

## APLICAÇÃO DE ALGUNS MODELOS QUIMIOMÉTRICOS À ESPECTROSCOPIA DE FLUORESCÊNCIA DE RAIOS-X DE ENERGIA DISPERSIVA

Fernando Schimidt, Maria Izabel M. S. Bueno\* e Ronei J. Poppi

Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13081-970 Campinas - SP

Recebido em 21/8/01; aceito em 27/2/02

APPLICATION OF SOME CHEMOMETRIC METHODS TO ENERGY DISPERSIVE X-RAY FLUORESCENCE SPECTROMETRY. The objective of this work was to accomplish the simultaneous determination of some chemical elements by Energy Dispersive X-ray Fluorescence (EDXRF) Spectroscopy through multivariate calibration in several sample types. The multivariate calibration models were: Back Propagation neural network, Levenberg-Marquardt neural network and Radial Basis Function neural network, fuzzy modeling and Partial Least Squares Regression. The samples were soil standards, plant standards, and mixtures of lead and sulfur salts diluted in silica. The smallest Root Mean Square errors (RMS) were obtained with Back Propagation neural networks, which solved main EDXRF problems in a better way.

keywords: chemometrics; multivariate calibration; X-ray fluorescence.

### INTRODUÇÃO

A fluorescência de raios-X é uma técnica espectroscópica de análise multielementar muito versátil, podendo ser aplicada a amostras sólidas e líquidas, sem a necessidade de uma preparação complexa. Nesta técnica, o tratamento de dados é usualmente feito através de curvas de calibração univariadas<sup>1,2</sup>, onde se integram determinadas linhas espectrais correspondentes às concentrações dos elementos de interesse. Quando se deseja relacionar mais de uma linha espectral ao mesmo tempo é, então, necessária a utilização de outros métodos matemáticos mais sofisticados, como por exemplo, modelos quimiométricos<sup>3</sup>.

As intensidades medidas em fluorescência de raios-X geralmente não são proporcionais às concentrações das espécies presentes em uma amostra, devido aos efeitos de matriz<sup>1,2</sup>. Sem considerar as amostras não-homogêneas e as particuladas, isto ocorre quando a emissão de raios-X do analito é significativamente afetada pelas variações de concentração dos outros elementos na amostra. Estes efeitos são denominados interelementares e dependem das frações de massa dos elementos na amostra e da relação entre seus coeficientes de absorção para radiação primária e secundária, sendo que neste último caso, ocorre o aparecimento da fluorescência secundária<sup>1,2</sup> (Figura 1). Neste caso, considerando-se uma amostra homogênea que contenha dois elementos químicos A e B na *mesma proporção* e com coeficientes

de absorção de massa semelhantes, ocorre a fluorescência secundária quando o elemento B recebe fótons de raios-X simultaneamente da fonte e do elemento A, sendo que o espectro desta amostra terá a linha espectral de B bem maior do que a de A<sup>4</sup>.

Um outro grande problema em análise multielementar por EDXRF é a sobreposição de linhas espectrais dos elementos, característica comum a elementos vizinhos na Tabela Periódica, em função das proximidades das energias de excitação dos elétrons. Todos estes efeitos dificultam muito a modelagem de espectros, tornando necessária a utilização de métodos matemáticos mais complexos para a sua resolução<sup>1-3</sup>.

### Redes neurais

Recentes avanços em neurofisiologia e novas técnicas experimentais têm proporcionado um aumento do conhecimento da anatomia e funcionamento do cérebro humano. Valendo-se destes novos conhecimentos, modelos matemáticos e algoritmos computacionais têm sido projetados para simular este funcionamento<sup>5</sup>. Assim, por definição, *redes neurais* é o nome dado ao conjunto de métodos matemáticos e algoritmos computacionais desenvolvidos no sentido de simular o processamento de informação e aquisição do conhecimento pelo cérebro humano<sup>3,6</sup>. A rede neural pode ser considerada como uma “caixa de processamento”, que pode aceitar uma série de dados de entrada e produzir uma ou mais saídas, como é representado na Figura 2.

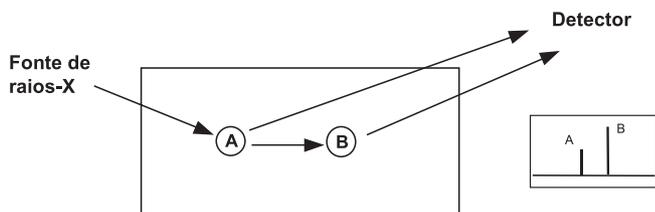


Figura 1. Representação simplificada da fluorescência secundária

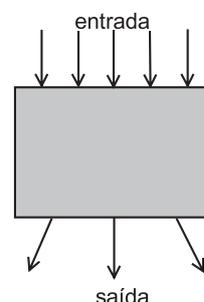
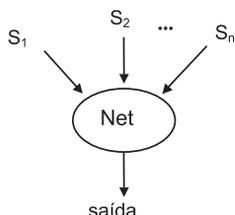


Figura 2. Representação da rede neural como uma “caixa de processamento”, mostrando a entrada e a saída de dados

\*e-mail: bell@iqm.unicamp.br

Como as redes neurais podem ser utilizadas para diversas aplicações, os dados de entrada da Figura 2 poderiam ser de um espectro de um composto, dados clínicos de um paciente, ou ainda, dados de uma imagem manchada de um objeto. A saída poderia ser, respectivamente, a estrutura molecular do composto, o diagnóstico para a doença ou a imagem nítida do objeto<sup>5</sup>. A rede neural imita os neurônios humanos, recebendo e enviando sinais. Ela pode aceitar uma série de dados de entrada e produzir a partir destes uma ou mais saídas, através da busca de uma função que descreva o sistema. Esta analogia feita com os neurônios humanos é baseada na transmissão de sinais elétricos. De maneira simplificada, o neurônio humano pode ser entendido como um corpo celular com duas ramificações: dendritos e axônios. A transferência de sinais elétricos é feita através do axônio de um neurônio para os dendritos de um outro neurônio vizinho, em um processo chamado sinapse.

Estas sinapses representam barreiras que, de um certo modo, modulam os sinais que passam através delas. Esta modulação representa a *força* da sinapse. Em um neurônio matemático, a força da sinapse é chamada de peso  $w$ . Como os neurônios possuem um grande número de dendritos, eles podem receber sinais de vários neurônios simultaneamente. O sinal recebido pelo  $i$ -ésimo dendrito de um neurônio será chamado de  $s_i$  com seu respectivo peso  $w_i$ . O sinal total *Net*, que entra em um neurônio, será dado pela soma de todos os produtos de  $s_i$  com seus respectivos pesos  $w_i$ :



$$\text{Net} = w_1s_1 + w_2s_2 + \dots + w_1s_1 + \dots + w_ns_n \quad (1)$$

O sinal total refere-se apenas à entrada de dados em um neurônio. A saída de um neurônio é uma função da entrada, pela qual a saída é calculada, denominada *função de transferência*.

Para a aprendizagem de uma rede neural, é necessário utilizar um conjunto de calibração. Em função do tipo do algoritmo escolhido, os dados de entrada induzirão a correção dos pesos matemáticos em cada neurônio e em cada camada. A saída calculada pela rede será comparada com a do conjunto de calibração, e em função do erro estipulado, todos os pesos matemáticos serão novamente corrigidos até que a saída forneça um resultado estabelecido. O erro máximo estipulado, o número de iterações e a velocidade de treinamento são parâmetros que são definidos pelo usuário antes de iniciar o algoritmo da rede. A etapa de treinamento de uma rede neural encerra-se quando o erro estipulado atingir o critério de convergência previamente estabelecido.

O procedimento chamado de *otimização da arquitetura da rede neural*, realizado após a etapa de treinamento, consiste na observação do erro RMS no treinamento da rede neural, em função da adição de neurônios à camada intermediária. O ponto onde ocorre o menor erro é considerado como a estrutura ideal de treinamento da rede para o conjunto de calibração utilizado. Isto permite que o algoritmo da rede possa ser utilizado em outros conjuntos de espectros de mesma natureza, sem o risco de haver um *super-ajuste* (do inglês, "overfitting") ou um *sub-ajuste* (do inglês, "underfitting").

Os algoritmos estudados neste trabalho são o "Back Propagation" (BP), "Levenberg-Marquardt" (LM), sendo que ambos diferem entre si na metodologia da correção de pesos, e por último, o "Radial Basis Functions" (RBF), que é um outro tipo de rede neural.

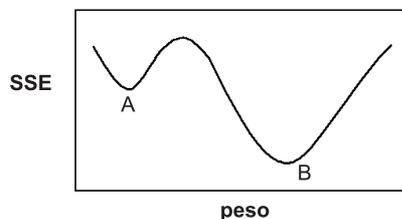
## Rede neural BP

A rede BP é um algoritmo para a correção de pesos com várias camadas. Cerca de 90% das aplicações de redes neurais em química empregam este método<sup>4</sup>. Uma das principais aplicações do BP é o modelamento de dados não lineares, onde o comportamento matemático entre as variáveis não é representado por funções de ordem definida. A principal característica desta rede neural é a equação de correção de pesos, que pode ser escrita resumidamente como:

$$\Delta w^k = \eta \cdot D + \mu \cdot \Delta w^{k-1} \quad (2)$$

onde:

- $\Delta w^k$  representa a diferença entre os pesos inicial e final para a iteração  $k$ ;
- $\eta$  é chamado de "learning rate" ou velocidade de treinamento. Sua função é ponderar as correções dos pesos, de modo a evitar grandes correções, principalmente no final do processo, quando boa parte dos pesos já foi corrigida e uma nova correção muito acentuada pode comprometer todas as correções já efetuadas;
- $D$  é a derivada do erro calculado através das funções de transferência escolhidas, conhecida como *método do gradiente descendente*, e fornece a direção que se deve caminhar na superfície de erros em direção ao mínimo;
- $\mu$  é chamada de *momentum*, uma constante que é multiplicada pela variação de pesos da iteração anterior ( $\Delta w^{k-1}$ ); estabelece uma relação matemática com a variação de pesos da iteração corrente, promovendo assim um acréscimo no valor numérico de  $\Delta w$  obtido no 1º termo da equação 2 ( $\eta \cdot D$ ). A adição do 2º termo à equação 2 ( $\mu \cdot \Delta w^{k-1}$ ) evita que ocorram convergências para mínimos locais, pois o mínimo obtido pelo primeiro termo da mesma equação pode não ser realmente global. Se o mínimo for local, o processo de busca pelo menor erro prossegue (Figura 3). Resumindo, esta constante  $\mu$  evita a possibilidade de se achar um mínimo local;
- $\Delta w^{k-1}$  representa a diferença entre os pesos inicial e final para a iteração  $k-1$ .



**Figura 3.** Representação gráfica da busca do mínimo global. A variação do erro SSE em função dos valores dos pesos gera um gráfico com um comportamento de "altos" e "baixos", como se fossem montanhas e vales; o mínimo é então representado por um "vale". A posição **A** na figura acima representa um mínimo local, e a posição **B** representa um mínimo global

## Rede neural LM

O algoritmo LM faz a correção de pesos através da aproximação do método de Gauss-Newton, chamada de Levenberg-Marquardt<sup>7</sup>. Resumidamente, a equação de correção de pesos pode ser escrita como:

$$\Delta w = (J^T J + \alpha \cdot I)^{-1} \cdot J^T \cdot \varepsilon \quad (3)$$

onde

- $\Delta w$  representa a diferença entre os pesos inicial e final;
- $\alpha$  é um escalar que controla a derivação dos erros, permitindo que o termo ( $J^T J$ ) possa ser invertido;

- $J$  é chamado de jacobiano da matriz derivada dos erros. Cada elemento desta matriz representa uma derivada parcial de um elemento da matriz de erros com o seu correspondente peso;
- $I$  é a matriz identidade. Como ela é multiplicada pela constante  $\alpha$ , pode-se gerar uma nova matriz contendo como elementos apenas valores de  $\alpha$ ;
- $\varepsilon$  é um vetor de erros calculados.

### Rede neural RBF

A rede RBF usa funções de base radial como funções de transferência entre as camadas. A mais conhecida RBF é a gaussiana, apresentando a forma típica  $f(x) = e^{-x^2}$ . A rede RBF é treinada através da adição de neurônios na camada intermediária, até que se encontre o menor erro de calibração. O processo de treinamento é feito através do ajuste dos parâmetros das funções de base radial, os quais são a posição do centro das funções e o raio das mesmas<sup>8</sup>. O algoritmo desta rede utiliza neurônios com a função RBF:

$$f(n,b) = e^{-(b-n)^2} \quad (4)$$

sendo

$$n = \left[ \sqrt{(p-w)^2} \right] \cdot b \quad (5)$$

onde  $p$  é o vetor de entrada,  $w$  é um vetor de pesos, e  $b$  é uma constante.

Neste algoritmo, é calculada a distância geométrica entre o vetor de entrada  $p$  e o vetor de pesos  $w$ , sendo então multiplicada pela constante  $b$ , que está diretamente relacionada com o raio de influência  $r$  (Figura 4). O ajuste de  $r$  é feito manualmente, em função do conjunto de calibração escolhido.

Este termo *raio de influência* diz respeito à distribuição ou ao formato da gaussiana, e está diretamente relacionado à distância entre o pico da função e a sua meia altura<sup>8</sup>. A otimização da arquitetura é o processo de busca do menor erro em função da adição de neurônios à camada intermediária. Neste algoritmo, esta otimização é feita sem interferência externa, através da adição destes neurônios até encontrar-se o menor erro preestabelecido. Isto pode ser considerado como

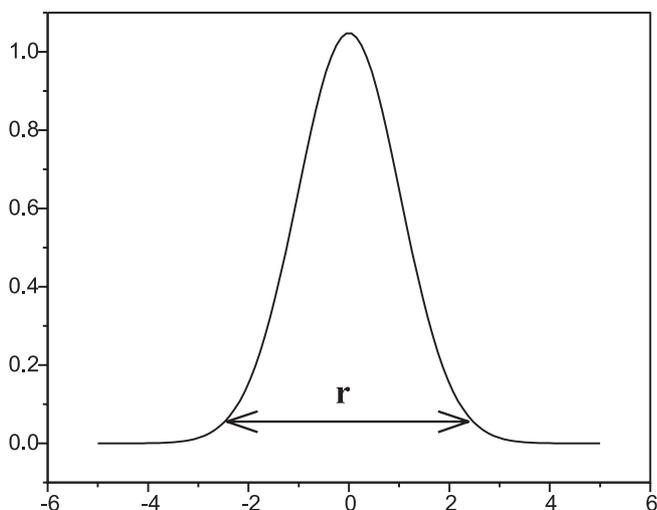


Figura 4. Representação do raio de influência  $r$  de uma função gaussiana RBF

uma vantagem deste tipo de rede, uma vez que, para os outros dois algoritmos, é necessário este procedimento de adição de neurônios.

A variação do parâmetro *raio de influência* (também chamado de “spread”) é um número escolhido na região entre 0,01 e 1,5. Esta faixa é determinada experimentalmente através de vários testes e representa os limites inferior e superior, respectivamente, em que a forma da função gaussiana pode ser alterada e produzir os menores valores dos erros de treinamento para a estimativa das concentrações.

### Modelo PLS

Este modelo faz uso de dois tipos de matriz de dados,  $\mathbf{R}$ , que é a matriz de respostas instrumentais (por exemplo, áreas de um espectro ou de um cromatograma) e  $\mathbf{C}$ , que é a matriz de concentrações<sup>9,10</sup>. O modelo PLS estima matrizes de “scores” e “loadings” tanto para  $\mathbf{R}$  como para  $\mathbf{C}$ :

$$\mathbf{R} = \mathbf{T} \mathbf{P}' + \mathbf{E} = \sum \mathbf{t}_h \mathbf{p}'_h + \mathbf{E} \quad (6)$$

$$\mathbf{C} = \mathbf{U} \mathbf{Q}' + \mathbf{F} = \sum \mathbf{u}_h \mathbf{q}'_h + \mathbf{F} \quad (7)$$

Os elementos  $\mathbf{T}$  e  $\mathbf{U}$  das equações 6 e 7 são chamados “scores” das matrizes  $\mathbf{R}$  e  $\mathbf{C}$ ; os elementos  $\mathbf{P}$  e  $\mathbf{Q}$  são chamados de “loadings” destas mesmas matrizes;  $\mathbf{E}$  e  $\mathbf{F}$  são as matrizes de resíduos (não explicados pelo modelo) de  $\mathbf{R}$  e  $\mathbf{C}$ , respectivamente. A correlação entre os dois blocos  $\mathbf{R}$  e  $\mathbf{C}$  é simplesmente uma relação linear obtida pelo coeficiente de regressão linear, tal como descrito na equação 8:

$$\mathbf{u}_h = \mathbf{b}_h \mathbf{t}_h \quad (8)$$

para “ $\mathbf{h}$ ” variáveis latentes, sendo que os valores de  $\mathbf{b}_h$  são agrupados na matriz diagonal  $\mathbf{B}$ , que contém os coeficientes de regressão entre a matriz de “scores”  $\mathbf{U}$  de  $\mathbf{C}$  e a matriz de “scores”  $\mathbf{T}$  de  $\mathbf{R}$ . A melhor relação linear possível entre os “scores” desses dois blocos é obtida através de pequenas rotações das variáveis latentes dos blocos  $\mathbf{R}$  e  $\mathbf{C}$ .

A matriz  $\mathbf{C}$  pode ser calculada de  $\mathbf{u}_h$ ,

$$\mathbf{C} = \mathbf{T} \mathbf{B} \mathbf{Q}' + \mathbf{F} \quad (9)$$

e a concentração de novas amostras previstas a partir dos novos “scores”,  $\mathbf{T}^*$ , substituídos na equação (9), o que fornece

$$\mathbf{C} = \mathbf{T}^* \mathbf{B} \mathbf{Q}' \quad (10)$$

Neste trabalho, o modelo PLS foi utilizado apenas para comparação dos resultados. Esta ferramenta matemática é bastante conhecida e possui uma vasta gama de aplicações na literatura, sendo muito simples de ser utilizada<sup>9,10</sup>.

### Modelo Fuzzy

A teoria fuzzy provém da teoria avançada de conjuntos e foi desenvolvida por Lotfi Zadeh na década de 60. A aplicação de tratamento fuzzy a um sistema permite interpretar valores em um conjunto de dados de entrada e estabelecer correlações com um conjunto de saída, através de regras condicionais do tipo *se-então*<sup>11</sup>. A definição desta base de regras é a parte mais importante da construção do modelo.

Neste modelo, os dados de entrada e saída são associados a funções matemáticas fuzzy conhecidas como funções de pertinência. No caso específico deste trabalho, os dados de entrada foram representa-

dos por funções gaussianas e os de saída, por funções lineares<sup>12</sup>.

A base de regras para o modelo fuzzy é construída através da aplicação do algoritmo de estimação de grupos (“cluster estimation”)<sup>13</sup>, que é baseado na solução de uma equação matricial algébrica. Para um sistema com duas entradas (por exemplo, dois picos cromatográficos) e uma saída (por exemplo, a concentração de uma espécie), a equação matricial é escrita como:

$$\mathbf{C} = \mathbf{G} \mathbf{R} + \mathbf{H} \quad (11)$$

Onde  $\mathbf{C}$  é a matriz de dados de saída,  $\mathbf{R}$  é a matriz de dados de entrada,  $\mathbf{G}$  e  $\mathbf{H}$  são matrizes de constantes. Sua estrutura condicional fuzzy pode ser escrita como:

*Se  $r_1$  é igual a  $A_1$  e  $r_2$  é igual a  $A_2$  então  $c_1$  é igual a  $B_1$*

Onde  $r_i$  são variáveis de entrada,  $c_i$  são variáveis de saída,  $A_i$  são funções de pertinência do tipo gaussianas e  $B_i$  são funções de pertinência do tipo lineares. Através de um conjunto de calibração são determinados os valores das matrizes  $\mathbf{G}$  e  $\mathbf{H}$  e pode-se estimar as concentrações das espécies em  $\mathbf{C}^{13}$ .

## PARTE EXPERIMENTAL

Neste trabalho, os dados de entrada apresentam-se como dois conjuntos diferentes: “scores”, obtidos através da aplicação de Análise de Componentes Principais (PCA) aos espectros, e *integrais*, pois, nos espectros de EDXRF, as linhas espectrais correspondentes às emissões características dos elementos permanecem sempre em uma mesma região para um determinado elemento químico, independentemente dos diferentes tipos de amostras. As integrais representam a área sob o pico e mostram uma certa proporcionalidade da concentração de um elemento na amostra, mesmo que esta seja afetada pelos efeitos de matriz e interelementares.

A aplicação dos modelos fuzzy neste trabalho foi feita através da utilização de funções específicas para este fim, presentes no “Fuzzy Logic Toolbox”<sup>14</sup>. Esta aplicação possui algumas características: as funções de pertinência de entrada são sempre do tipo gaussianas, as de saída são sempre do tipo lineares e o método de defuzzificação utilizado é uma média ponderada entre os valores obtidos das funções de saída. Todos os resultados obtidos com este modelo são apresentados em uma forma gráfica, em que se verifica a variação do erro RMS em relação ao parâmetro chamado também de raio de influência, sendo este último um número que varia entre 0 e 1 e deve ser ajustado.

Todos os resultados obtidos com as redes neurais BP e LM (cálculo das concentrações) foram feitos em triplicatas, pois a escolha dos pesos iniciais para os neurônios em cada camada é feita, pelo “Toolbox do Matlab<sup>TM</sup>”, através de uma função aleatória. Devido a isto, o primeiro cálculo de erro de treinamento feito pela rede pode gerar um resultado ruim, então todo o procedimento é repetido mais duas vezes.

A escolha do número de “scores” como conjunto de dados de entrada dos modelos foi feita e determinada através da aplicação do algoritmo “Validação Cruzada” (VC), variante “deixa-um-fora”<sup>10,15</sup>.

Foram utilizados padrões certificados de solos de várias partes do mundo fornecidos pelo Instituto de Geociências da Unicamp. Os padrões de diversos tipos de plantas (a maior parte de comestíveis) foram fornecidos pelo Instituto Agrônomo de Campinas.

Os resultados foram avaliados através do cálculo do erro RMS (“Root Mean Square”) ou erro quadrático médio, para as concentrações normalizadas (entre 0 e 1) dos elementos, segundo a equação<sup>16</sup>:

$$RMS = \sqrt{\frac{\sum_{i=1}^n (C_i - \hat{C}_i)^2}{n}} \quad (12)$$

onde,

$C_i$  é a concentração verdadeira das amostras de previsão;

$\hat{C}_i$  é a concentração calculada das amostras de previsão;

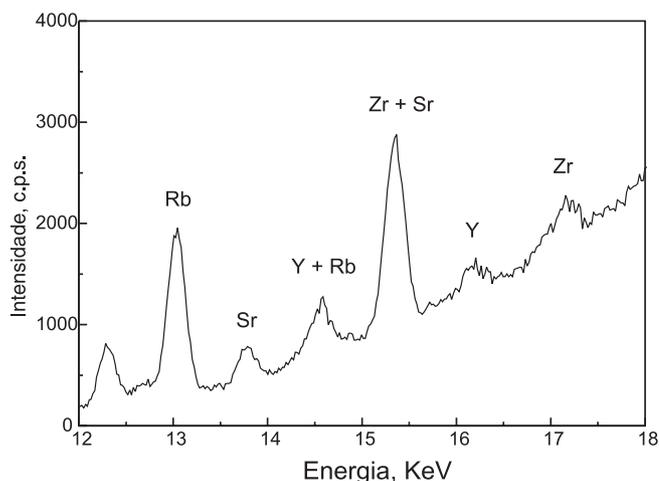
$n$  é o número de amostras de previsão.

A equação 12 é bastante difundida em trabalhos de Quimiometria<sup>16</sup> e é uma representação dos desvios para as concentrações, calculados para um conjunto de previsão.

## RESULTADOS E DISCUSSÃO

### Amostras de solos

A utilização de padrões certificados de solos foi feita com o objetivo de verificar a eficiência da análise multivariada na determinação simultânea de elementos que são vizinhos na Tabela Periódica, especificamente Rb, Sr, Y e Zr. A Figura 5 mostra um espectro típico de EDXRF com a sobreposição das linhas espectrais da região de interesse desses 4 metais. Para estes elementos ocorrem sobreposições das linhas  $K\alpha$  do Y com  $K\beta$  do Rb e  $K\alpha$  do Zr com  $K\beta$  do Sr. Este tipo de amostra foi escolhido por tratar-se de uma amostra natural complexa, sendo possível a observação nos espectros de linhas de mais de 10 elementos químicos detectáveis, sendo todos grandes fontes de fluorescência secundária. Nenhum tipo de pré-tratamento foi feito nestes espectros para redução ou eliminação dos ruídos; foi feita apenas uma subtração de uma amostra “branco” a fim de se ajustar à linha base. Esta amostra utilizada como branco foi uma amostra de solo em que as concentrações dos quatro elementos eram muito pequenas.



**Figura 5.** Parte de um espectro de EDXRF típico, mostrando as linhas espectrais sobrepostas dos quatro elementos estudados

Especificamente para estas amostras de solos, a aplicação dos modelos foi feita considerando-se duas situações: (1) espectro completo (de 1 a 40 keV) e (2) região correspondente às linhas dos 4 elementos. Em função de resultados obtidos em experimentos semelhantes com este tipo de amostra, optou-se por verificar se existe contribuição das linhas espectrais dos outros elementos na estimativa das concentrações dos 4 referidos metais.

Primeiramente, considerou-se a região onde aparecem as principais linhas espectrais situadas entre 12,8 e 15,9 KeV, que apresentaram uma variação significativa. Na análise VC dos espectros das amostras, foi determinado um número de 4 componentes principais representado por 4 “scores”. A análise mostrou uma variância capturada total de 99,78% para o bloco dos espectros e de 69,50% para o bloco das concentrações. As integrais foram calculadas apenas para as 4 linhas espectrais situadas entre 12,8 e 15,9 KeV, uma vez que as linhas  $K\beta$  isoladas de Y e Zr são muito pequenas em todos os espectros. Foram utilizados 48 espectros de padrões, separando-se 34 para calibração, 10 para previsão, e foram retirados 4 “outliers”. Os resultados das aplicações dos modelos estão na Tabela 1.

**Tabela 1.** Erros RMS calculados para os modelos utilizados para amostras de solos, considerando-se o número de “scores” igual a 4 e o número de integrais igual a 4, para a região correspondente as quatro linhas

	“scores”	integrais
rede neural BP	0,0741	0,0650
rede neural LM	0,0800	0,0850
rede neural RBF	0,177	0,0893
Modelo Fuzzy	0,107	0,0790
PLS	0,107	0,113

As redes neurais são treinadas adicionando-se neurônios à camada intermediária (de 1 a 7 neurônios) e testando um conjunto de validação. Analisam-se então os resultados dos erros RMS para todos os elementos, obtendo-se um erro médio. O ponto onde o erro é o mais baixo considera-se como sendo a configuração ideal da rede. Este procedimento permite a melhor escolha de configuração evitando que ocorra o sub-ajuste ou o super-ajuste da rede.

A rede BP (com arquitetura otimizada 4-3-4) apresentou o menor erro RMS quando foram utilizadas as integrais como conjunto de entrada. Um teste F ao nível de 95% de confiança mostra uma diferença significativa entre os resultados obtidos pela rede BP e os obtido pelo modelo PLS, para os dois tipos de dados de entrada (considerando-se  $n = 44$ ,  $F_{crit} = 1,35$ )<sup>17</sup>.

No segundo tipo de aplicação, considerando o espectro completo, a análise VC das amostras determinou um número de 6 componentes principais representado por 6 “scores”. A análise mostrou uma variância capturada total de 99,88% para o bloco dos espectros e de 98,21% para o bloco das concentrações. O aumento da variância capturada para as concentrações, em relação à aplicação anterior, já mostra que realmente existe uma maior correlação entre os dados de entrada e saída. Foram utilizados os mesmos conjuntos de dados, 48 espectros de padrões, separando-se 34 para calibração, 10 para previsão, e foram retirados 4 “outliers”. Para esta aplicação não foram utilizadas as integrais comparativamente. Os resultados das aplicações dos modelos estão na Tabela 2.

**Tabela 2.** Erros RMS calculados para os modelos utilizados para amostras de solos, considerando-se o número de “scores” igual a 6, para o espectro completo

	“scores”
rede neural BP	0,0346
rede neural LM	0,0319
rede neural RBF	0,0598
modelo Fuzzy	0,0452
PLS	0,0643

A rede LM (com arquitetura 6-3-4) apresentou o menor erro RMS para esta segunda aplicação. Um teste F ao nível de 95% de confiança mostra uma diferença significativa entre os resultados obtidos pelas redes BP e LM e os obtidos pelo modelo PLS ( $n = 44$ ,  $F_{crit} = 1,35$ )<sup>17</sup>. O melhor resultado obtido pela rede neural LM (0,0319) utilizando todo o espectro é significativamente melhor que o obtido somente com a região das 4 linhas com a rede BP (0,0650) segundo o teste F:

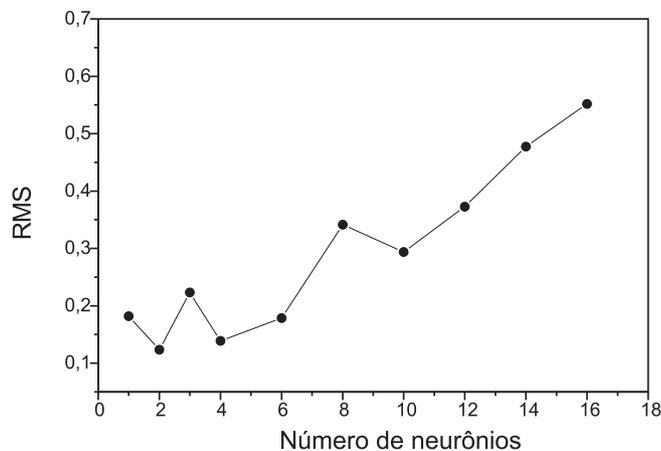
$$F = \frac{(RMS_{BP})^2}{(RMS_{LM})^2} = \frac{(0,0650)^2}{(0,0319)^2} = 4,15$$

A Tabela 2 mostra uma diminuição, de maneira geral, nos erros encontrados para todos os modelos quando se aplica a PCA a todo o espectro, em relação à aplicação da região das 4 linhas (Tabela 1). Pode-se concluir que existe, para este caso, uma contribuição das outras linhas espectrais de outros elementos no processo de estimação das concentrações do Rb, Sr, Y e Zr, ou seja, a caracterização da fluorescência secundária. As linhas espectrais dos outros elementos presentes neste tipo de amostra interferem na correlação entre os dados de entrada e saída para os quatro elementos estudados.

Os erros relativos calculados para as concentrações obtidas pela rede LM, que foi o melhor modelo, são menores que 20%. Neste erro está embutida a existência da fluorescência secundária (a maior contribuição), problemas na preparação das amostras, além dos erros instrumentais nas medidas.

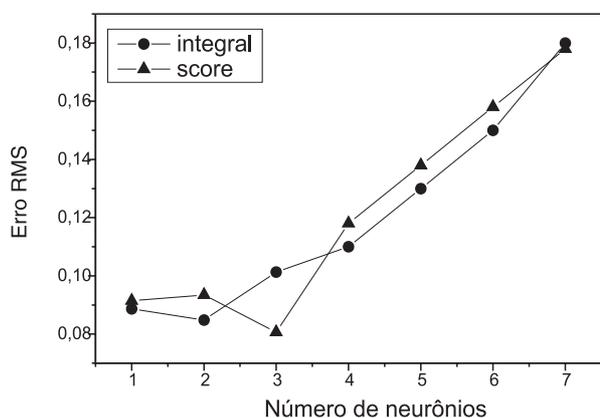
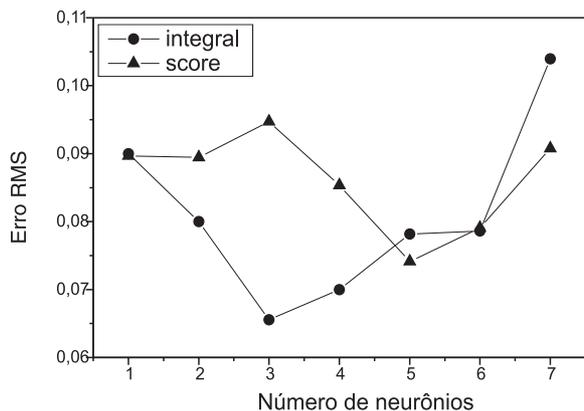
Para estas amostras de solos, apenas a rede neural BP foi aplicada também de uma terceira maneira distinta, em relação ao conjunto de dados de entrada, em que foram utilizados diretamente pontos dos espectros com os respectivos valores de intensidade de emissão, ou seja, 131 pontos (131 neurônios na camada de entrada), correspondentes à região entre 12,8 keV e 15,9 keV dos espectros. Foram adicionados neurônios na camada intermediária de 1 até 16.

Os resultados obtidos para esta aplicação são mostrados na Figura 6. Eles não foram muito bons, mostrando um erro maior do que os que estão mostrados na Tabela 2. Este estudo já havia sido feito com outros conjuntos de amostras e os resultados foram igualmente inferiores em relação ao uso dos “scores” ou integrais como dados de entrada. Provavelmente para um conjunto muito grande de dados de entrada como este (131 pontos), a rede acaba também modelando ruídos e/ou informação que não é relevante. A rede neural utilizada também foi otimizada, através da adição de neurônios à camada intermediária.



**Figura 6.** Erros RMS calculados contra o número de neurônios na camada intermediária. A melhor configuração da rede BP é 131-2-4. O menor erro é de 0,123, bem superior ao encontrado utilizando-se “scores” ou integrais na entrada

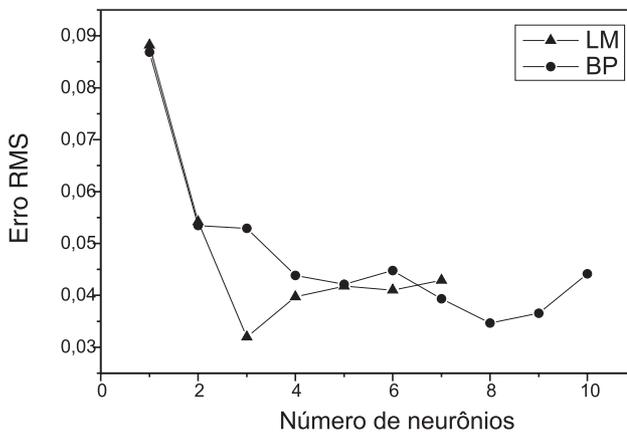
Os resultados da aplicação das redes neurais para *scores* e integrais como dados de entrada para a região dos quatro picos apresentam um comportamento muito semelhante ao observado na Figura 6, e estão mostrados na Figura 7.



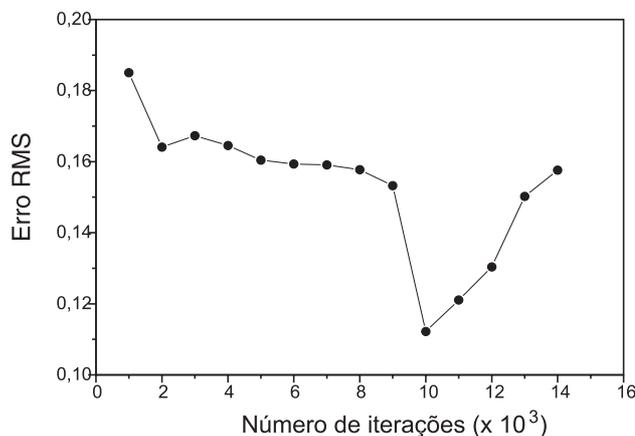
**Figura 7.** Erros RMS calculados contra o número de neurônios na camada intermediária para as redes BP e LM, utilizando-se "scores" e integrais como conjunto de dados de entrada, considerando-se a região das linhas dos 4 metais

Existe um padrão de comportamento em que o erro começa alto para um neurônio na camada intermediária, depois ele diminui até o número ótimo e então volta a crescer. Pode-se que concluir que este comportamento mostra que as redes neurais passam pelos estados de sub-ajuste e super-ajuste, antes e depois de atingirem a melhor configuração, respectivamente.

Este tipo de comportamento foi observado para todas as redes neurais, para todos os diferentes tipos de amostras utilizados neste trabalho. Experimentos realizados com redes neurais BP para modelar espectros em vários tipos de amostras mostraram que o número ótimo de iterações é 10000. Este é um dos parâmetros que deve ser escolhido para iniciar tanto a rede BP com a LM e deve ser testado também, em função do tipo de amostra. Após obtido este número, a possibilidade de ocorrer super-ajuste é muito grande. O seu critério de escolha deve ser feito com muito cuidado, devendo ser testados diversos valores a fim de obter-se o menor erro de treinamento da rede neural. Como exemplo, foi feito um estudo da variação do erro RMS com o número de iterações para a rede neural BP, apresentado na Figura 6 e de configuração 131-2-4, o resultado está mostrado na Figura 9. O comportamento da variação do erro apresentado nesta figura ocorre também de forma semelhante, quando se utiliza os



**Figura 8.** Erros RMS calculados contra o número de neurônios na camada intermediária para as redes BP e LM, utilizando-se "scores" como conjunto de dados de entrada, considerando-se o espectro completo



**Figura 9.** Variação do erro RMS com o número de iterações, para uma rede neural 131-2-4

"scores" ou as integrais como conjunto de dados de entrada, ou seja, independente do tipo de dado de entrada.

### Misturas de sais de chumbo e enxofre

A aplicação dos modelos quimiométricos para este tipo de amostra foi feita através de 38 misturas sintéticas de sais contendo chumbo ( $\text{Pb}(\text{NO}_3)_2$ ) e enxofre ( $\text{K}_2\text{SO}_4$ ) em matriz de sílica. Este sistema é um caso típico de sobreposição de linhas espectrais correspondentes as linhas  $\text{L}\alpha$  do chumbo (10,02 a 10,51 keV) e a sobreposição das linhas  $\text{K}\alpha$  do enxofre e  $\text{M}\alpha$  do chumbo (2,07 a 2,50 keV), como mostrado na Figura 10. Para estas amostras, foram utilizados 24 espectros para calibração, 7 para validação e foram retirados 7 "outliers", para a determinação simultânea de Pb e S.

Na análise VC dos espectros das amostras foi determinado um número de 6 componentes principais representando por 6 "scores". Esta análise mostrou uma variância capturada total de 99,98% para o bloco dos espectros e de 95,50 % para o bloco das concentrações. As integrais foram obtidas de todos os espectros entre as regiões de 2,07 a 2,50 keV, e de 10,02 a 10,51 keV. Os resultados da aplicação dos modelos estão mostrados na Tabela 3.

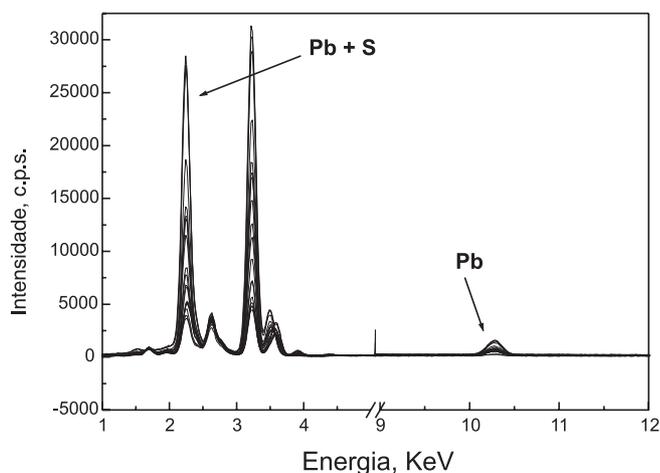


Figura 10. Parte de espectros de EDXRF utilizados neste trabalho, mostrando as linhas espectrais dos elementos Pb e S

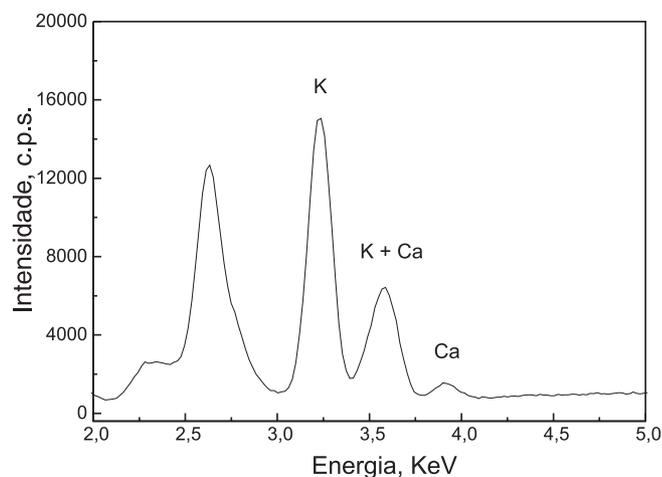


Figura 11. Espectro de EDXRF típico para uma amostra de planta mostrando a sobreposição de linhas espectrais de Ca e K

**Tabela 3.** Erros RMS calculados para os modelos utilizados para misturas de sais de chumbo e enxofre, considerando-se o número de “scores” igual a 6 e o número de integrais igual a 2

	“scores”	integrais
rede neural BP	0,0632	0,0229
rede neural LM	0,0501	0,0297
rede neural RBF	0,0428	0,0472
modelo Fuzzy	0,0943	0,0275
PLS	0,0863	0,0559

De acordo com a Tabela 3, os erros menores foram obtidos através da utilização de integrais para a rede BP com arquitetura da rede 2-4-2. Um teste F aplicado aos valores obtidos pelo modelo fuzzy e pelos modelos PLS em relação à rede BP (integrais), indica haver diferença significativa ao nível de 95% apenas em relação ao segundo modelo (considerando-se neste caso,  $n = 38$  então  $F_{crit.} = 1,71$ )<sup>17</sup>:

$$F_1 = \frac{(RMS_{MF})^2}{(RMS_{BP})^2} = \frac{(0,0275)^2}{(0,0229)^2} = 1,44 \quad F_2 = \frac{(RMS_{PLS})^2}{(RMS_{BP})^2} = 5,96$$

Embora os resultados dos erros RMS para integrais estejam bem próximos, a rede neural BP resolveu melhor os problemas relativos a este tipo de amostra. Isto pode ser comprovado através dos erros relativos calculados para as concentrações, que estão em torno de 4,6%.

O comportamento dos erros RMS para as redes neurais apresenta-se bastante semelhante ao que foi observado para as amostras de solos.

### Amostras de plantas

Um outro estudo foi feito utilizando-se padrões de diversos tipos de plantas com ênfase na determinação simultânea das concentrações de Ca e K. Neste tipo de amostra ocorre a sobreposição das linhas  $K\alpha$  do cálcio com a  $K\beta$  do potássio, conforme mostrado na Figura 11. Assim como as amostras de solos, estas amostras possuem mais de 10 elementos químicos detectáveis nos espectros, contribuindo para o aumento da fluorescência secundária. Também neste caso não foi feito nenhum tipo de pré-tratamento, ou mesmo ajuste de linha base.

De um total de 32 espectros, foram selecionados 24 espectros para calibração, 6 para previsão e foram retirados 2 “outliers”. Na análise VC dos espectros das amostras, foi determinado um número de 10 componentes principais representados por 10 “scores”. A análise mostrou uma variância capturada total de 100% para o bloco dos espectros e de 95,86% para o bloco das concentrações. As integrais foram calculadas para as linhas espectrais situadas entre 3,0 e 4,2 KeV (3 linhas). Os resultados das aplicações dos modelos estão mostrados na Tabela 4.

**Tabela 4.** Erros RMS calculados para os modelos utilizados para amostras de plantas, considerando-se o número de “scores” igual a 10 e o número de integrais igual a 3

	“scores”	Integrais
rede neural BP	0,117	0,0247
rede neural LM	0,133	0,0461
rede neural RBF	0,113	0,0222
modelo Fuzzy	0,135	0,0424
PLS	0,153	0,0507

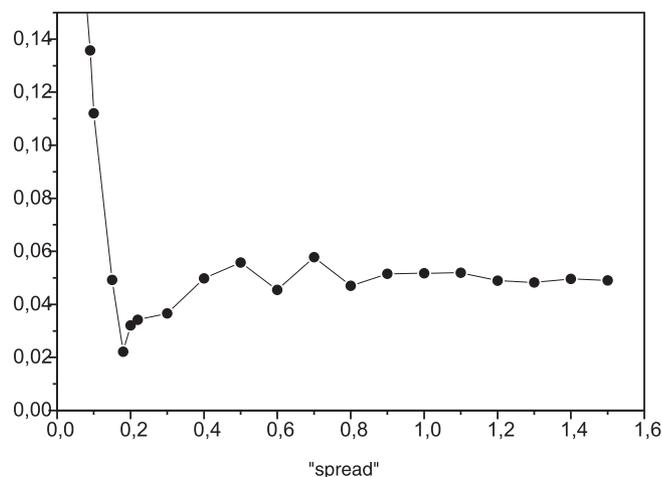


Figura 12. Comportamento do erro RMS em função da adição de neurônios à camada intermediária, para a rede neural RBF

Para este tipo de amostra, a utilização das integrais como dados de entrada foi surpreendentemente melhor do que o uso de "scores". O melhor resultado obtido foi através da aplicação da rede RBF com integral com configuração 3-10-2, embora a rede BP também tenha apresentado um bom resultado. O teste F aplicado entre estes dois resultados mostra que não existe diferença entre eles ( $F_{\text{crit.}} = 1,90$  para  $n = 32$ )<sup>16</sup>; apenas existe diferença em relação ao modelo PLS.

$$F = \frac{(RMS_{BP})^2}{(RMS_{RBF})^2} = \frac{(0,0247)^2}{(0,0222)^2} = 1,24 \quad F = \frac{(RMS_{PLS})^2}{(RMS_{RBF})^2} = \frac{(0,0507)^2}{(0,0222)^2} = 5,22$$

Neste tipo de amostra, os erros relativos calculados para as concentrações para a rede RBF ficaram em torno de 9,4%. A Figura 12 mostra o comportamento do erro para a rede RBF utilizada neste tipo de amostra.

## CONCLUSÃO

Neste trabalho podemos concluir que a aplicação de quimiometria à fluorescência de raios-X, especialmente para análise multielementar, mostrou ser muito importante, pois permite a estimativa das concentrações dos elementos presentes em uma amostra com uma economia de tempo e reagentes gastos para separação dos componentes, em relação às análises tradicionais, embora às vezes com erros um pouco maiores.

De maneira geral, os modelos quimiométricos utilizados resolvem de maneira satisfatória os maiores problemas encontrados na espectroscopia EDXRF (fluorescência secundária e sobreposição de linhas espectrais), desde que o conjunto de calibração escolhido para os modelos seja adequado (espectros que representem as proporções de concentração dos elementos estudados em uma amostra real), sendo este fator de fundamental importância, pois determina a ordem de grandeza dos erros para os modelos.

A utilização de redes BP e LM para a modelagem dos espectros de fluorescência de raios-X mostrou-se mais eficiente para a determinação simultânea de elementos, quando comparada com os outros métodos utilizados também para modelagem de dados. Estas redes BP e LM estabelecem uma melhor correlação entre os conjuntos de

dados de entrada e saída, embora estes modelos apresentem um gasto computacional maior. Para as redes neurais, a observação do erro RMS em função da adição de neurônios às camadas intermediárias permite que se encontre o menor erro de uma maneira segura e eficiente.

As redes neurais RBF e os modelos fuzzy não mostraram um bom desempenho em relação aos outros modelos. Ambos os modelos têm apenas um parâmetro a ser definido para escolha do menor erro de treinamento e possuem uma extensa teoria matemática como fundamento. Seria interessante um estudo mais aprofundado para outras futuras aplicações.

## REFERÊNCIAS

1. Leyden, D.E.; *Fundamentals of X-Ray Spectrometry as Applied to Energy-Dispersive Techniques*, Tracor X-Ray; Mountain View: California, 1984.
2. Bertin, E.P.; *Principles and Practice of X-Ray Spectrometric Analysis*, Plenum Publishing Corporation: New York, 1970.
3. Nagata, N.; Bueno, M. I. M. S.; Peralta-Zamora, P. G.; *Quim. Nova* **2001**, *24*, 531.
4. Schimidt, F.; *Tese de Doutorado*; Universidade Estadual de Campinas, Brasil, 2000.
5. Zuppan, J.; Gasteiger, J.; *Neural Networks for Chemists: An Introduction*, VCH: New York, 1993.
6. Zuppan, J.; Gasteiger, J.; *Anal. Chim. Acta* **1991**, *248*, 1.
7. Facchin, I.; Mello, C.; Bueno, M.I.M.S.; Poppi, R.J.; *X-Ray Spectrom.* **1999**, *28*, 173.
8. Cabral Jr., E.F., ed.; *Redes neurais artificiais : um curso teórico e prático para engenheiros e cientistas*, Editora da USP: São Paulo, 1999.
9. Thomas, E.V.; *Anal. Chem.* **1994**, *66*, 795A.
10. Otto, M.; *Chemometrics : Statistics and Computer Application in Analytical Chemistry*, Wiley-VCH Verlag: Weinheim, 1999.
11. Pedrycz, W.; Gomide, F.; *An Introduction to Fuzzy Sets: Analysis and Design*, MIT Press: Cambridge, 1998.
12. Tagaki, T.; Sugeno, M.; *IEEE Transactions on Systems, Man and Cybernetics* **1985**, *15*, 116.
13. Chiu S.; *Journal of Intelligent & Fuzzy Systems* **1994**, *2*, 267.
14. *Fuzzy Logic Toolbox User's Guide*, versão 2.0, The Mathworks Inc., 1998.
15. Ferreira, M.C.; Antunes, A.M.; Melgo, M.S.; Volpe, P.L.O.; *Quim. Nova* **1999**, *22*, 724.
16. Wang, Z.; Hwang, J.N.; Kovalski, B.R.; *Anal. Chem.* **1995**, *67*, 1497.
17. Skoog, D.A.; West, D.M.; Holler, F.J.; *Fundamentals of Analytical Chemistry*, Saunders College Publishing: Florida, 1992.