Automatic Classification of Steels by Processing Energy-Dispersive X-ray Spectra with Artificial Neural Networks

Jorge F. Magallanes*,†,‡ and Cristina Vazquez†

Unidad de Actividad Química, Centro Atómico Constituyentes, Comisión Nacional de Energía Atómica, Avenida Del Libertador 8250, 1429 Buenos Aires, Argentina, and Departamento de Química Inorgánica, Analítica y Químicafísica, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Buenos Aires, Ciudad Universitaria de Nuñez, Pabellón 2, 1428 Buenos Aires, Argentina

Received December 17, 1997

The automatic classification of steels has been studied. The chemical compositions of 19 certificate steels were correlated with its energy dispersion X-ray fluorescence spectra. Twelve relevant elements of these samples were selected for data processing through artificial neural networks (ANNs). A Kohonen type ANN of 8 × 8 × 1 1 dimension was used. This net architecture allows on-line classification with 100% efficiency, that is, without errors.

INTRODUCTION

Energy-dispersive X-ray spectrometry (EDX) is a widespread analytical technique frequently used for materials characterization. Qualitative and quantitative information of the sample can be obtained quickly.[1] Many modern instruments can automatically perform the calculations for the final results. In the case of solid materials, the pretreatment of the sample is simple and wet chemistry can be avoided.

On the other hand, artificial neural networks have shown the capability of performing classification tasks.[2,3] For problems involving a big number of variables, ANNs appears to be more efficient than standard chemometric techniques such as hierarchical clustering and principal component analysis.[5,9]

In order to combine these two capabilities, we have explored in this work the possibility of analyzing an EDX spectrum through a Kohonen net in order to classify automatically different types of steels. The main goal of this work is to get a method for an efficient and fast classification of unknown steel samples capable of being used "on line" with a minimum of errors and operator intervention.

EDX MEASUREMENTS

The instrument consists of an excitation source of ^{241}Am radioisotope. Its radiation impinges on a Sn secondary target. This characteristic secondary radiation excites the analyte of the sample. A Si(Li) detector, electronic amplifiers, and a multiple channel analyzer complete the detection system. Many function and data acquisition operations are performed by an associated computer system.

The spectra were analyzed by the AXIL program (Analysis of X-ray Spectra by Iterative Least Squares Fitting).[6,7] The emitted characteristic lines are represented as Gaussian function shapes, and the background level radiation is given as straight lines or a polynomial curve. The net area of the peak is obtained by a nonlinear least squares calculation using a modified Marquardt algorithm.

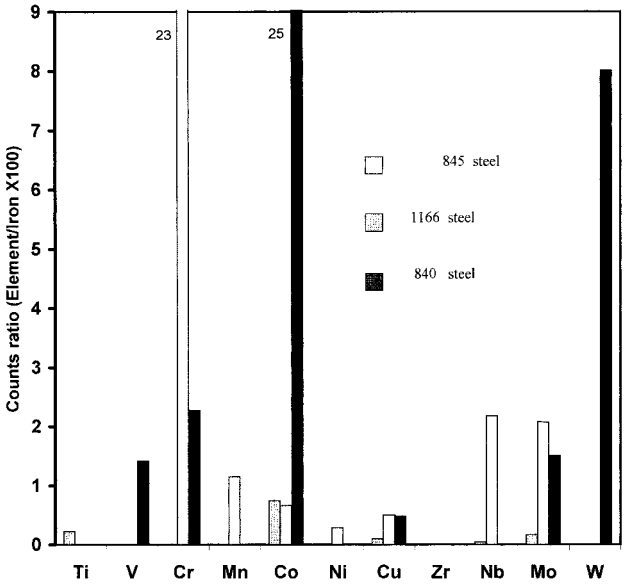

Figure 1. : Three EDX expectra of the three standard types (one type each). The eleven normalized variables are shown as relative intensities proportional to the iron matrix intensities. The factor 100 is included only for a better appreciation of the bars.

Three different sets of certified standards steels were measured. One set is identified with numbers from 836 to 841, and they are tool steels. Another set, numbered from 845 to 850, is one of stainless steels (see Table 1). The last one is a set of low-alloy steels, numbered from 1161 to 1167. All of the standards are from the National Institute of Standards and Technology (NIST), USA. Within Table 1, there are uncertified elements concentrations; they have been measured to complete our data matrix. Figure 1 shows three superposed spectra, each one belonging to a different steel class. The heights of the bars are proportional to the iron intensity.

The XRF spectrums of 19 certified standard steels have been obtained. Each standard was measured 10 times, all of them at the same experimental conditions but on different days. Then, we obtained a collection of 190 spectra, that is, 190 samples. Twelve elements were chosen from all of

† Centro Atómico Constituyentes Comisión Nacional de Energía Atómica.
‡ Universidad Nacional de Buenos Aires.

S0095-2338(97)00114-5 CCC: $15.00
Published on Web 07/03/1998

Table 1. Certified Standard Steels Data[a]

standard	ID letter	Mn	Cu	Ni	Cr	V	Mo	Ti	Zr	Nb	Co	W	Fe
1161	G	0.36	0.34	1.73	0.13 ND	0.024	0.30	(0.01) ND	<0.005	0.011	0.26	0.012 ND	96.50
1162	H	0.94	0.20	0.70	0.74	0.058	0.080	0.037 ND	0.063	0.096	0.11	0.053 ND	95.55
1163	I	1.15	0.47	0.39	0.26	0.10	0.12	0.010 ND	0.20	0.19	0.01	0.105 ND	96.16
1164	J	1.32	0.094	0.13	0.078	0.295	0.029	0.004 ND	0.010	0.037	0.02	0.022 ND	97.23
1165	K	0.032 ND	0.019	0.026 ND	0.004 ND	0.002 ND	0.005	0.20	(0.002) ND	(0.001) ND	0.008	(0.001) ND	99.42
1166	L	0.11 ND	0.033	0.051	0.011 ND	0.007ND	0.011	0.057	<0.005 ND	0.005	0.04	(0.006) ND	99.53
1167	M	0.27	0.067	0.038	0.036	0.041	0.021	0.26	0.094	0.29	0.07	0.20	98.06
836	N	0.21	0.075	NI−ND	6.02	0.63	2.80	NI−ND	ND−NI	NI−ND	M−NI	9.7	79.2
837	O	0.48	M−NI	NI−ND	7.79	3.04	1.50	NI−ND	ND−NI	NI−ND	2.9	2.8	79.7
838	P	0.20 ND	0.17	NI−ND	4.66	1.17	8.26	NI−ND	ND−NI	NI−ND	4.9	1.7	77.6
839	Q	0.18 ND	0.12	NI−ND	2.72	1.50	4.61	NI−ND	ND−NI	NI−ND	7.8	5.7	76.0
840	R	0.15 ND	0.059	NI−ND	2.12	2.11	0.070	NI−ND	ND−NI	NI−ND	11.8	13.0	69.1
841	S	0.27	0.072	NI−ND	4.20	1.13	0.84	NI−ND	ND−NI	NI−ND	M−NI	18.5	73.6
845	A	0.77	0.065	0.28	13.31	0.05	0.92	0.03 ND	ND−NI	0.11	M−NI	0.42	83.2
846	B	0.53	0.19	9.11	18.35	0.03 ND	0.43	0.34	ND−NI	0.60	M−NI	0.04 ND	68.8
847	C	0.23	0.19	13.26	23.72	0.03 ND	0.059	0.02 ND	ND−NI	0.03	M−NI	0.06 ND	61.8
848	D	2.13	0.16	0.52	9.09	0.02 ND	0.33	0.23	ND−NI	0.49	M−NI	0.14 ND	85.3
849	E	1.63	0.21	6.62	5.48	0.01 ND	0.15	0.11	ND−NI	0.31	M−NI	0.19 ND	84.2
850	F	M−NI	0.36	24.8	2.99	0.006 ND	M−NI	0.05 ND	ND−NI	0.05	M−NI	0.21	70.8

[a] NI, noninformed; ND, nondetected; M, measured for this work; ( ), uncertified, informed as indicative value.

the components shown in the spectrum. The chemical criteria for selecting these twelve elements were the importance of the element in steel compositions and the sensing ability of EDX to get accurate precision. With these elements we tried to embrace all the different classes of steels of our bunch of standards and perhaps these sets are enough to include many more types of steels. For each standard some of these elements are nondetectable. These cases have been considered as zero concentration in the data matrix. Instead of working directly with the net photon counts of elements, the ratio between counts of each element with respect to iron were used. The iron count is the highest, so the ratio will always be less than 1. Then, the twelve measured elements become eleven normalized variables because a degree of freedom is lost in the calculation of ratios. We denote x_{new} as the new variable definition, x_{old} as the counts for one of the elements, and x_{Fe} as the counts for iron; then the arithmetic relationship is $x_{new} = 10x_{old}/x_{Fe}$. As the iron count is much higher than the rest of the elements, the factor 10 is included for the purpose of approaching x_{new} to 1. It must be taken into account that this is not just a concentration ratio related to iron, because the fluorescence intensities for each element are quite different. The data processing is facilitated by normalizing the variables. Furthermore the normalization helps to typify very different classes of steels by relating each element fluorescence intensity to the main component of the matrix.

STATISTICAL METHODS

In a previous work,[8] the classification of steels was proposed by using an algorithm based in the properties of the sample and applying a cluster classification with Euclidean distance and *k* nearest neighbor methods. With over 100 samples and with use of 10 parameters the efficiency obtained was approximately 90%.

Cluster analyses of our data have been tried. As is explained below, the 190 samples were randomly divided into two set of 95 samples each. Many kinds of linkage methods were studied, as follows: single, average, centroid, and ward methods, but all of them lead to the same conclusion, with only little differences in the dendograms obtained. When the calculation is forced to look for 19 clusters, the result is that a series for two low-alloy steels, 1161 and 1165, are grouped together. Steel 848, at the same level of similarity, is divided into two subgroups. This means that at higher similarities than this, the cluster for the 1161 and 1165 series cannot be distinguished before other series being artificially divided. Figure 2 shows a dendogram of a typical cluster analysis output.

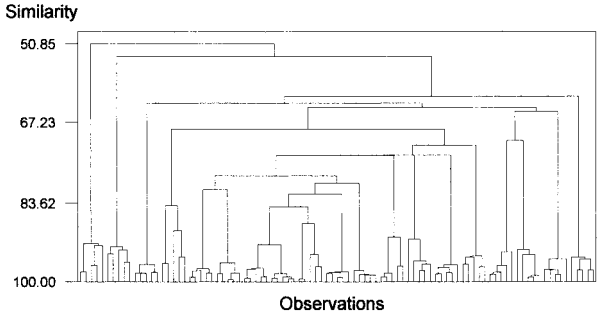

Figure 2. Dendogram of a hierarchical cluster of observations. Standardized variables with single linkage and Euclidean distance.

In a previous EDX work for mineral classification[5] principal component analysis (PCA) has been applied. In this case each element composition can be expressed as a sum of contributions of the chemical compounds that constitute the sample. In our case of steel classification, the sample is composed of pure elements and its concentration cannot be expressed as a sum of independent contributions. Furthermore the interelement effect in an EDX experiment can enhance or absorb the emission intensity of each element, depending on the sample composition. As a result of these considerations, the basic condition in order to apply PCA is not fulfilled. That is, the intensities cannot be expressed as a sum of independent products, as in $m = \sum r_i x_i$. Furthermore, for the case of mineral classification previously mentioned, the authors reported a relatively poor separation ability of this method.

KOHONEN CLASSIFICATION

Basically, a Kohonen net consists of a plane surface of $n \times n$ cells. Each cell is connected to an input layer and to

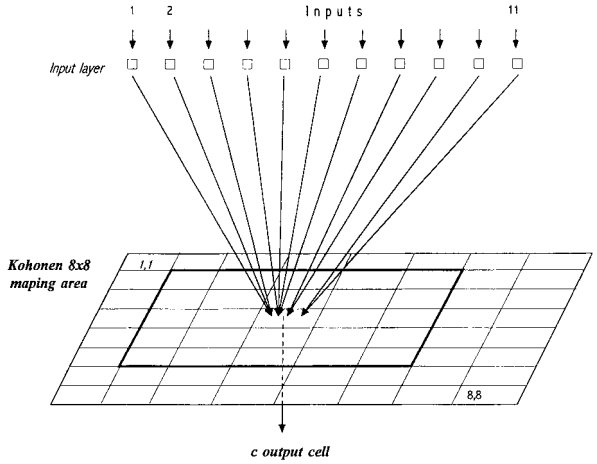

Figure 3. Architecture of the Kohonen net. For the sake of clarity, the connections of only one output *c* cell is shown.

its neighbor cells (see Figure 3). Each neuron of the input layer receives a signal component (a variable) of the vector sample. The neurons of the plane net have *m* weights input, each one received from one input of the input layer. The total input to a neuron is

$$\text{net} = \sum_{i=1}^{m} w_i x_i \qquad (1)$$

where *w* are weights and *x* are the variables of a vector sample.

Only one cell of the plane is selected as output following an imposed criteria, as in eq 3 in our case. For the selected cell the weights are corrected with the intention that the next time it will be more similar to the criteria imposed. The weights of the neurons into a subarea formed by the neighbors of that cell are corrected to a lesser degree than the central cell. By this way a subarea of the plane is adapted to receive objects according to the selected criteria. Then, the next object enters into the net in order to continue with the learning process. The net must be trained with many objects (samples) to optimize all the weights. After that, it is ready to perform the classification of the objects in several subareas of the plane. The opposite sides of the plane are joined between them, resulting in an actual toroid surface. For complete details of the calculation method the reader is addressed to refs 2−4.

In our case, a Kohonen net with eleven inputs per neuron shall be necessary, because we have eleven variables. After trying with different net sizes, a $8 \times 8 \times 11$ net architecture was shown to be the most efficient one. Usually the trial and error method is used to optimize the net size; one of the measured errors could be the squared error criterion for measuring the clustering quality; that is,

$$E = \sum_{p} ||Wc - Xp||^2 \qquad (2)$$

where ||...|| is the Euclidean distance operator, *c* is the selected neuron when the input *Xp* is entered and *Wc* is the center of the cluster corresponding to neuron *c*.

Figure 3 shows the architecture of the net with the 8×8 feature plane map surface. In this figure, each cell represents a neuron, all of them are connected with the input layer,

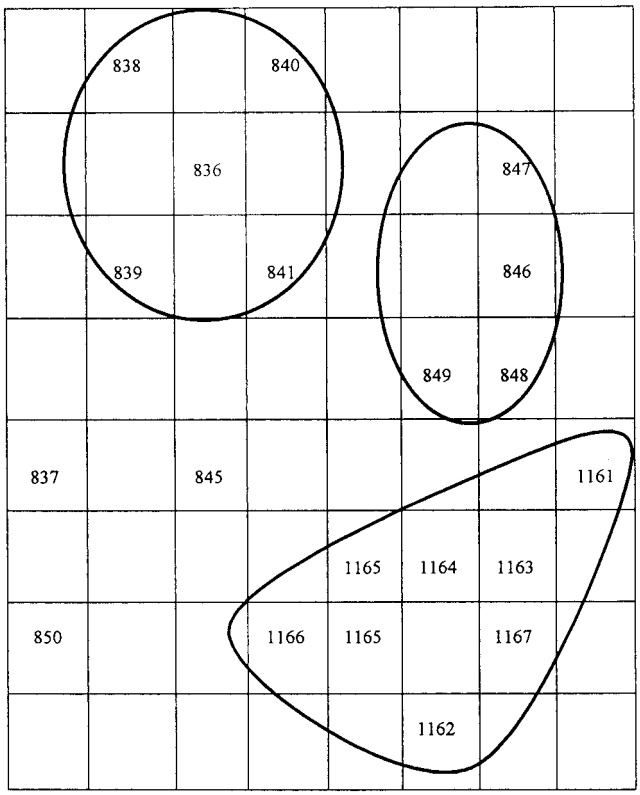

Figure 4. Output top map representation of the Kohonen "toroid" surface. The classification of the 19 different steels can be appreciated. The surrounding areas include members of similar steel type.

Table 2. Parameters for Data Processing

surface mapping geometry	toroid	max correction factor	0.90
scaling function shape	triangle	min correction factor	0.30
topological distance	2	no. of epochs	2

and for clearness the connections for only one cell are shown. Also, it shows the neighborhood correction area associated with the *c* output cell. The 190 samples were randomly divided into two sequences of 95 samples each. One set was used for training and the other for prediction. For the learning stage, the computer running parameters selected for data processing are described in Table 2. A value of 1 for all of the set of 704 initial weights [(8 × 8 neurons) × (11 input levels)] were selected. The efficiency of the data processing does not change significantly by decreasing this initial value. The selection of the central neuron *c* was done according to the criteria of the "weight vector *Wj* most similar to the input signal", the following equation takes this into account:

$$c \leftarrow \min[\sum_{i=1}^{m} (X_{si} - w_{ji})^2] \qquad (3)$$

X being an input variable; *s*, a sample; *i* is one of the $m = 11$ neuron weights, and *j*, a particular neuron.

CONCLUSION

We explain at first the results of the learning stage. After training with a first series of 95 samples, an output map of the 8×8 top area is obtained, and it is shown as a plane in Figure 4. This map is undependable for the order in which

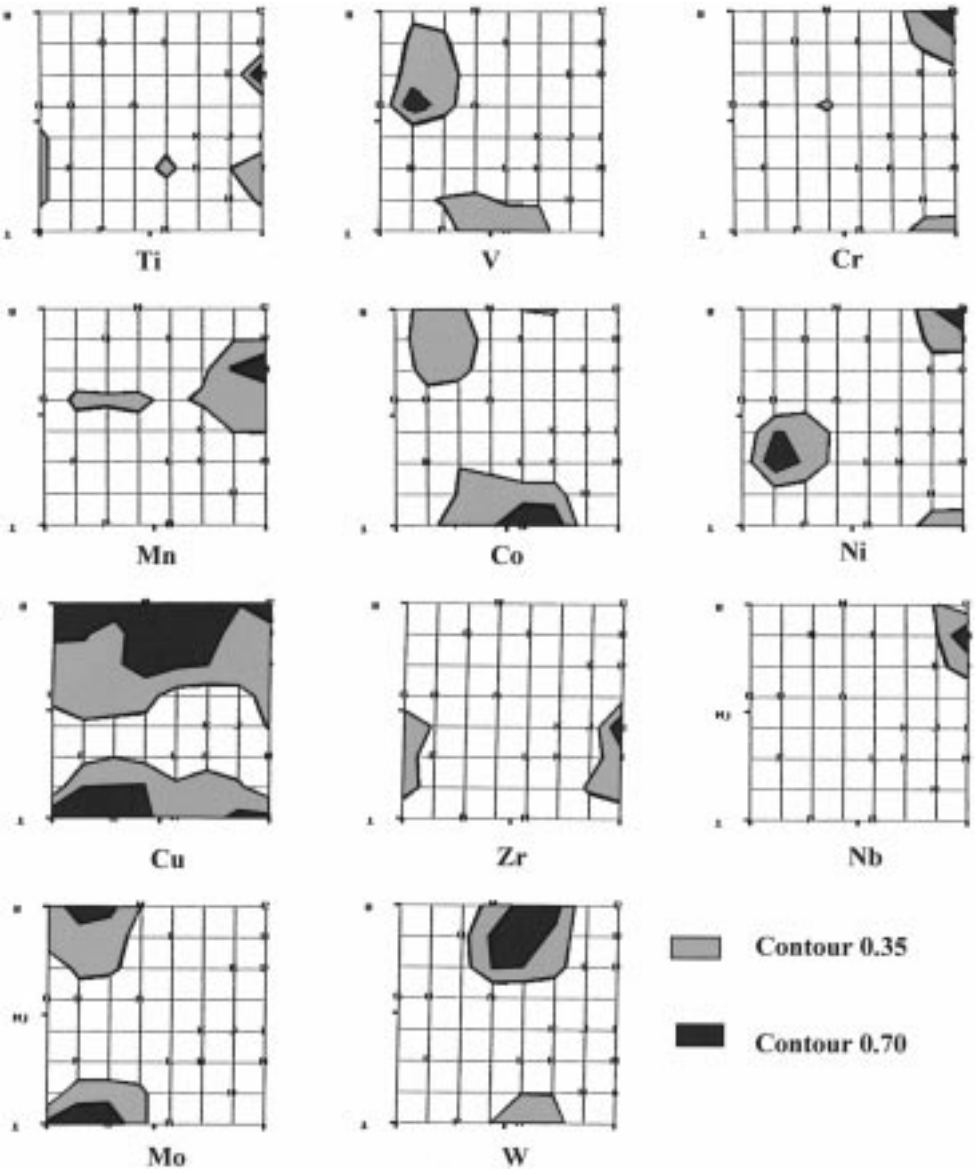

Figure 5. Eleven variable levels, each one representing an element. Two contour lines are shown, one of them for 0.35 and the other for 0.70 of the maximum weight value of each level (see text for more details).

the samples have been fed. In said figure there are numbers identifying the steel and its location in the neural net plane. As can be seen, each specific steel is located in a very reduced area of neighbor neurons. Actually only one cell, except for the 1165 steel. Furthermore, steels of similar types are grouped together. It can be observed that the low-alloy series 1161−1167 encircled is located at the bottom right corner of the plane, most of the tool steels series is encircled in the top left area and the other subarea embraces the stainless steel types. There are exceptions due to the fact that some members are quite different from the rest of the family. For instance, one member of the set 836−841 and two members of the set 845−850 were located far away from the rest of their groups. Their chemical compositions are different enough from the rest of the sequence. Steel 837 drops out of the area of the 836−841 set. Its contents of Mn, Si, and V are the highest of the group. Particularly, the content of V is 2.32 times higher than the average of the rest of the group (mean value = 1.31, $s = 0.54$). Similar explanations can be found for steels 845 and 850 by looking at Table 1. To sum up, curve lines surround areas for steels of similar compositions, and these subareas do not overlap between them. So, this classification rapidly indicates as much the differences as the similarities between the objects.

Figure 5 helps us to understand the variable separations through the contour maps. As shown in the figure, when the training has finished, each element of the steel is represented for one of the 11 variable levels of our $8 \times 8 \times 11$ net architecture. The values relative to the maximum weight of the level are represented, indicating the distribution of the most important weight zone. There are drawn only two contour lines, 0.35 and 0.70, of all the possible ones. The light gray areas cover weight values greater than 0.35 of the maximum, being 0.35 exactly in the contour line. The dark gray areas have a contour line of 0.70. In the figure there are superposed letters identifying each steel type on the top map (same as that of Figure 4). The superposition of the 11 levels gives us the 11 weights components that characterize the steel type indicated for the letter on the top map.

Looking at the level corresponding to Cu, it can be seen that this element does not help so much in the steels

separation. Its area involves many classes of steels which superpose with areas of other levels. Instead, the rest of the levels show reduced areas that superposed only partially with others and determine the possibility of a successful classification.

It should be highlighted that during the learning stage there are zero errors on the objects, so an exactitude of 100% is obtained in the classification. This means that each output cell is repeatedly excited by only one particular steel.

For the prediction stage we used another similar set of data containing 95 samples. Now, the weights are kept constant with the values obtained during the training stage, so during the prediction stage the input order of the samples has no importance. Again, there were zero errors for prediction. The mapping of the 8×8 net surface is exactly the same as that for the learning stage. The Kohonen method is specifically appropriate for solving classification tasks. Other ANN techniques, for instance, the feed-forward method, could perhaps solve this problem with the same exactitude,[9] but it is known beforehand that its data processing is too slow to be used as an "on-line" method when it is coupled with an automatic instrument. These results showed that classification of steels can be done in a rapid and safe manner with the ANN technique. The Kohonen ANN was applied without combination of any other mathematical or logical method. Its simplicity offers an excellent program running time for coupling it to an EDX equipment and for processing the data in real time.

ACKNOWLEDGMENT

This work was performed as part of CNEA-CAC-UAQ Projects 95-Q-02-01, 96-Q-02-06, and 96-Q-05-01. Thanks are expressed to Professor Jure Zupan for his advise and revision of the manuscript. We also thank Ariel Verboso from Unidad de Combustibles Nucleares for helping us with the EDX measurements.

REFERENCES AND NOTES

(1)) Bertin, E. P. Principles and Practice of X-ray Spectrometric Analysis, 2nd ed.; Plennum Press: New York, 1975.

(2) LiMin, Fu. *Neural Networks in Computer Intelligence*; McGraw-Hills Series in Computer Sciences, International Edition; McGraw-Hill: Singapore, 1994.

(3) Zupan, J.; Gasteiger, J. *Neural Networks for Chemists*; VCH Publishers: New York, 1993.

(4) Veelenturf, L. P. J. *Analysis and Applications of Artificial Neural Networks*; Prentice Hall: London, 1995.

(5) Ruisanchez, I.; Potokar, P.; Zupan, J. Classification of Energy Dispersion X-ray Spectra of Mineralogical Samples by Artificial Neural Networks. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 214−220.

(6) Van Espen, P.; He, F. *AXIL, Quantitative X-ray Fluorescence Analysis*; Technical Note; Antwerp (Wilrijk)-Belgium.

(7) Van Espen, P.; Janssens, K.; Swenters, I. *AXIL, X-ray Analysis Software Users Manual*; Antwerp (Wilrijk)-Belgium, 1990.

(8) Rozycki; Cezary. Sample Classification in the Case of Known Classes. Use of the X-ray fluorescence spectra. *Chem. Anal. (Warsaw)* **1986**, *31* (3−6), 751−761.

(9) Cleij, P.; Hoogergrugge, R. Linear Data Projection Using a Feedforward Neural Network. *Anal. Chim. Acta* **1997**, *348*, 495−501.

CI9701143