Wednesday Morning, October 22, 2008

Applied Surface Science
Room: 207 - Session AS-WeM

Advanced Data Analysis for Surface Characterization
Moderator: S. Pachuta, 3M

8:00am **AS-WeM1 Near Real-time Analysis of XPS Data**, *A.S. Lea, D.R. Sisk, M.H. Engelhard,* Pacific Northwest National Laboratory*, J.E. Castle,* University of Surrey, UK*, D.R. Baer*, Pacific Northwest National Laboratory

The application of x-ray photoelectron spectroscopy (XPS) to analyze different types of materials appears to be growing rapidly around the world. In our Department of Energy National User Facility the Environmental Molecular Sciences Laboratory (EMSL) we currently have user requests for thousands of hours per year of XPS based experiments and more routine sample analysis. The real-time (or near real-time) analysis of the XPS data as it is collected has significant potential advantages to scientists and instrument operators in that it has the potential to qualitatively alter the way experiments are done. Such endeavors have the potential, not only to improve the turn-around time for data analysis and the sophistication of data analysis reportable to the User, but also to reduce the labor involved in data analysis, resulting in significant time (and cost) savings. We are developing a prototypical real–time data analysis capability to analyze x-ray photoelectron spectroscopy (XPS) data generated in EMSL. Our goals are to provide atomic composition, carbon contamination layer thickness and component segregation (layer) information as soon as adequate data is collected on a specimen. Because of the high user need and previous efforts to design an XPS expert system, XPS is a good candidate for development of a near real time data analysis capability. Our real-time data analysis package follows many elements of the expert system approach proposed by Castle.[1] A set of rules and algorithms are used to address a well defined series of analysis objectives (based on the needs of the XPS analysis) to characterize the surface in terms of atomic concentration, layer sequences, and enrichment/depletion of elements as a function of depth. For each goal, rules have been developed to determine whether the stated goals are met. The real-time analysis package evaluates the rules by automated interpretation of the wide-scan (survey) spectra. Once the goals have been met, the real-time analysis package generates a report that includes, in addition to the above stated goals: a first level approximation of surface composition and metadata associated with this analysis, which will be stored with the datafile produced by the data analysis package.

[1] Castle, J.E., J. Vac. Sci. Technol. A (2007) 25, 1-27.

8:20am **AS-WeM2 Data Scaling for Quantitative Imaging XPS**, *J. Walton,* The University of Manchester, UK*, N. Fairley,* Casa Software

Multivariate analytical techniques are seeing increasing use in surface analysis due to the ability of current instrumentation to acquire multispectral data sets. Their use can provide a significant improvement in signal/noise, and simplify analysis of the large amount of data present by reducing its dimensionality. In the case of XPS this enables quantification by measurement of photoelectron peak areas, and chemical state determination using curve fitting to the spectrum at every pixel in an image. Selection of the most appropriate technique is dependent upon the characteristics of the data, and for XPS, where there are relatively few components compared with the number of objects in the data set, Non-linear Iterative Partial Least Squares (NIPALS) affords a significant saving in computational requirements, as the procedure may be terminated after the appropriate number of components has been calculated. A key aspect in the use of these techniques is their ability to order the data so that the chemical information is easily separated from the noise. For data acquired by pulse counting, which is governed by Poisssonian statistics and where the variance scales as the data, ordering the data by variance may result in noise from high intensity photoelectron peaks dominating low intensity chemical information. Effective separation of noise and chemical information therefore necessitates prescaling the data, so that the noise is evenly distributed. Individual objects, whether images or spectra, may be scaled by the square root of the variance. This provides good separation of the chemical information from the noise, but leads to a signal level below which information cannot be extracted. This is not important where the inelastic background is greater than this base level, but leads to significant errors in quantification where the background is equal to, or below this level, and data is lost from photoelectron peaks. This is particularly important for photoelectron peaks occurring at low binding energy. It will be shown that scaling to the square root of the mean variance in both image and spectral domains, known as optimal scaling, avoids this limitation and allows use of the NIPALS procedure for fully quantitative imaging XPS.

8:40am **AS-WeM3 Using Multivariate Analysis and Modeling for Structure-to-Property Relationships Built from XPS Data**, *K. Artyushkova, J.E. Fulghum, P. Atanassov*, The University of New Mexico
INVITED

This talk will discuss new approaches in accelerating material development and design by building models describing structure-to-property relationship based on X-ray Photoelectron data. The science of designing of heterogeneous materials has benefited from an understanding of the chemical, surface and materials phenomena at the nanoscale. Among the more complex nano-structured functional materials that will be discussed in this talk are electrocatalysts and biocatalysts. Understanding the structure of catalysts, and linking this structure to performance is essential for identification of the active catalytic sites, for optimization of catalyst performance, and elucidation of failure mechanisms. XPS is one of the most widely utilized surface spectroscopic techniques for analysis of catalyst structure. The ability to discriminate between different surface oxidation states and chemical environments is one of the primary advantages of the use of XPS in the characterization of catalyst structures. It is critical that the XPS spectra are interpreted and quantified with a high confidence level, as this information will be a central link between structure and performance. Although, the majority of XPS analyses of catalysts are focused on identifying the oxidation state and overall speciation, the ambiguity in peak assignment from overlapping peak components in XPS spectra is still a significant problem. Multivariate statistical methods of data analysis (MVA) are of critical importance in developing unambiguous methods of XPS data interpretation. Correlation of XPS structural data to any other property, such as derived from BET porosity, microscopic images and performance characteristics, represents a multivariate problem. Initially, Principal Component Analysis and Correlation maps will be used to study qualitative correlations between amounts of chemical species detected by XPS and variety of relevant for particular system macroscopic properties such as surface area, pore size distribution, electrochemical performance, corrosion rate, etc. In order to learn about relationship between several independent variables and a dependent variable and to determine the magnitude of those relationships, a variety of Regression Models are widely used. Multiple Linear Regression along with Genetic Algorithm for Variable selection will be discussed in attempt to build a predictive model between XPS, macroscopic parameters and performance characteristics.

9:20am **AS-WeM5 Identification and Quantification of ToF-SIMS Images with Topography using Multivariate Analysis**, *J.L.S. Lee, I.S. Gilmore,* National Physical Laboratory, UK*, I.W. Fletcher,* Intertek MSG, UK*, M.P. Seah*, National Physical Laboratory, UK

Surface topography is a crucial issue for the analysis of many innovative devices such as microfluidic systems, fibres, composite materials, sensors, organic electronics and biomedical devices. The strength and durability of these components is critically dependent on their nanoscale surface chemistry and molecular interactions. However, quantitative characterisation of surfaces with topography remains a significant challenge due to the lack of systematic and validated measurement and data analysis methods. Previously, we presented a systematic study of the effects of surface topography on ToF-SIMS and provided guidance to practical analysts for identifying and reducing topographical effects.[1] Here, we investigate the robust use of multivariate methods for the identification and quantification of ToF-SIMS images with surface topography using principal component analysis (PCA) and multivariate curve resolution (MCR). Multivariate analysis simplifies the description of data and is powerful for identifying trends and highlighting chemically significant areas on images. However, many challenges remain with its application to complex images obtained in practical analysis, especially where sample topography or detector saturation[2] gives rise to large non-linear intensity variations in the data. In this study, we use several model samples, including polymer fibres with multi-organic coatings, natural starch grains and human hair, to investigate the merits of different multivariate analysis strategies for samples with topography. The emphasis is placed on the accurate identification and quantification of surface chemistry using careful application of multivariate methods, combined with suitable data selection and preprocessing and valid interpretation of the results. This study extends from our previous work on flat samples[3] and provides helpful guidance in the rapid, unbiased analysis of high-resolution raw spectral data in ToF-SIMS images of increasingly complex multi-organic surfaces and biomaterials.

[1] J L S Lee, I S Gilmore and M P Seah, Appl. Surf. Sci. in press
[2] M R Keenan, V S Smentkowski, J A Ohlhausen and P G Kotula, Surf. Interface Anal. 40 (2008) 97-106
[3] J L S Lee, I S Gilmore and M P Seah, Surf. Interface Anal. 40 (2008) 1-14.

9:40am AS-WeM6 Methods for Reducing Topographical Influences in ToF-SIMS Spectral Images, *J.A. Ohlhausen*, *M.R. Keenan*, Sandia National Laboratories

Time-of-Flight Secondary Ion Mass Spectrometry is used to analyze surfaces and surface volumes of many kinds of materials. Many times the surfaces are rough in texture and contain variable materials compositions, thus affecting secondary ion responses. Data interpretation can be influenced unknowingly by topographical differences that might be present. In 3D analysis, data are typically represented as idealized cubes. Actual data will vary from the ideal cube in some unknown way. Spatial distortion from the idealized cube can be caused by several events including starting topography and differential sputtering rates. By understanding morphological and sputtering deviations, a 3D profile can be adjusted to reflect the actual shape of the removed volume. Strategies for measuring topographies and correcting images for 2D and 3D analysis will be presented. Pitfalls and shortcomings will be discussed.

10:40am AS-WeM9 The Effects of Pre-Processing of Secondary Ion Mass Spectrometry (SIMS) Image Data on Self-Modeling Image Analysis, *W. Windig*, *B.M. Wise*, Eigenvector Research, Inc., *M.R. Keenan*, Sandia National Laboratories

SIMS imaging is a powerful technique for surface analysis. The data from SIMS results in hundreds or thousands of images corresponding to ions of different masses. In order to facilitate data analysis, data reduction techniques are required. One of the tools to reduce the massive amounts of data is self-modeling mixture analysis, which expresses the SIMS image data in a few images representing pure components and their associated mass spectra. This paper will focus on the pure variable approach. A pure variable has contributions from only one component in the mixture data set (i.e. a value of m/e to which only one chemical component contributes) and thus can be used as a relative concentration estimate to resolve the mixture data into pure component spectra and their contributions ("concentrations') in the form of images. Similarly, pure pixels can be selected to resolve the mixture data. Image data are often of a noisy nature. Therefore, pre-processing of the data is often used to improve the results. A popular pre-processing for TOF-SIMS data is based on the Poisson nature of the data. This paper will show a modification of Poisson scaling procedure of the data, which makes it less susceptible to noise. Another way to enhance data analysis is using correlation based techniques to minimize the influence of outlying pixels. This paper will show how the data analysis results, as obtained with the pure variable/pixel approach, can be improved using the proper pre-processing tools, using data sets of actual samples of several chemical mixtures and a fused metal sample.

11:00am AS-WeM10 Comparison of MAF and PCA for Processing 3-D ToF-SIMS Images of Organic and Biological Samples, *B.J. Tyler*, University of the West Indies, Trinidad and Tobago

Recent technological advances have facilitated 3-D SIMS imaging of organic and biological samples. To fully realize the potential of this technology, new tools are needed to aid in the image analysis. Even in two-dimensions, obtaining clear contrast between chemically similar regions, distinguishing between chemical and topographical effects and identifying chemical species from a complex ToF-SIMS data set can be a formidable challenge. These challenges become even greater as the size and complexity of the data sets increase due to both the third dimension and the higher useful mass range commonly obtained with cluster ion sources. In the past, we have found that Maximum Autocorrelation Factors (MAF) provides significant improvement over PCA for enhancing image contrast, reducing spectral complexity and facilitating compound identification.[1,2] We have investigated several approaches to generalizing the MAF approach for use in 3 dimensions. These 3-D MAF algorithms have been tested on synthetic images and on a variety of organic and biological 3-D SIMS images. Results have been compared to converntion single-peak data analysis and to PCA results using various scaling options. MAF, which includes information on the nearest neighbors to each pixel, shows clear advantages over PCA, particularly for identifying spase of subtle components in the images. Additionally, MAF is insensitive to pre-processing choices that can dramatically influence PCA results.

[1] Tyler, B.J. Applied Surface Science, Volume 252, Issue 19,30 July 2006, Pages 6875-6882
[2] Tyler, B.J.,Rayal G, Castner D.G., Biomaterials 2007, May 28(15):2412-23.

11:20am AS-WeM11 SVD + Factor Rotation: A Powerful Alternative to PCA for Spectral Image Analysis, *M.R. Keenan*, Sandia National Laboratories

Factor analysis has proven an effective approach for distilling high dimensional spectral-image data into a limited number of components that describe the spatial and spectral characteristics of the imaged sample. Principal Component Analysis (PCA) is the most commonly used factor analysis tool; however, PCA constrains both the spectral and abundance factors to be orthogonal, and forces the components to serially maximize the variance that each accounts for. Neither constraint has any basis in physical reality; thus, principal components are abstract and not easily interpreted. The mathematical properties of PCA scores and loadings also differ subtly, which has implications for how they can be used in abstract factor "rotation" procedures such as Varimax. The Singular Value Decomposition (SVD) is a mathematical technique that is frequently used to compute PCA. In this talk, we will argue that SVD itself provides a more flexible framework for spectral image analysis since spatial-domain and spectral-domain singular vectors are treated in a symmetrical fashion. We will also show that applying an abstract rotation in our choice of either the spatial or spectral domain relaxes the orthogonality requirement in the complementary domain. For instance, samples are often approximately orthogonal in a spatial sense, that is, they consist of relatively discrete chemical phases. In such cases, rotating the singular vectors in a way designed to maximize the simplicity of the spatial representation yields physically acceptable and readily interpretable estimates of the pure-component spectra. This talk will demonstrate that this approach can achieve excellent results for difficult-to-analyze data sets obtained by a variety of spectroscopic imaging techniques.

11:40am AS-WeM12 A Comparison of Multivariate Statistical Analysis Protocols for ToF-SIMS Spectral Images, *V.S. Smentkowski*, *S.G. Ostrowski*, General Electric Global Research Center, *M.R. Keenan*, Sandia National Laboratories

Time of Flight Secondary Ion Mass Spectrometry (ToF-SIMS) instruments produce raw data sets with a tremendous quantity of data. Multivariate Statistical Analysis (MVSA) tools are being used to boil the massive amount of chemical information into a smaller set of components which are easier to interpret and understand due to species association. Standard Principal Component analysis (PCA) is the most heavily used MVSA algorithm used in the ToF-SIMS community. Other algorithms such as Multivariate Curve Resolution (MCR) have also gained popularity over the past few years. In this work, we compare the as-measured ToF-SIMS spectrum and ion images with four MVSA data analysis protocols; standard PCA, image-rotated PCA, spectra-rotated PCA, and MCR. Image-rotated PCA and spectra-rotated PCA are variations of standard PCA that involve abstract rotation of the principal components, and are designed to enhance either spatial contrast or spectral contrast in the components, respectively. We will show that the four MVSA protocols provide essentially the same information, but accentuate different aspects of the sample's composition and lateral distribution, and that taken together these methods provide a more complete understanding of the sample. We will demonstrate that the component spectra provided by MVSA protocols assists the analyst in understanding species correlation which would have been difficult, if not impossible, using univariate analysis protocols. Since each component image is represented by an associated spectrum (and not just a single peak) enhanced signal-to-noise and contrast is obtained. For the data set described here, MVSA tools identified unexpected species, which were not obvious in the as measured data.

Authors Index

Bold page numbers indicate the presenter

— A —
Artyushkova, K.: AS-WeM3, **1**
Atanassov, P.: AS-WeM3, 1
— B —
Baer, D.R.: AS-WeM1, 1
— C —
Castle, J.E.: AS-WeM1, 1
— E —
Engelhard, M.H.: AS-WeM1, 1
— F —
Fairley, N.: AS-WeM2, 1

Fletcher, I.W.: AS-WeM5, 1
Fulghum, J.E.: AS-WeM3, 1
— G —
Gilmore, I.S.: AS-WeM5, 1
— K —
Keenan, M.R.: AS-WeM11, **2**; AS-WeM12, 2; AS-WeM6, 2; AS-WeM9, 2
— L —
Lea, A.S.: AS-WeM1, **1**
Lee, J.L.S.: AS-WeM5, **1**
— O —
Ohlhausen, J.A.: AS-WeM6, **2**

Ostrowski, S.G.: AS-WeM12, 2
— S —
Seah, M.P.: AS-WeM5, 1
Sisk, D.R.: AS-WeM1, 1
Smentkowski, V.S.: AS-WeM12, **2**
— T —
Tyler, B.J.: AS-WeM10, **2**
— W —
Walton, J.: AS-WeM2, **1**
Windig, W.: AS-WeM9, **2**
Wise, B.M.: AS-WeM9, 2