Partial Least Squares Techniques in the Energy-Dispersive X-ray Fluorescence Determination of Sulfur–Graphite Mixtures

Jurgen Swerts and Piet Van Espen*

Department of Chemistry, Micro and Trace Analysis Centre, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Belgium

Paul Geladi

University of Umeå, Department of Organic Chemistry, S901 87 Umeå, Sweden

Energy-dispersive X-ray fluorescence (ED-XRF) was used to determine sulfur in graphite. Because of interferences and abnormal scattering of the excitation radiation (diffraction effects), spectrum evaluation using least squares fitting failed. The partial least squares (PLS) method on the other hand is able to predict the S concentrations with an accuracy better than 5% in the concentration range of 2–60%. The various artifacts observed in the spectra can be explained with the PLS model.

INTRODUCTION

Energy-dispersive X-ray fluorescence (ED-XRF) is an important analytical technique for the determination of trace, minor, and major elements in a variety of matrices. A drawback of this method is the limited resolution of the energy-dispersive semiconductor detector. With a typical resolution of 150 eV at MnKα, frequent overlap between the characteristic lines of adjacent elements occurs. As a consequence, the method requires the use of sophisticated spectrum evaluation methods to obtain interference-free and background-corrected net peak areas of the analytically important fluorescence lines. These spectrum evaluation methods mainly rely on the least squares principle; analytical functions describing the features in the spectrum[1] or reference spectra of pure elements[2] are used as a model to describe the observed spectrum. After obtaining net peak areas, a variety of quantization procedures are in use, ranging from simple linear calibration curves,[3] over empirical and semiempirical influence coefficients[4,5] to full fundamental parameter approaches.[6]

Although the procedure of spectrum evaluation followed by an independent quantification has proven to work very well in many cases, it remains a complicated process that requires a high degree of experience and knowledge from the operator. A major source of errors is related to the spectrum evaluation. This step becomes especially difficult if distortions and artifacts are observed in the spectrum. These phenomena occur when the spectrometer is operated in a high countrate regime or when unknown peak-like structures such as diffraction lines or scattered excitation lines interfere with the analytical lines. In such cases the spectra become "dirty" and are nearly impossible to describe analytically.

On the other hand multivariate techniques such as principal component regression (PCR) and partial least squares (PLS) have proven to be applicable for quantification in cases where the spectral data as well as the underlying physical process that generated them are very complex.[7-9]

In this paper, we investigate the possibility of using PLS for the quantification of sulfur in a graphite matrix. In this case, the spectrum evaluation and quantization are done in a single pass. Moreover, it will be shown that a number of parameters that normally degrade the quality of the analytical results are automatically taken care of when using the PLS method. This offers interesting possibilities for integrating quality control into the analysis process.

EXPERIMENTAL SECTION

Samples were prepared by pressing well-homogenized sulfur (Janssen Chimica, 99.5%) and graphite (Ringsdorff-Werke GMBH, 99.99%) mixtures into pellets of 3.2 cm diameter at a pressure of 10 tons.

The ED-XRF equipment used was a Tracor Spectrace 5000 instrument (Tracor, Northern USA). The samples were irradiated in vacuum by the unfiltered X-ray spectrum of an Rh-anode X-ray tube operating at 12 kV. X-rays were detected with a Si(Li) semiconductor detector with a resolution of 160 eV (at MnKα).

Table I summarizes the sample composition and the measurement conditions. The samples were irradiated at different tube currents in order to induce the effect of variation in count rate. The measurement time was chosen so that the product of time and current was constant (same number of exciting photons). In Table I the first character of each label refers to the applied tube current, the second character (a or b) is used to distinguish replicates measured under identical experimental conditions, and the third character refers to the sample number. The measurements at 0.01 mA, 500 s, and 0.02 mA, 250 s, were done twice in order to check the reproducibility.

THEORY

Spectrum Evaluation. One of the most widely used methods to obtain net peak areas from an energy-dispersive spectrum is the least squares fitting of the spectral data with a mathematical function. The parameters of this function are calculated such that the weighted sum of squared difference between the observed spectrum and the model (the χ^2 value) is minimum. Since the analytical function is

(1) Van Espen, P.; Nullens, H.; Adams, F. *Nucl. Instrum. Methods* **1977**, *142*, 243–250.
(2) Schamber, F. H. *X-ray Fluorescence Analysis of Environmental Analysis*; Dzubay, T., Ed.; Ann Arbor Science Publ.: Ann Arbor, MI, 1977.
(3) *Introduction to X-ray Spectrometric Analysis*; Bertin, E. P., Ed.; Plenum Press: New York, 1978.
(4) Rousseau, R. M. *Adv. X-ray Anal.* **1989**, *32*, 69–75.
(5) Lachance, G. R. *Adv. X-ray Anal.* **1988**, *31*, 471–478.
(6) He, F.; Van Espen, P. *Nucl. Instrum. Methods* **1990**, *A299*, 580–583.
(7) Wang, Y.; Zhao, X.; Kowalski, B. *Appl. Spectrosc.* **1990**, *44*, (6), 998–1002.
(8) Karstang, T. V.; Eastgate, R. J. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 209–219.
(9) Cahn, F.; Compton, S. *Appl. Spectrosc.* **1988**, *42*, 865–872.

0003-2700/93/0365-1181$04.00/0

Table I. Identification of Measured X-ray Spectra: Sulfur Contents of Samples and Experimental Conditions

	tube current and measuring time					
sulfur contents (% w)	0.02 mA, 250 s	0.01 mA, 500 s	0.02 mA, 250 s	0.01 mA, 500 s	0.03 mA, 167 s	0.04 mA, 125 s
1.04	2a1	1a1	2b1	1b1	3a1	4a1
3.10	2a2	1a2	2b2	1b2	3a2	4a2
5.07	2a3	1a3	2b3	1b3	3a3	4a3
8.15	2a4	1a4	2b4	1b4	3a4	4a4
10.26	2a5	1a5	2b5	1b5	3a5	4a5
19.06	2a6	1a6	2b6	1b6	3a6	4a6
49.72	2a7	1a7	2b7	1b7	3a7	4a7
80.15	2a8	1a8	2b8	1b8	3a8	4a8

most often nonlinear in one or more parameters, this minimization is done in an iterative way.

The reduced χ^2 is defined as

$$\chi^2 = \frac{1}{n-m}\sum_{i=n_1}^{n_2}\frac{1}{\sigma_i^2}[y - y(i, a_1, ..., a_m)]^2 \quad (1)$$

where n is the number of channels in the spectrum being evaluated (from n_1 to n_2), and m is the number of parameters in the model. The weight σ_i^2 is equal to the channel content of the measured spectrum y_i (Poisson statistics). $y(i, a_1, ..., a_m)$ represents the fitting model.

For a perfect fit, χ_ν^2 follows a χ^2 distribution with the expected value of one. For visual evaluation of the quality of the fit, the residuals defined as

$$R_i = \frac{y_i - y(i, a_1, ..., a_m)}{\sqrt{y_i}} \quad (2)$$

are plotted versus the channel number i. For a good fit 99% of the R_i values should be between the limits −3 and +3 and should not show any structure (randomly distributed).

Partial Least Squares Regression (PLS). The goal of chemical analysis is to find a relationship between a known concentration (called y-variable) and some measurable property (called x-variable), using a set of samples called "calibration set".[10] The assumption is then that the function found for this calibration set will work for future samples, allowing the estimation of concentration from the measured property. In the case of sulfur in graphite analysis by XRF, known sulfur concentrations and measurement conditions have to be related to measured ED-XRF spectra or portions of them.

In the case of sulfur in the coal data set, the x-variables, the spectral variables can be described as a matrix **X** of size 48 (rows) and 600 (columns). The 48 S concentrations would form a vector **y** of size 48 by 1. The relationship can then be written in the form:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{f} \quad (3)$$

where **b** is a 600 × 1 vector of regression coefficients (one for each energy channel in the subset) and **f** is a 48 × 1 vector of residuals, explaining misfit between the model and the data.

If there is more than one y-variable, the above equation can be written:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{F} \quad (4)$$

This equation is more general and will be used subsequently. See also Figure 1. The matrix **Y** has J columns where J is the number of y-variables. **F** is of the same size as **Y**. The matrix **B** contains J columns of regression coefficients. In the present example, where both the concentration of S and the irradiation intensity (expressed as X-ray tube current in mA) are used as y-variables, the matrix **B** consists of two vectors. Usually X and Y are pretreated before the relationship between them is calculated.[11] The pretreatments used are mentioned in the Results section.

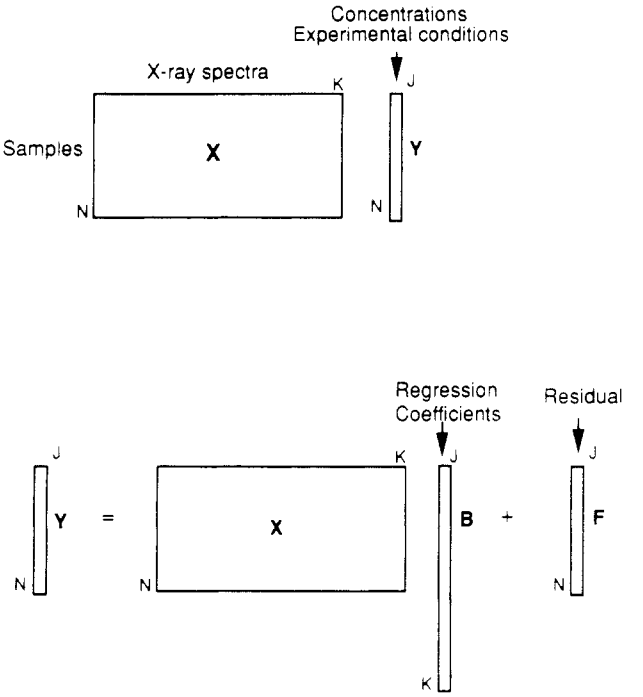

Figure 1. Matrix representation of a linear regression model. This model is built between the matrix **X** of measured X-ray spectra and **Y** of known concentrations and experimental conditions. The model is a linear one. The residual describes the data in **Y** that cannot be modeled by the *X*-data.

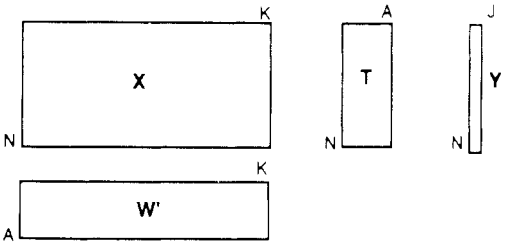

Figure 2. Important elements of the PLS2 model are the *A* score vectors and the *A* loading weight vectors. *A* is the dimensionality of the model.

There are many ways of calculating **B**. Not all of them are equally good in situations where there are more variables than samples and where the variables are expected to be very collinear. The method of partial least squares regression has proven to be able to deal with many-variable cases and collinearity very well.[11-13] PLS and similar latent variable regression methods do this by creating a limited number of latent variables that are used as regressors.[11-13] An introduction to the theory of PLS can be found in refs 11–13. Earlier applications of PLS to calibration problems for spectra are found in refs 14–17. A more profound theoretical discussion is found in refs 18–21.

(10) Sharaf, M. A.; Illman, D. L.; Kowalski, B. R. *Chemometrics*; John Wiley & Sons: New York, 1986; Vol. 82.

(11) Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley & Sons: New York, 1989.

(12) Wold, S.; Ruhe, A.; Wold, H.; Dunn, W., III *J. Sci. Stat. Comput.* **1984**, *5*, 735–743.

(13) Geladi, P.; Kowalski, B. *Anal. Chim. Acta* **1986**, *185*, 1–17.

(14) Lindberg, W.; Persson, J. A.; Wold, S. *Anal. Chem.* **1983**, *57*, 643–648.

(15) Otto, M.; Wegscheider, W. *Anal. Chem.* **1985**, *57*, 63–69.

(16) Haaland, D.; Thomas, E. *Anal. Chem.* **1988**, *60*, 1193–1202.

(17) Haaland, D.; Thomas, E. *Anal. Chem.* **1988**, *60*, 1202–1208.

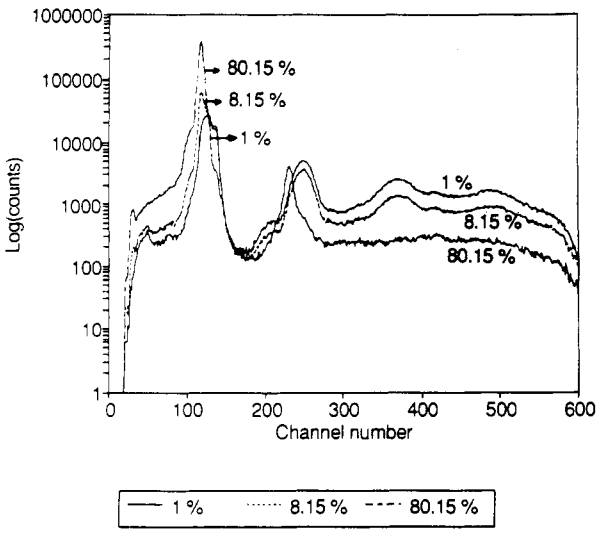

Figure 3. X-ray spectra of the samples containing 1, 8, and 80.15 wt % sulfur in graphite.

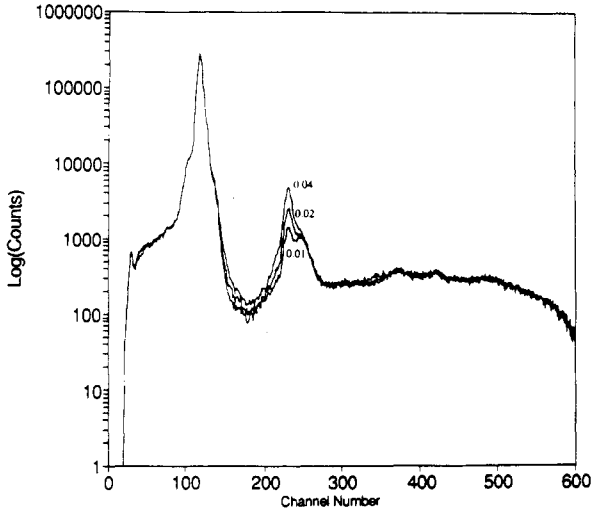

Figure 4. X-ray spectra of the sample containing 50 wt % sulfur measured at a tube current of 0.01, 0.02, and 0.04 mA.

For now, it suffices to explain the elements in Figure 2. The scores T and loading weights W are used not only to calculate the regression coefficient B but also as a means of interpreting the model. Score vectors can be plotted against each other as scatter plots (score plots). Loading weight vectors can be plotted against channel number in a spectrum-like fashion. In this way the contribution of each PLS dimension to the model can be interpreted spectroscopically.

RESULTS AND DISCUSSION

In Figure 3 the X-ray spectra of samples containing 1%, 8.15%, and 80.15% of sulfur measured at a current of 0.02 mA are displayed. One observes a clear increase in the intensity of the S-Kα peak located at channel number 115. The Rh-L lines from the tube spectrum are scattered by the sample and interfere considerable with the analytically important S-Kα line. Some other peak-like structures above channel 150 are also observed. This will be discussed further.

Figure 4 shows X-ray spectra of the sample containing 50% sulfur measured respectively at 0.01, 0.03, and 0.04 mA. Small differences at the high energy tail of the peak are observed as well as an increase of the peak structure above channel 150. These phenomena are caused by pileup effects in the pulse processing electronics of the instrument. The magnitude of this effect increases with count rate (and thus with tube current).

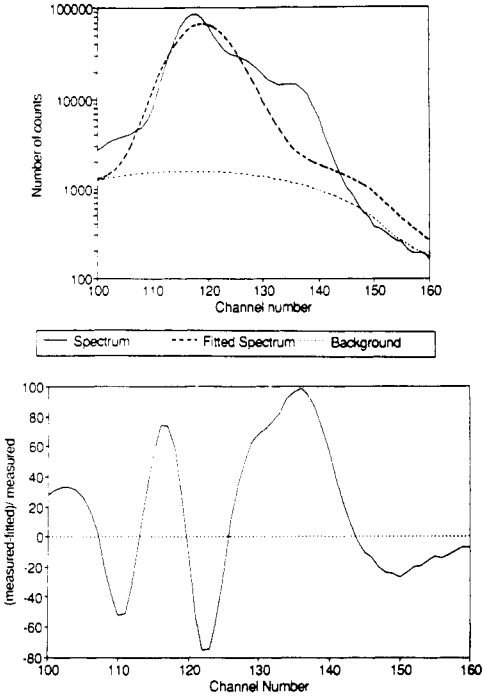

Figure 5. Original and fitted spectrum of the sample containing 10 wt % sulfur and measured at 0.02 mA, together with the residual plot.

First, an attempt was made to obtain the net peak area of the S-Kα,S-Kβ line and the Rh-L lines by least squares fitting of the spectral region from channel 100 to channel 160 on a second degree polynomial background. The spectrum evaluation was done using the program AXIL (Analysis of X-ray spectra by Iterative Least squares).[22] In Figure 5, the spectrum of the sample containing 10% S measured at a tube current of 0.02 mA is shown together with the fit and the residuals. The reduced χ^2 value of this least squares fit is 321, indicating that the model is unable to describe the observed spectrum. This is mainly due to the fact that incomplete charge collection in the detector causes the shape of the observed X-ray lines to deviate considerable from a Gaussian (low energy tailing). Also, the incoherently and coherently scattered Rh-L lines and the background are difficult to model. Some other effects may contribute to the observed spectrum which are not modeled by the least squares approach as will be discussed later.

From this observation, it becomes clear that spectrum evaluation by classical least squares fitting is nearly impossible and that any quantitative method relying on the net peak areas calculated via this method will fail.

In order to evaluate this problem using a multivariate approach, the program Unscrambler was used.[23,24] The **X** matrix contains the different spectra, and the **Y** matrix contains the concentration of sulfur and the tube current (mA) used for every object. The square root of the X-data was taken in order to give a homogeneous variance. Indeed since the channel content follows a Poisson distribution (counting statistics, $s^2_x = x$), the square root of the channel

(18) Helland, I. *Commun. Stat., Simul. Comp.* **1988**, *17*, 581–607.
(19) Lorber, A. Wangen, L.; Kowalski, B. *J. Chemom.* **1987**, *1*, 19–31.
(20) Hoskuldsson, A. *J. Chemom.* **1988**, *2*, 211–228.
(21) Manne, R. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 187–197.
(22) Van Espen, P.; Janssens, K.; Nobels, J. *Chemom. Intell. Lab. Syst.* **1986**, *1*, 109–114.
(23) Unscrambler 2, Software for multivariate data analysis applying PCA, PCR and PLS, CAMO A/S.
(24) Helland, K. *J. Chemom.* **1991**, *5*, 413–415.

Table II. Percentage of Variance Explained for Each Matrix with, Respectively, Prediction Errors Using 1–4 Factors

	% matrix variance explained				square error of modeling prediction	
factor	X	Y_{tot}	Y_{conc}	Y_{mA}	X	Y
0	0	0	0	0	218	1.0
1	98.8	48.6	97.2	0.2	2.94	0.55
2	99.5	63.3	98.2	28.4	1.40	0.38
3	99.8	74.6	98.4	51.0	0.45	0.30
4	99.9	97.8	98.6	96.8	0.40	0.03

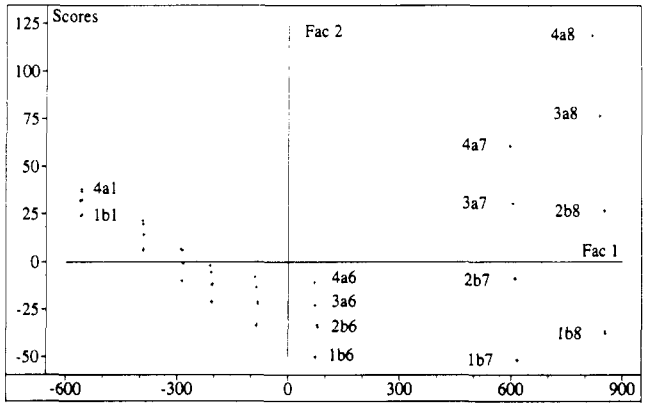

Figure 6. Scores plot of factor 1 against factor 2.

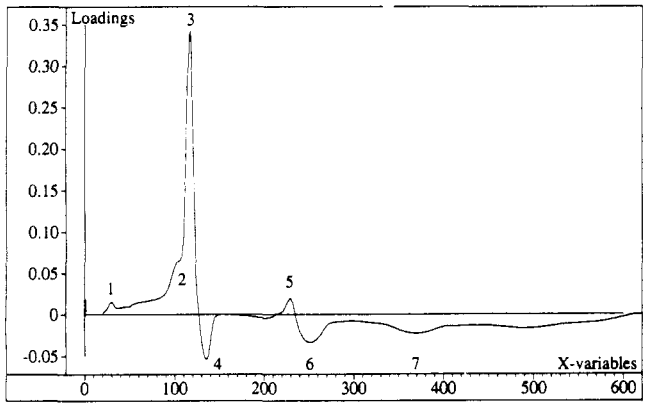

Figure 7. Loading plot of the first factor (first eigenspectrum).

contents has a variance of 1/2. The same was done for the **Y** matrix.

Mean centering was applied, and the data of the **Y** matrix were weighted using the reciprocal of the standard deviation. This weighting is performed because concentration and tube current do not have the same dimensions. Cross-validation was used in order to find the number of relevant factors. Table II displays the percentage of variance explained of the **X** and **Y** matrixes and the square error of modeling prediction, together with the percentage of variance explained for each Y-variable. One observes that four factors account for 99.9% of X-data and 97.8% of the Y-data. The first factor explains the variability of the concentration, while factors 2, 3, and 4 explain the influence of the tube current on the spectra.

If one looks at the score plot of factor 1 against factor 2 (Figure 6), one can see that the first factor separates the different concentrations (projection of the objects on the horizontal axis, the first component) and that the second factor is related to the applied tube current.

Figure 7 displays the "eigenspectrum" of the first factor. Seven features can be observed. Peak 1 corresponds to the escape peak of sulfur Kα. Peak 2 is due to the effect of the incomplete charge collection of the detector. Peak 3 corresponds with the K lines of sulfur. Peak 4 is related to the

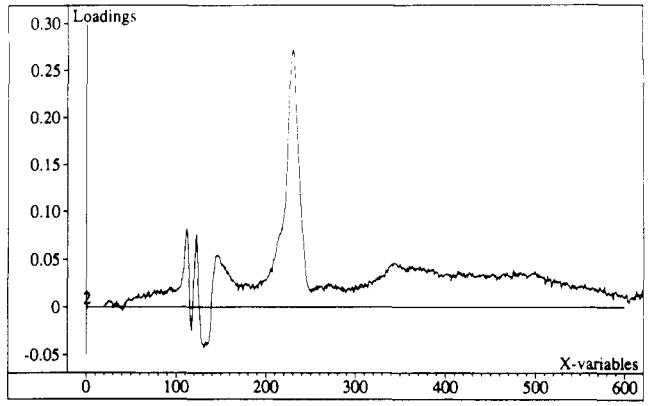

Figure 8. Loading plot of the second factor (second eigenspectrum).

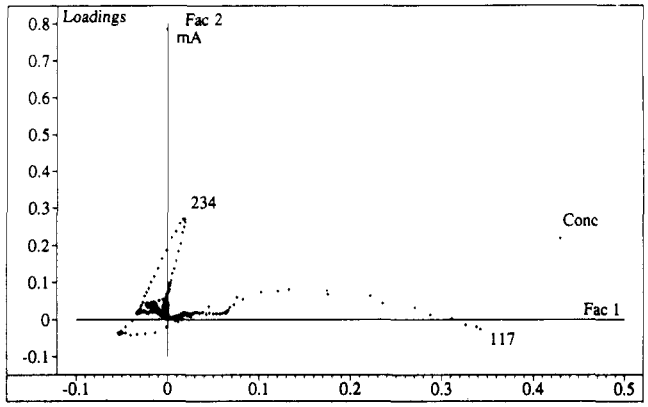

Figure 9. X and Y loadings of the first against the second component.

scattered L lines of rhodium and the first order diffraction effect of graphite. Peak 5, with the maximum in channel 234, is the sum peak of the sulfur K lines. Peaks 6 and 7 are more difficult to explain.

One possible explanation is that those are respectively second- and third-order diffraction effects. These observations can be interpreted in the following way. If the sulfur content in the samples increases, the peak structures 1, 2, 3, and 5 increase (positive loadings). Indeed, it is evident that the S-K line's intensity is directly related to the concentration of sulfur, and it is well-known that the escape (peak 1) and incomplete charge collection (peak 2) are directly proportional to the intensity of these sulfur K lines.

The increase of the sum peak is also understandable from the relation:

$$N_{11} = \tau N_1 N_1 \qquad (5)$$

where N_{11} is the count rate in a sum peak due to the coincidence of two pulses with the same energy; in our case, N_1 is the count rate of the sulfur Kα peak. τ is the pulse-pair resolution time.

Structures 4, 6, and 7 decrease with an increasing sulfur content as well as the background region between channels 260 and 600. This is due to the lower scattering cross-section of sulfur compared to carbon. The peak-like structures observed (4, 6, and 7) can be explained by considering the Bragg law $n\lambda = 2d \sin\theta$ with 2d = 0.6708 nm for graphite and $\theta = 45°$. Diffraction phenomenon can be expected for energies of 2.61, 5.22, and 7.84 keV corresponding to channel numbers 130, 261, and 392. This diffraction phenomena gives rise to an increased intensity of the scattered continuum radiation and is observed as a broad peak-like structure in the spectrum. The first-order diffraction is more difficult to observe directly as it is a coincidence with the energy of the Rh-L lines.

Figure 8 displays the second eigenspectrum. In this spectrum, the sum peak at channel 230 is related to milli-

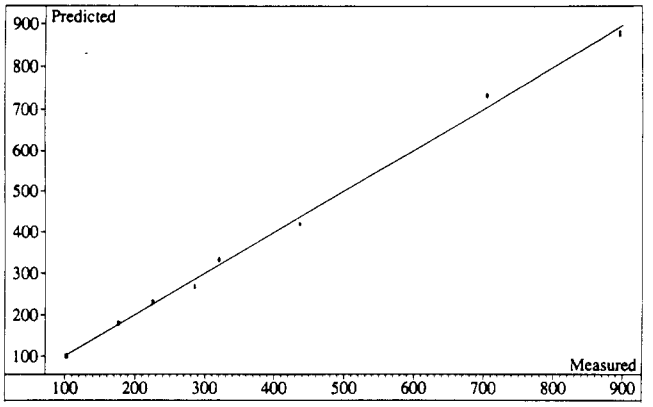

Figure 10. Plot of the true against the predicted concentrations of the objects of the calibration model using the PLS model with four factors.

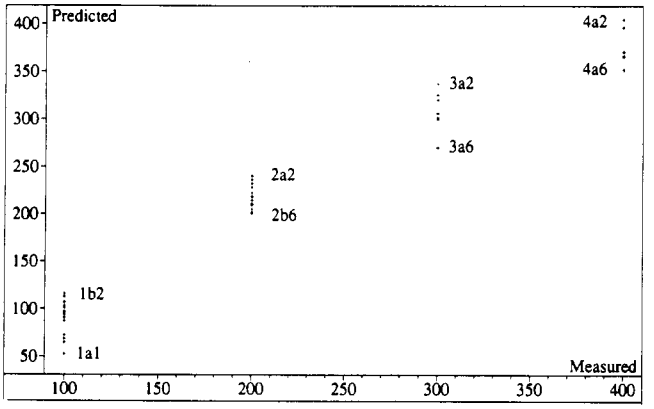

Figure 11. Plot of the true against the predicted tube current of the objects of the calibration model using the PLS model with four factors.

Table III. True and Predicted Sulfur Concentrations of Pellets Not Used To Build Calibration Model, Together with the Percentage of Concentration Difference and True and Predicted Tube Current Used.

pellet	[S] in %	predicted [S]	% difference	mA used	mA predicted
2u1	2.12	2.06	2.8	2	2.56
2u2	16.33	17.64	-8.0	2	2.25
2u3	29.24	30.9	-5.7	2	1.96
2u4	62.06	63.04	-1.6	2	2.56
3u1	2.12	1.96	7.5	3	3.72
3u2	16.33	17.38	-6.4	3	2.96
3u3	29.24	31.14	-6.5	3	2.72
3u4	62.06	63.20	-1.8	3	3.53

amperes (~ count rate). The oscillations around channel 150 are caused by the shift of the position of the sulfur peak in the X-ray spectrum at higher count rate.

Figure 9 displays the loadings of the first component against the loadings of the second component. One sees that the concentration is mainly determined by the channels around 117, and the tube current is determined mainly by the channels around 234.

The third and fourth eigenspectra are not shown here, because we cannot yet retrieve the complete information or relation to the experimental factors. They contain information on base-line shifts and non-linearity phenomena, beside other, yet unknown effects.

Figure 10 shows the predicted and the true concentration of sulfur in the calibration set using the PLS cross-validated model with four factors. In Figure 11, the predicted tube current is compared with the true value.

New samples were made. Measurements were done at 0.02 and 0.03 mA, respectively. The results are summarized in Table III. The mean percentage of difference between true and predicted concentrations is 2.5% S. It appears that an excellent prediction of the S contents was made on the concentration range 2–60% for both measurement conditions. The prediction of the applied tube current is less accurate, although the general trend is evident. This can be explained by the fact that the milliamperes are only predicted by the sum peaks and the shift of the spectrum. Moreover, the current in the calibration set is only varied over a very narrow range, 0.01–0.04 mA.

CONCLUSION

When the experimental conditions in ED-XRF are such that the observed X-ray spectra are very complex, e.g., when interferences from scattered peaks, diffraction phenomena, etc. occur, the classical approach of spectrum evaluation by least squares fitting followed by quantization based on the net peak area of the fluorescent lines fails. Spectrum evaluation can be carried out in an interactive, supervised manner, but in many cases some kind of automation is desirable. The spectra studied in this work do not easily allow for this automation. PLS, on the other hand, handles the cases very well, and it is able to give very accurate estimates of the concentrations. The PLS calculations require no human interaction whatsoever. When a good calibration model is available, the PLS model can be used as a "black box" in the concentration calculations, with possibly an outlier alert mechanism included (not tested for this paper). This may simplify automation substantially.

Another advantage of the PLS method is that the underlying physical phenomena observed in the spectra and their relation with the experimental conditions can be explained. This opens possibilities for the use of PLS as a part of a quality control procedure where the experimental conditions and abnormal behavior of the ED-XRF spectrometer are monitored by the different PLS components. More research is, however, required to fully exploit this approach.

RECEIVED for review August 6, 1992. Accepted January 15, 1993.