemical compound to a total overlapped curve is resolved by making some general physically meaningful assumptions about the underlying peak shape, for example non-negativity of spectral and intensity profiles. The Simple-to-use Interactive Self-modeling Mixture Analysis (Simplisma) [39] finds the pure variables in the data set and uses the intensities of the pure variables to estimate the concentration profiles of the analyte. A pure variable is a wavelength or energy, while a pure object is an analysis time, at which only one of the compounds is
present in the system. Resolution of components in the Simplisma method is based on purity which is a ratio between the standard deviation $\sigma_j$ and mean $\mu_j$ of each variable (spectrum):

$$p_{ij} = \sigma_j / (\mu_j + \alpha) \quad (4)$$

The first pure variable has the largest relative variation with respect to the time axis and is found by finding the variance, i.e. $\sigma_j^2$, that maximizes the purity $p_{ij}$. To avoid attributing a high purity value to a spectrum with low intensity, i.e. noise, the offset $\alpha$ is introduced in the denominator. Survey and high resolution spectra acquired as a function of time from the blend are used for Simplisma evaluation.

(4) Multivariate self-modeling curve resolution (MCR) refers to a group of techniques which recover the spectral and time or intensity profiles of components in an unresolved and unknown mixture obtained from an evolutionary process for which no prior information is available about the nature and composition of these mixtures [40]. The performance of the method is based on two resolution theorems, stating that (1) if all interfering compounds that appear inside the concentration window of a given analyte also appear outside the window, it is possible to calculate the concentration profile of the analyte, and (2) if for every interference the concentration window of the analyte has a subwindow where the interference is absent then it is possible to calculate the spectrum of the analyte.

MCR utilizing an alternating least squares (ALS) algorithm uses information from pure variables (Simplisma), local rank (EFA) and local physical constraints. It is based on a linear model law similar to Lambert Beers law:

$$D = CS^T + E \quad (5)$$

Starting with initial values of $C$ and $S$, an iterative algorithm computes new $C$ and $S$ ($S^T = C^*D$, $C = D(S^T)^*$) at constraints that are applied at each iteration. Thus, the method requires initial intensity or spectral profiles. Spectra acquired as a function of time from the blend, in which PVC degrades, are used for MCR.

(5) It is also possible to treat a set of correlated data matrices as one data set. $K$ different data sets of size $m \times n$ can be combined into (i) one multi-way array of size $m \times n \times k$ [41] or (ii) into an augmented two-way data set of size $(m \times k) \times n$ by column-wise setting each of the independent matrices on top of one another [42]. In parallel factor analysis (PARAFAC), elements $d_{m \times n \times k}$ of a three-way array $D$ are expressed as the sum of products of elements from factor loading matrices $A$, $B$ and $C$ within experimental error $E$:

$$d_{m \times n \times k} = \sum d_{mn} b_{nr} c_{kr} + e_{m \times n \times k} \quad (6)$$

An obvious advantage of the three-way PARAFAC model over the two-way factor analysis model is the uniqueness of the solution. For trilinear data, the underlying spectra will be found if the right number of components is used and the signal-to-noise ratio is appropriate [43]. The mathematical meaning of uniqueness is that the estimated PARAFAC model cannot be rotated without a loss of fit, as opposed to two-way analysis where one may rotate scores and loadings without changing the fit of the model.

Applying MCR to an augmented data matrix also allows for factor analysis rotational ambiguities to be solved completely for trilinear data and for non-trilinear data under physical constraints [42].

Data sets combining spectra acquired as a function of time from the blend and spectra from the pure polymers are used for multiway MCR and PARAFAC analyses.

3.2. Multivariate image analysis

Large image data sets can be analyzed using multivariate image analysis (MIA). MIA can be used for extracting the significant information and dividing this information into orthogonal components [44]. MIA methods, such as PCA, transform multivariate images into a number of factorial (or latent or principal component or score) images, which carry information related to these orthogonal components. The goal of MIA is to find images which are globally correlated or anticorrelated and to visualize the pixels which are responsible for the correlations. The correlation information can be displayed as scores of the different maps, while the pixel information is displayed as component images. This often simplifies data analysis as compared to studying the pure variable images.
4. Experimental

4.1. Sample preparation

Poly(vinyl chloride), PVC, of MW 77.3 kDa and poly(methyl methacrylate), PMMA, of MW 100 kDa were obtained from Scientific Polymer Products, Inc.

As-received materials were used in 2% w.v. solutions of PVC and PMMA in HPLC grade tetrahydrofluoran (THF). Solutions containing a 50/50 mixture of the two polymers were allowed to sit for 24 h before deposition onto teflon watch glasses using pipettes. The resulting films were air-dried for 24–48 h. The films were peeled from the teflon substrate before analysis by XPS. PVC and PMMA standards were prepared in a similar manner. The air-facing side of the samples was analyzed by XPS.

4.2. XPS Instrumentation

The XPS spectra and images were acquired on a Kratos AXIS Ultra photoelectron spectrometer using a monochromatic Al Kα source operating at 300 W. The base pressure was $2 \times 10^{-10}$ Torr, and operating pressure was $2 \times 10^{-9}$ Torr. Charge compensation was accomplished using low energy electrons. Standard operating conditions for good charge compensation are $-2.8$ V bias voltage, $-1.0$ V filament voltage and filament current of 2.1 A. Previous studies in our lab have shown that differential charging is not a problem for heterogeneous insulating samples [45].

4.3. Data acquisition

In order to evaluate a variety of different algorithms for mixture analysis, XPS data sets with evolution processes in time or energy were obtained. Spectral degradation experiments on both the pure polymers, PVC and PMMA, and blends were performed. The samples were exposed to monochromatic Al Kα X-rays for approximately 12 h. Survey and high resolution spectra were acquired continuously at pass energies of 80 and 20 eV with 1 and 0.1 eV step intervals, respectively, for 2 min for each spectrum. These experiments resulted in 41 spectra of each type, in which changes due to PVC degradation are detected.

Images are acquired as a function of binding energy in an images-to-spectra experiment. Small area spectra can be obtained from anywhere within the images by plotting image pixel intensity for a single pixel or a group of pixels versus binding energy. In this work, the images-to-spectra data set is used as a multivariate image data set for the evaluation of the performance of MIA methods. Medium magnification images were acquired over the binding energy range from 289 to 279 eV with a 0.2 eV step. This corresponds to a binding energy range from 291 to 281 eV after charge correction. A pass energy of 80 eV and acquisition time of 2 min per image was employed.

Prior to the images-to-spectra experiment, medium magnification O 1s, Cl 2p and C 1s images were acquired from the same area on the sample for 3 min each at a pass energy of 80 eV.

4.4. Data pretreatment

All spectral and image data files were transferred from the KRATOS format to a format supported by MATLAB [46] and ENVI [47], which is ASCII and a simple ‘flat binary file’, respectively. In order to compare the profiles resolved by MVA methods to conclusions based on direct interpretation of the photoelectron spectra, the spectral preprocessing prior to the multivariate analysis was consistent with processing usually employed in quantitative XPS. MVA utilizes the entire survey spectrum, and no background was subtracted. For evaluation of high resolution C 1s spectra, straight line backgrounds were subtracted prior to multivariate analysis. All high resolution spectra in a single data set were calibrated by the same shift, which was determined from the first spectrum in the set, so that the hydrocarbon peak is at 285.0 eV. This approach was found to provide a uniform calibration, avoiding the introduction of extraneous components into the data. This procedure also allows for the combination of multiple data sets into one, as required for multivariate and augmented analysis.

In general, smoothing makes peaks wider and changes the noise structure, complicating detection of pure components. Smoothing can also introduce
additional artifacts, making, for example, noise look like a peak. Hence, no smoothing was done prior to any regression analysis.

Intensity variations may also be the part of the changes occurring in the system during the degradation experiment, and normalization may destroy this valuable information. Therefore, no normalization was applied when a single data matrix was analyzed. For the analysis of combined three-way or augmented data matrices, normalization is required for the data analysis routines.

Bereton et al. have demonstrated that mean-centering of spectroscopic data can result in a decrease in the number of principal components, causing a loss of information [48]. Hence, for all the methods applied, no centering or scaling of data was utilized.

The high resolution C 1s spectra were calibrated and background subtracted in SpXZeigR 3.0.6r prior to MVA analysis [49]. Survey spectra of blends and pure polymers were combined into separate data matrices of size 41 (spectra) × 383 (data points), and high resolution C 1s spectra were combined into matrices which are 41 (spectra) × 71 (datapoints). The images-to-spectra data set was also combined in a separate three-way matrix of size 256 (pixels) × 256 (pixels) × 50 (matrices).

4.5. Programs used

All routines for spectral analysis were run in the MATLAB environment, while image analysis was performed using both MATLAB [46] and ENVI [47].

PCA and EFA routines are available in the PLS_Toolbox 2.0.4 [50]. The TFA routine is taken from Malinowsky and modified to simultaneously resolve both profiles [35]. A demonstration version of the SIMPLISMA algorithm is taken from the report by Windig [39]. The routine for MCR is available for free download from http://www.ub.es/gesq/mcr/mcr.htm. The PARAFAC routine is provided by free download by Chemometrics Group, Food Technology at http://www.models.kvl.dk/source/. Image processing methods are applied in ENVI. An additional image PCA routine is applied in PLS_Toolbox in MATLAB to extract spectral profiles/loadings from images-to-spectra data.

5. Results and discussion

5.1. Analysis of blend composition based on spectroscopy

For the survey and high resolution C 1s degradation data sets, the goal is to resolve two or three components that represent the constituents of the blend: PMMA, PVC and the hydrocarbon (HC) which is a product of PVC dehydrochlorination. Representative degradation survey spectra from PMMA, PVC and blend are shown in Fig. 1a. Both polymers contribute to the C 1s peak in the survey spectra, but information from other peaks correlate with the C 1s peak for both polymers, Cl 2p for PVC and O 1s for PMMA, which simplifies the resolution of components. This data set is thus viewed as a simple test case in the initial evaluation of the methods. Two component spectra are expected to be resolved from the survey data set: one containing peaks from O and C, the other from Cl and C. During the degradation a third component representing HC generation appears, so it is reasonable to expect a third component to be resolved from the data.

The high resolution C 1s spectrum is a convolution of overlapping components from PVC and PMMA. Representative C 1s degradation spectra for the blend and pure polymers are shown in Fig. 1b. During X-ray exposure, the shape of the PVC peak changes dramatically and becomes similar to that of PMMA. This complicates curve fitting, since curve shapes based on PMMA and PVC standards cannot be used to fit the C 1s photopeak. Resolution of components from the C 1s photopeak is thus a more challenging test of the multivariate techniques, since the chemical components from both polymers are similar. Ideally, we expect to resolve three chemical components — PVC, PMMA and a HC degradation product.

The scheme behind the soft modeling of the polymer blend spectral data and order of procedures applied are shown in Fig. 2. Data matrices containing survey and high resolution C 1s degradation spectra are analyzed following this scheme, and profiles of components resolved from the survey and high resolution degradation data sets are evaluated by comparing them to the survey and high resolution
spectra from pure PMMA and PVC (first spectrum at the beginning of degradation experiments shown in Fig. 1).

5.1.1. Spectral resolution

The number of independent components in the original data set \( D \), i.e. the rank of the data \( r \), is determined in PCA by two criteria. First, the knee in the plot of total variance captured by eigenvalues versus the number of eigenvalue, called a scree plot, indicates the number of components in the model. Alternatively, the original data matrix is reproduced with progressively more factors included until it is reconstructed within experimental error. Sufficient PCs have been included when the plot of the residuals (the difference between the original and reproduced data) shows random noise rather than spectra-like features, as illustrated in Fig. 3. Although the data reconstructed using one component (Fig. 3b) look almost identical to the original data (Fig. 3a), the residuals, which constitute 10% of the data, have a spectrum-like shape, indicating that the data are not modeled correctly. The residuals for the two-component model (Fig. 3c) represent noise in the data, indicating that this particular data set can be described by two principal components.

PCA applied to both the survey and high resolution C 1s spectral data sets shows that there are two components present in both data sets. The loadings (spectral profiles) and scores (intensity profiles) of these two resolved components are presented in Fig. 4. The shape of the first loading
from the survey data set (Fig. 4a) resembles the spectrum from the blend, while the second loading has a significant peak which can be assigned to HC along with small spikes indicating a decrease in Cl 2p, O 1s and C1s (at 285 eV). This suggests a two-component system containing the blend itself and the degradation product, which is mainly HC. The first resolved spectral profile from the high resolution C 1s data set (Fig. 4b) resembles the shape of the C1s spectrum from the blend. The second loading for this data set has negative parts around the peaks attributed to three components from PMMA (O–C=O at 288.2 eV, C–O at 286.2 eV and C–C–O at 285.7 eV) and a positive peak at the binding energy corresponding to the HC contribution. The second component thus corresponds to the degra-
Fig. 3. The number of components in PCA is determined by reproducing the original data with progressively more components and checking the residual shape. (a) Original data, (b) reproduced data with one component (upper) and residual (lower) and (c) reproduced data with two components (upper) and residual (lower). Note that residuals change shape from spectra-like for the model with one component to noise-like for the model with two components.

dation product in which the signal from the HC high resolution C 1s data set (Fig. 5c) is a broad correlates with the signal from the oxygen-containing part of PMMA. The identification of both components is also confirmed by the scores extracted from both survey and high resolution data sets. The first component is decreasing and the second component is increasing with time, as a result of the decreasing contribution from the blend and increasing contribution from the degradation product. These two components have contributions from both PVC and PMMA, and do not represent the pure polymers. EFA applied to the original data matrix $D_{m \times n}$ resolves only one set of $r$ profiles at a time, i.e. intensity profiles with dimensions $m \times r$. To resolve a profile in the second order, the spectral profiles of size $n \times r$, the data matrix is transposed to $D_{n \times m}$ and EFA is reapplied. Intensity profiles resolved by EFA from both data sets are similar, showing one component decreasing while the second is increasing with time (Fig. 5). The spectral profiles obtained from the survey data set are uninterpretable (Fig. 5a and b). The first spectral profile resolved from the high resolution C 1s data set (Fig. 5c) is a broad peak covering the binding energy range from 285 to 289 eV, which corresponds to the high resolution C 1s spectrum. The second profile contains a peak around 284 eV, which corresponds to the position of the HC component in the C 1s spectrum (Fig. 5c). If three components are utilized, identification is even more challenging, as demonstrated by the intensity and spectral profiles shown in Fig. 5b and d. The EFA method is based on tracking the number of significant components in the data over time, as some components leave or enter the system, as in the case of a chromatographic experiment. In the degradation XPS experiments, the changes that are occurring in the system are not as significant, and the rank of the data does not change with time, as no components are entering or leaving the system. Therefore, local rank based methods do not perform as well for these test data sets. Initial estimations of intensity profiles are still easily obtained using EFA with no external knowledge about the system. Three types of test target vectors are used in TFA:
(a) randomly generated vectors, (b) loadings and scores resolved by PCA and (c) spectral profiles from pure PMMA and PVC. These test vectors are iteratively rotated to resolve concentration and spectral profiles of pure components from the data matrix $D$. TFA using randomly generated intensity and spectral profiles as test vectors does not resolve meaningful components for either the survey or high resolution C 1s data sets (Fig. 6a). When loadings and scores obtained by PCA are used as targets, the TFA results do not improve the chemical component resolution (Fig. 6b). If survey spectra from pure PVC and PMMA are used as targets, TFA resolves two similar, but not identical, components (Fig. 6c(i)). When high resolution C 1s spectra from pure PVC and PMMA are used in TFA, two identical spectral profiles are obtained (Fig. 6c(ii)). Iterative refining of these profiles does not improve the results. Pure components are frequently unavailable in the analysis of heterogeneous samples, and this method is not applicable in these cases. TFA has been successfully used for resolving chemical components from XPS depth profiles [5,51], but this method was not successful in the analysis of our test data. Possible reasons for this failure include the gradual changes in component shape with X-ray exposure time, the significant overlap between the pure components and the complex degradation behavior of PVC within the blend. The degradation behavior is discussed in more detail in conjunction with image analysis.

Fig. 4. PCA model applied for (a) survey and (b) high resolution data sets. Loadings and scores for two-component models are shown. Note similar scores for both models. The shape of the loadings is discussed in the text.
Fig. 5. EFA model applied to survey (a) using two factors and (b) three factors, and to the high resolution data set (c) using two factors and (d) three factors. Intensity and spectral profiles are shown. Note similar intensity profiles. Identification is difficult due to spectral shape.

Simplisma is an interactive analysis method, in which the user decides both whether or not to select a pure spectrum and on the size of the offset, $\alpha$. The offset is required in order to avoid attributing a high purity value to the noise spectrum (Eq. (3)). Several different criteria can be used to estimate the number of components in Simplisma. If too many components are included, degenerate peaks, or peaks located outside of a reasonable binding energy range, are observed since a high purity value has been assigned to noise components. In addition, the quality of the resolution has a significant dependence on the number of components chosen. The model with the correct number of components has better resolved profiles than a model with too few or too many components. The resolved profiles of extra components are generally noisy. Increasing the offset does not influence the shape of spectral or intensity profiles, but it allows the user to obtain a model with resolved profiles, containing pure peak at reasonable positions, without negative intensities. If the spectra to be resolved contain overlapping components such that pure variables, or energies at which only one of the compounds is present in the system, do not exist for every component of the system, Simplisma is difficult to apply for resolution of the components directly from the data. This is clearly the case in the analysis of the spectra from the PVC/PMMA blend. For cases like this, taking the second inverse derivative before the application of Simplisma enhances the resolution and facilitates the detection of pure variables [39].

Intensity profiles resolved from survey and high resolution C 1s spectral data using low offset values ($\alpha = 1-3$) had negative regions. Increasing the offset value to $\alpha = 5$ results in the resolution of two components from both data sets (Fig. 7). The first spectral profile in the survey data set increases with time, and has peaks from HC and from O 1s (Fig. 7a). The second spectral profile from the same data
has a shape resembling the survey spectrum of the blend, and the score decreases with time. It is thus possible to chemically identify the first component as a product of polymer degradation and the second as the blend. It is interesting to note that HC, which is produced during PVC dehydrochlorination, is correlated with the oxygen signal from PMMA. The resolution of two components from the high resolution C 1s spectra (Fig. 7b) is slightly poorer than from the survey data (Fig. 7a). The first profile is similar to the last spectrum in the degradation experiment (Fig. 1b, last C 1s spectrum), which is a
blend spectrum with a high HC content. The second profile resembles the spectrum of the blend before degradation (Fig. 1b, first C 1s spectrum). The intensity profiles resolved from the high resolution spectral data set are different from that for the survey data set, probably as a result of the different degree of component resolution. It is important to note the negative values in the intensity profiles resolved from the high resolution spectral data.

MCR uses initial estimations of intensity and spectral profiles obtained from local rank analysis (EFA) or pure variable methods (Simplisma) to resolve the components from the data matrix $D$. Non-negativity constraints are applied in both spectral and intensity orders. Results from the application of MCR to the survey data depend upon whether initial profiles from EFA or Simplisma are utilized (Fig. 8a and b). Initial profiles from EFA resolve two almost identical components (Fig. 8a) while initial profiles from Simplisma result in more distinct components (Fig. 8b). The first is a HC component which is correlated with an O peak and the second is representative of the blend. Although MCR has the advantage of including non-negativity constraints, the resolved profiles are not a significant improvement over the Simplisma results shown in Fig. 7a.

Analysis of the high resolution C 1s spectral data using MCR gives similar results for both sets of initial profiles from EFA and Simplisma (Fig. 8c and d). Non-negativity constraints applied at each iteration force all of the negative values present in the intensity profiles resolved by Simplisma (Fig. 7b) to be positive and, therefore, provide a small improvement in the resolution of the spectral profiles (Fig.
Chemically, one of the spectral profiles resembles a blend with a high PVC content before degradation, while the second resembles the blend after degradation.

Applying any of the methods discussed above, PCA, EFA, TFA, Simplisma or MCR, to the survey and high resolution data sets from the PVC/PMMA blend thus resolves components representing combinations of the pure polymers and degradation product, and not the pure PMMA and PVC.

The last method for solving the rotational ambiguity problem existing in two-way factor analysis is to combine related matrices into one multi-way or augmented data set. Two-way data matrices of survey or high resolution C 1s spectra from pure PVC, pure PMMA and the blend were combined and analyzed by Simplisma, MCR-ALS and PARAFAC methods.

The determination of the number of components present in the data set depends on the method of data
set construction. If several matrices $D_{m \times n}$ are augmented into a single two-way matrix $D_{(m+n) \times k}$, then PCA and Simplisma can be applied to this matrix to determine the number of components. For multiway data sets, the data set is first collapsed. For a data matrix with dimensions $m \times n \times k$, the new collapsed matrix will be

$$m \times n = m \times n \times 1 + m \times n \times 2 \cdots + m \times n \times k$$

Next, the collapsed matrix is analyzed by PCA to determine the rank of the data.

PCA applied to both the collapsed and augmented matrices from the survey and high resolution data sets indicates three significant components in the data. This result means that the correlation and variation in the data sets combining spectral data from the blend and the two pure polymers requires one additional component in order to be adequately described.

Components that can be identified as PVC, PMMA, and a HC degradation product result from the application of Simplisma to the augmented data set of survey spectra (Fig. 9a(i)). The MCR-ALS algorithm is used with non-negativity constraints introduced in both spectral and concentration orders to resolve three components. Spectral profiles resolved by MCR with initial profiles from EFA (Fig. 9a(ii)) are similar to those resolved by Simplisma (Fig. 9a(i)), but contain small negative spikes. Non-negativity constraints are also applied during PARAFAC application. The algorithm was run several times, using random initialization, to check that the solution converges to the global minimum. The profiles resolved by the PARAFAC method were the

![Fig. 9. Multiway analysis. Three resolved spectral profiles are shown for (a) survey and (b) high resolution data sets using (i) Simplisma, (ii) MCR using initial profiles from EFA and (iii) PARAFAC. The three methods resolve identical pure components: pure PVC, pure PMMA and HC, as a product of PVC degradation.](image-url)
same for all random runs, although the number of iterations and percent of variance described were slightly different. PARAFAC resolves the same three components as Simplisma and MCR, corresponding to pure polymers and the HC degradation product (Fig. 9a(iii)). Unfortunately, the performance of this version of the PARAFAC algorithm [52] is much slower than MCR-ALS and Simplisma, making the application of several random initializations for checking the convergence to a global minimum impractical as a routine approach.

The application of Simplisma to the augmented data set of high resolution C 1s spectra also resolves three spectral profiles (Fig. 9b(i)). The first can be chemically identified as PMMA, the second is similar to PVC and the third is HC with a contribution from the blend. Negative values in the intensity profiles (not shown) suggest that this set of resolved profiles is not a chemically meaningful solution of the rotational ambiguity problem. It is difficult to achieve convergence to a solution containing profiles with only positive values using MCR since the initial estimates from Simplisma contain negative values. MCR applied to the same data set using initial profiles from EFA combined with non-negativity constraints results in a more chemically meaningful solution. The three spectral profiles which are resolved are similar to the results from the analysis of the survey data set (Fig. 9b(ii)). In this case, the third component is more obviously a PVC degradation product, as contrasted to the results obtained by Simplisma (Fig. 9(i)). Results from PARAFAC are similar to those from MCR and Simplisma (Fig. 9a(iii)), but algorithm convergence is slower. Combining several related matrices into one data set thus allows for resolution of the three chemical components which are present in the degraded polymer blend: PMMA, PVC and HC.

For the survey data set, Simplisma and MCR were able to resolve two components successfully — one corresponding to the blend and another to a degradation product in which the HC signal is correlated with O from PMMA. Including the standards in the multi-way or augmented data sets allows for the resolution of three components: the pure polymers, PMMA and PVC, and the HC degradation product. Simplisma provided the best results from the analysis of the survey augmented data set, as it gave the least ambiguous chemical component resolution in the shortest time with no application of external knowledge. Survey spectra contain narrow peaks, providing a data set which is well suited for this pure variable method.

For the high resolution C 1s spectra, however, the components were not well resolved from the blend data set by either method. Combining the data set from the blend with those from standards assisted in the resolution of three components. The MCR method using initial profiles from EFA gave the best performance in the shortest time. This approach requires initial estimations which can be easily obtained using local rank methods, but sometimes it can fail to converge. This was the case if intensity profiles from Simplisma, which contain negative values, were utilized.

5.2. Multivariate image analysis

The ability of MIA methods to identify clusters of pixels representing a single phase are tested by the analysis of the images-to-spectra C 1s data set to extract images of components and their spectral profiles. A correlation between the HC component and the O-containing component from PMMA was observed based on analysis of the spectral degradation data set. The MIA analysis was utilized to provide additional insight into this correlation.

Initially, the number of components present in the data sets is determined, and component images are extracted, using the ENVI PCA routine. The corresponding spectral profiles are resolved in the im agepca routine in PLS Toolbox 2.

Cl 2p, O 1s and C 1s images acquired from the same area on the sample at 198.0, 536.0 and 283.0 eV, respectively, are shown in Fig. 10. The Cl 2p image, in which bright areas are enriched in PVC, shows the highest contrast, and contains a large circular feature in the center of the image. The O 1s image, in which bright areas are enriched in PMMA, has lower contrast but still shows a good anticorrelation with the Cl 2p image. The C 1s image is from the hydrocarbon component, to which both polymers contribute, resulting in a featureless image. The spatial distribution of phases cannot be ascertained from this image.

PCA was applied to the multispectral images-to-
Fig. 10. Photoelectron Cl 2p (a), O 1s (b) and C 1s (c) images acquired from the same area on the sample.

Fig. 11. PCA results applied to image-to-spectra data set. Upper — scores; lower — loadings for left — first component and right — second component.
spectra data set, acquired at conditions described in the Experimental section. For this application, PCA results in component maps (scores) and corresponding spectral profiles (loadings), as shown in Fig. 11. It is possible to identify components based on the visual inspection of principal component images. For this purpose we use elemental C 1s, O 1s and Cl 2p photoelectron images (Fig. 10) as the reference for comparison. The principal component map of the first component is similar to the C 1s image acquired at 285 eV (Fig. 10c), in which peaks from both polymers overlap. There is also a pronounced feature in the map of the first component which is similar to the PVC-enriched area in the center of the Cl 2p image (Fig. 10a). Data for the images-to-spectra experiment took approximately 12 h to acquire for this particular system. During this time, PVC degrades through the dehydrochlorination mechanism described in the Introduction. The product of PVC degradation, hydrocarbon (HC), will contribute to the binding energy range of the C 1s peak. In C 1s images, therefore, the HC contribution will be most pronounced in the areas enriched with PVC, which is a bright circular feature in the Cl 2p image. Based on this comparison, the first extracted component can be identified as resulting from the 50/50 blend after pronounced PVC degradation. The map of the second extracted component is similar to the O 1s image (Fig. 10b) or an inverse of the Cl 2p image. Unambiguous identification of this component based on PC image appearance only is not possible. The spectral profiles (loadings) extracted by PCA (Fig. 11) allow for the identification of both components based upon comparison of their shape to high resolution C 1s spectra of the pure polymers, PVC and PMMA (Fig. 1). The shape of the first extracted loading, similar to the C 1s spectrum of the blend with a high HC content from degradation, is in good agreement with the identification of the first component from principal component image inspection. The loading of the second component is negative with respect to baseline peaks, so a chemical identification is not justified. Mathematically, though, two peaks are observed in the second spectral profile at approximately 284 and 288.5 eV, which can be attributed to HC, as a product of PVC degradation, and O−C=O from PMMA, respectively. The contribution from PMMA to the second loading may explain the similarity of the second component image to the O 1s image. An unexpected correlation between a product of the PVC dehydrochlorination, HC, and the O 1s signal from PMMA is observed, as in the analysis of the spectroscopic data. The authors initially expected that a component representative of the HC would be extracted by PCA, as the concentration of the degradation product, i.e. HC, is increasing with time in the blend system.

The combination of multivariate mixture analysis techniques applied to the XPS images and spectra indicates that the degradation behavior of the PVC within the blend is more complex than was expected. As discussed in the Introduction, a dependence of the rate of PVC degradation on a second polymer has been observed previously [24,25,53]. In a blend of PVC with PMMA, free chlorine radicals generated during PVC dehydrochlorination can react with adjacent C=O bonds. There is evidence that in a blend of PVC with PMMA, low temperature degradation of the second polymer is promoted by Cl radicals [54]. In addition, released HCl may be captured by PMMA moieties, and reaction may occur in which HCl converts some ester side groups into anhydride rings [53]. The product of blend degradation under X-ray exposure is, thus, not only hydrocarbons generated during dehydrochlorination of PVC, but also products of PMMA degradation, promoted by intermediates of PVC dehydrochlorination. Degradation of PMMA would occur first of all in PVC-enriched areas, where its promotion is the strongest. With this additional information, identification of the two components present in the blend as determined by MIA is as follows: (1) the 50/50 blend with high HC content due to degradation of PVC, and (2) a degradation product, resembling the image of Cl 2p, representing PVC correlated with oxygen, representing PMMA. It is apparent that without application of MVA to the spectra and images, the degradation behavior of this polymer blend would not be understood completely. A more detailed discussion of these phenomena has been published [55].

6. Conclusions

This study has evaluated a number of different MVA algorithms that can be applied to XPS spectra
and images to enumerate the species existing in a multicomponent system and resolve chemically meaningful components. The performance of multivariate methods in the resolution of chemically meaningful pure components was tested by comparing the extracted information to the chemical information obtained from direct interpretation of the XPS data.

PCA was used to resolve mathematical components and provide an initial identification. None of the methods used to rotate the mathematical components to a chemically meaningful solution resolve the chemical components representing the pure polymers present in the blend when single data sets containing spectra varying with time or energy are analyzed. For example, although the system does vary with time, the changes in spectra associated with polymer degradation are not significant enough to enable component resolution using evolving factor analysis. Target factor analysis also fails in resolving components as reasonable target vectors could not be selected for this system. Pure polymers could not be used as target vectors as a result of the complex PVC degradation behavior.

Simplisma resolved chemically meaningful components from the survey data set. Survey spectra contain narrow peaks, providing a data set which is well suited for this pure variable method. For this data set, it provided chemical component resolution in the shortest time with no application of external knowledge. The resolution of components from high resolution C 1s data sets which contain overlapping features is not as good, and the intensity profiles which are recovered contain negative values.

Chemically meaningful components could be obtained from both the survey and high resolution C 1s data sets using an MCR method with non-negativity constraints and initial profiles from local rank or pure variable methods. In all cases, however, the components which are resolved represent combinations of the pure polymers and the degradation product, rather than pure PVC or PMMA.

Combining several related matrices into one data set for Simplisma, MCR and PARAFAC, allowed for the resolution of three chemical components, i.e. PVC, PMMA and the HC degradation product. The MCR method using initial profiles from EFA gave the best resolution in the shortest time. Unfortu-

nately, this approach can fail to converge, as was the case if intensity profiles from Simplisma, which contain negative values, were utilized. Results from PARAFAC are similar to those from MCR and Simplisma, but algorithm convergence is slower.

Multivariate image analysis for extraction of morphological information from complex image data sets was tested on XPS images. It was possible, based on the PCA analysis, to elucidate the existence of phase-separated regions from a set of images acquired at constant binding energy without any additional information or experiments.

The PVC/PMMA blend was initially chosen as the test system for the evaluation of various algorithms since the authors considered it to be a well-studied system. The combination of multivariate mixture analysis techniques applied to the XPS images and spectra indicates that the degradation behavior of PVC within the blend is more complex than was expected. The product of blend degradation under X-ray exposure is not only hydrocarbons generated during dehydrochlorination of PVC, but also products of PMMA degradation, promoted by intermediates of the PVC dehydrochlorination. Without application of MVA to the spectra and images, the degradation behavior of this polymer blend would not be understood completely. This study thus also demonstrates that the appropriate application of surface chemometrics can extract ‘hidden’ information from complex data.

Acknowledgements

This work was partially supported by NSF ALCOM (DMR89-20147). The XPS was funded by a grant from the Keck Foundation and by NSF CHE-9613880.

References

[50] PLS Toolbox 2.0, Eigenvector Research, Inc., Manson, WA.