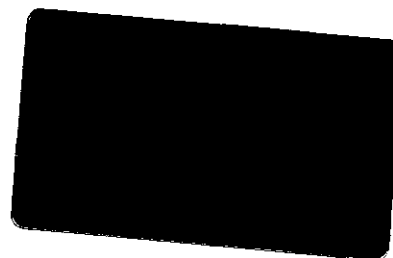


Universidade Estadual de Campinas

Instituto de Química

Tese de Doutorado



*“Aplicação de alguns Modelos Quimiométricos à Espectroscopia de
Fluorescência de Raios-X de Energia Dispersiva”*

Aluno: Fernando Schimidt

Orientadora: Profa. Dra. Maria Izabel M. S. Bueno

Co-orientador: Prof. Dr. Ronei J. Poppi

"A coisa mais bela que o homem pode experimentar é o mistério. É esta a emoção que está na raiz de toda ciência e arte. O homem que desconhece este encanto, incapaz de sentir admiração e estupefação, esse já está, por assim dizer, morto e tem os olhos extintos."

"Albert Einstein"

(1879 - 1955)

"Na ciência, existem questões ingênuas, questões entediantes, questões apresentadas de modo inadequado. Mas cada questão é um grito para entender o mundo. Não existe pergunta estúpida."

"Carl Sagan"

(1934 - 1996)

Agradecimentos

- A Bell, pela orientação, amizade, carinho, dedicação e apoio dado ao longo deste trabalho;
- Ao Ronei, pela co-orientação, pelas críticas construtivas e sugestões que foram muito importantes neste trabalho;
- Aos professores Fernando Gomide (FEE - Unicamp) e Stephen Chiu (Rockwell Science Center), pela ajuda fundamental na aprendizagem e aplicação da Teoria Fuzzy;
- Ao Eng. Paolo Antoni (Instituto Politécnico de Milão) pelas dicas sobre o Fuzzy Logic Toolbox;
- As pessoas que implementaram no IQ a rede Unix e o acesso a Internet e proporcionaram a utilização de uma das ferramentas tecnológicas e de comunicação mais fantásticas dos últimos tempos;
- Aos companheiros de grupo, Noemi, Regina, Helena e Antenor, pelo apoio, amizade e companheirismo;
- As colegas de laboratório, Marta, Aloísia, Vilarinho, Júlio, Anderson, Marcilene, Hyam, Janaína, Gisele, Paulo e Rafael, pela amizade e convivência agradável;
- Aos meus pais, pelo apoio, carinho e compreensão, a eles dedico este trabalho;
- A Mariângela, pelo amor e carinho, que foram fundamentais ao longo deste trabalho;
- A Deus, pela saúde, paciência e pela vida.

Resumo

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, 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 necessário a utilização de outros métodos matemáticos mais sofisticados, como por exemplo modelos quimiométricos.

O objetivo deste trabalho foi o de utilizar alguns modelos quimiométricos, como as redes neurais "Back Propagation" (BP-ANN), "Levenberg-Marquardt" (LM-ANN) e "Radial Basis Function" (RBF-ANN) e também modelagem fuzzy, além dos modelos PLS e PCR (para comparação dos resultados) em dados de fluorescência de raios-X de energia dispersiva (EDXRF). O intuito deste artifício é o de realizar a determinação simultânea de alguns elementos químicos em vários tipos de amostras, principalmente nos casos onde ocorrem fluorescência secundária e sobreposição de linhas espectrais, problemas comuns relacionados ao emprego desta técnica. Um segundo objetivo deste trabalho foi fazer um estudo comparativo entre os modelos, na tentativa de buscar um modelo mais adequado para resolução da maioria dos problemas e também tentar identificar a forma do melhor tipo de conjunto de dados de entrada ("scores" ou integrais) para os modelos. Foram utilizadas amostras com muitos elementos químicos, sem nenhum tipo de pré-tratamento nos dados. Foram elas (a) padrões de solos, (b) padrões de plantas, (c) aço, (d) misturas sintéticas sólidas de óxidos de terras raras e (e) sais de chumbo e enxofre diluídos em sílica.

Os menores erros quadráticos médios ("Root Mean Square" - RMS) obtidos foram: 0,0229 para as misturas de sais de chumbo e enxofre com o uso de BP-ANN; 0,0753 para amostras de solos com RBF-ANN; 0,0222 para amostras de plantas com RBF-ANN; 0,166 para amostras de aços com BP-ANN e 0,200 para amostras de terras raras com LM-ANN. Com relação ao estudo dos tipos de conjuntos de dados de entrada, as integrais apresentaram erros um pouco menores do que os "scores".

Conforme os resultados obtidos, pode-se concluir que os principais problemas de EDXRF são melhores resolvidos com as redes neurais BP-ANN e LM-ANN do que com a rede neural RBF-ANN e também com modelagem fuzzy.

Abstract

X-ray fluorescence spectrometry is a very versatile and multielemental analytical technique that can be applied to solid and liquid samples, without the need of complex preparations. Data treatment is usually done through univariate calibration curves, where some spectral lines corresponding to elemental concentrations are integrated. When it is necessary to relate more than one spectral line at a time, then it is necessary to make use of more complex mathematical methods, such as those presented by chemometrics.

The objective of this work was mainly to apply some chemometric models to Energy Dispersive X-Ray Fluorescence (EDXRF) data. These data involved the simultaneous spectrometric determination of some chemical elements in several sample types, mainly where secondary fluorescence and overlapping of spectral lines are intense, common problems related to EDXRF. A second aim was to perform a comparative study among the models, looking for a more generalized one to solve most EDXRF problems. A third objective was to identify the best type of input data (scores or integrals) for the models. The analyzed samples contained many chemical elements, without any pre-treatment. They were (a) soil standards, (b) plant standards, (c) steels, (d) solid synthetic mixtures such as rare earth oxides and (e) mixtures of lead and sulfur salts diluted in silica.

Artificial neural networks (ANN) in the following modes "Back Propagation" (BP-ANN), "Levenberg-Marquardt" (LM-ANN) and "Radial Basis Function" (RBF-ANN) were evaluated. Besides ANN, fuzzy modeling, Partial Least Square (PLS) and Partial Component Regression (PCR) models were also considered.

The smallest Root Mean Square (RMS) errors obtained were: 0.0229 for lead and sulfur mixtures with BP-ANN; 0.0753 for soil samples with RBF-ANN; 0.0222 for plant samples with RBF-ANN; 0.166 for steel samples with BP-ANN, and 0.200 for rare earth samples with LM-ANN. The use of integrals showed smaller errors than using scores, in the study of types of input data.

The overall result indicated BP-ANN and LM-ANN as the best chemometric tools to solve these problems in EDXRF.

Curriculum

FERNANDO SCHIMIDT

Formação Acadêmica:

- Mestre em Química Analítica pela UNICAMP (1997)
- Bacharel em Química Tecnológica pela UNICAMP (1994)

Produção Científica:

- Trabalho apresentado na forma de poster no 8º COLACRO, Congresso Latino Americano de Cromatografia e Técnicas Afins em 13/04/2000, em Buenos Aires - Argentina, com o título "Quantification of overlapping HPLC peaks in multiresidue pesticides analysis using a neural network".
- Artigo publicado na revista "Arzneimittel-Forschung / Drug Research", como co-autor, com o título "Structure-activity relationship analysis of 4'-bromo-[1,1'-biphenyl] 4-X-phenyl methanone derivatives and activity against Mycobacterium tuberculosis", vol. 49(2), 1999, pag. 1025-1029.
- Foi submetido um artigo a Revista de Educação Tecnológica da Escola Técnica Federal de Goiás, intitulado "O que são Redes Neurais Artificiais", no prelo.
- Dois trabalhos apresentados na forma de poster no 10º Encontro Nacional de Química Analítica em 03/09/99 em Santa Maria - RS, com os títulos "Determinação de alguns metais em solos por fluorescência de raios-x de energia dispersiva com modelagem fuzzy" e "Determinação de Ca e K em plantas por fluorescência de raios-x de energia dispersiva e calibração multivariada".
- Trabalho apresentado na forma de poster no 39º Congresso da Associação Brasileira de Química em 30/09/99 em Goiânia - GO, com o título "Determinação de Ca e K em plantas por fluorescência de raios-x de energia dispersiva e redes neurais".
- Trabalho com apresentação oral e publicado nos Anais do 4º Congresso Brasileiro de Redes Neurais em 20/07/99 em S.José dos Campos - SP, com o título "Determinação de alguns

metais em solos por fluorescência de raios-x de energia dispersiva e redes neurais”.

- Trabalho apresentado na forma oral, com publicação de artigo (ISSN 1515-1565) nos Anais do VI Seminário Latinoamericano de Análisis por Técnicas de Rayos X, em 17/11/98 em Huerta Grande, Província de Córdoba, Argentina, com o título “Determinação de fenóis por EDXRF e redes neurais artificiais”.
- Trabalho apresentado na forma de poster na 21ª Reunião Anual da Sociedade Brasileira de Química em 05/98 em Poços de Caldas - MG, com o título “Determinação de misturas de fenóis por fluorescência de raios-x de energia dispersiva com tratamento dos dados por redes neurais”.
- Participação no VIII Workshop Anual de Usuários do Laboratório Nacional de Luz Síncrotron (LNLS) em 11/97 em Campinas - SP.
- Defesa da dissertação de mestrado em 30/04/97 no IQ-UNICAMP com o título “Desenvolvimento de um programa computacional para tratamento de dados de imagens digitalizadas e sua aplicação em química analítica”.
- Trabalho apresentado na forma de poster na 19ª Reunião Anual da Sociedade Brasileira de Química em 06/96 em Poços de Caldas - MG, com o título “Determinação de pH através de imagens digitalizadas de indicadores imobilizados em papel”.
- Trabalho apresentado na forma de poster na 18ª Reunião Anual da Sociedade Brasileira de Química em 06/95 em Caxambu - MG, com o título “Tratamento de dados de imagens digitalizadas de cromatografia e eletroforese”.

Participação em Cursos:

- 39º Congresso da Associação Brasileira de Química em 09/99 em Goiânia - GO, cursos “Computação aplicada ao ensino superior de química: métodos para solução de problemas químicos” e “Espectroscopia no Infravermelho Próximo (NIR) aplicada à Química Analítica”.
- 5ª Escola de Redes Neurais, S. José dos Campos - SP, em 19/07/99, cursos “Aprendizados por reforço” e “Lógica Nebulosa”.

Índice

Índice de figuras.....	xvii
Índice de tabelas.....	xxi
1. Introdução.....	1
1.1. Absorção e Emissão de Raios-X.....	1
1.2. Espalhamento de Raios-X.....	4
1.3. Instrumentação para EDXRF.....	4
1.4. Efeitos de Matriz.....	6
1.5. Utilização de EDXRF.....	7
1.6. Modelos Quimiométricos.....	9
1.6.1. Modelos PLS e PCR.....	10
1.6.1.1. O método "Cross Validation".....	14
1.6.2. Redes Neurais.....	16
1.6.2.1. Treinamento da rede neural.....	20
1.6.2.2. Rede neural BP.....	21
1.6.2.3. Rede neural LM.....	22
1.6.2.4. Rede neural RBF.....	23
1.6.2.5. "Overfitting" e "Underfitting".....	25
1.6.2.6. Uso de redes neurais em fluorescência de raios-X.....	27
1.6.3. Modelagem Fuzzy.....	30
1.6.3.1. Definição de conjuntos fuzzy.....	31

1.6.3.2. Definição de operadores fuzzy.....	33
1.6.3.3. Definição de regras fuzzy.....	35
1.6.3.4. Defuzzificação.....	35
1.6.3.5. Modelo fuzzy para o "problema da gorjeta".....	36
1.6.3.6. O algoritmo "cluster estimation".....	37
2. Objetivos.....	42
3. Parte experimental.....	43
4. Resultados e discussão.....	55
4.1. Misturas de sais de chumbo e enxofre.....	57
4.2. Amostras de solos.....	63
4.3. Amostras de plantas (K e Ca).....	71
4.4. Amostras de aços.....	76
4.4.1. Reescalonamento.....	81
4.5. Amostras de terras raras.....	88
4.6. Amostras de plantas (Si, P, S, K e Ca).....	92
4.6.1. Reescalonamento.....	92
4.6.2. Determinação de elemento único.....	102
5. Conclusão.....	108
6. Referências bibliográficas.....	111
7. Apêndice.....	113

Índice de figuras

Figura 1 - Gráfico em escala logarítmica do coeficiente de absorção de massa em função do comprimento de onda do raio-X.....	3
Figura 2 - Espectro de EDXRF típico para vários metais pré-concentrados em resina de troca catiônica.....	5
Figura 3 - Tipos de efeitos interelementares em espectroscopia de fluorescência de raios-X.....	7
Figura 4 - Representação simplificada da fluorescência secundária.....	9
Figura 5 - Representação das matrizes R e C.....	11
Figura 6 - Representação gráfica dos "scores" e "loadings".....	11
Figura 7 - Representação da rede neural como uma "caixa de processamento".....	16
Figura 8 - Representação dos neurônios.....	17
Figura 9 - Representações simplificadas de um neurônio biológico e de uma sinapse.....	18
Figura 10 - Representação da sinapse de um neurônio matemático.....	18
Figura 11 - Funções de transferência usadas em redes neurais.....	19
Figura 12 - Representação de uma rede neural 4-3-2.....	20
Figura 13 - Representação gráfica da busca do mínimo global.....	22
Figura 14 - Representação do raio de influência r de uma função gaussiana RBF.....	24
Figura 15 - Representação dos pontos de uma função de ordem complexa.....	25
Figura 16 - Representação dos pontos de uma função de ordem complexa, com perfeito ajuste.....	26
Figura 17 - Representação do estado de "underfitting" de uma rede neural.....	26
Figura 18 - Representação do estado de "overfitting" de uma rede neural.....	27
Figura 19 - Função de pertinência para a variável alta(x).....	32
Figura 20- Diagrama mostrando o tratamento fuzzy para o "problema da gorjeta".....	37
Figura 21 - Parte de espectros de EDXRF utilizados neste trabalho, mostrando as linhas espectrais dos elementos Pb e S.....	57

Figura 22 - Erros RMS calculados para a rede neural BP, para integrais e "scores".....	59
Figura 23 - Erros RMS calculados para a rede neural LM, para "scores" e integrais.....	60
Figura 24 - Erros RMS calculados para o parâmetro "spread", para a rede neural RBF, para "scores" e integrais.....	60
Figura 25 - Erros RMS calculados contra o parâmetro "raio de influência", para o modelo fuzzy.....	60
Figura 26 - Erros RMS calculados para a rede neural BP e para a rede neural LM.....	62
Figura 27 - Erros RMS calculados para o modelo fuzzy e para a rede neural RBF.....	63
Figura 28 - Parte de um espectro de EDXRF típico de amostra de solo.....	64
Figura 29 - Erros RMS calculados com partes do espectro para rede BP.....	68
Figura 30 - Erros RMS calculados para as redes BP e LM, utilizando-se "scores" e integrais.....	69
Figura 31 - Erros RMS calculados para as redes BP e LM.....	69
Figura 32 - Variação do erro RMS com o número de iterações.....	70
Figura 33 - Erros RMS calculados para o modelo fuzzy, para "scores" e integrais.....	70
Figura 34 - Parte de espectros de EDXRF para amostras de plantas.....	71
Figura 35 - Espectro de EDXRF típico para uma amostra de planta.....	72
Figura 36 - Erros RMS calculados para a rede neural BP, para integrais e "scores".....	74
Figura 37 - Erros RMS calculados para a rede neural LM, para "scores" e integrais.....	74
Figura 38 - Erros RMS calculados para a rede neural RBF, para "scores" e integrais.....	74
Figura 39 - Erros RMS calculados para o modelo fuzzy, para "scores" e integrais.....	75
Figura 40 - Parte dos 16 espectros de EDXRF de amostras de aço, sem reescalonamento.....	76

Figura 43 - Espectros de EDXRF para as amostras de aço extrapolando-se a escala de intensidade de emissão.....	77
Figura 42 - Erros RMS calculados para a rede neural BP, para os dois grupos de validação.....	79
Figura 43 - Erros RMS calculados para a rede neural LM, para os dois grupos de validação.....	80
Figura 44 - Erros RMS calculados para a rede neural RBF, para os dois grupos de validação.....	80
Figura 45 - Erros RMS calculados para o modelo fuzzy, para os dois grupos de validação.....	80
Figura 46 - Espectros reescalados das amostras de aço.....	81
Figura 47 - Parte de espectros de fluorescência de raios-X por radiação síncrotron.....	82
Figura 48 - Erros RMS calculados para a rede neural BP, para os dois grupos de validação usando "scores"	85
Figura 49 - Erros RMS calculados para a rede neural BP, para os dois grupos de validação usando integrais	86
Figura 50 - Erros RMS calculados para a rede neural LM, para os dois grupos de validação usando integrais.....	86
Figura 51 - Erros RMS calculados para a rede neural LM, para os dois grupos de validação usando "scores".....	86
Figura 52 - Erros RMS calculados para a rede neural RBF, para os dois grupos de validação, utilizando os "scores".....	86
Figura 53 - Erros RMS calculados para a rede neural RBF, para os dois grupos de validação, utilizando integrais.....	87
Figura 54 - Erros RMS calculados para o modelo fuzzy, para os dois grupos de validação, utilizando os "scores".....	87
Figura 55 - Erros RMS calculados para o modelo fuzzy, para os dois grupos de validação, utilizando integrais.....	87
Figura 56 - Parte de espectros sobrepostos de EDXRF de amostras puras de Pr, Nd e Sm.....	88
Figura 57 - Parte de espectros sobrepostos de EDXRF de amostras de terras raras.....	89

Figura 58 - Erros RMS calculados para as redes neurais BP e LM.....	91
Figura 59 - Erros RMS calculados para a rede neural RBF e para o modelo fuzzy.....	91
Figura 60 - Região de interesse dos 32 espectros de EDXRF para amostras de plantas.....	92
Figura 61 - Erros RMS calculados para as redes neurais BP e LM.....	95
Figura 62 - Erros RMS calculados para a rede neural RBF e para o modelo fuzzy.....	95
Figura 63- Espectros reescalados das amostras de plantas.....	96
Figura 64- Erros RMS plotados para as redes BP e LM.....	100
Figura 65 - Erros RMS para a rede RBF e para o modelo fuzzy utilizando apenas os "scores".....	100
Figura 66 - Erros RMS para a rede RBF e para o modelo fuzzy.....	100
Figura 67 - Erros RMS plotados para as redes BP e LM.....	101
Figura 68 - Erros RMS para a rede RBF e para o modelo fuzzy, utilizando apenas integrais.....	101
Figura 69 - Erros RMS plotados para as redes BP e LM utilizando apenas integrais.....	101

Índice de tabelas

Tabela 1 - Dados estatísticos obtidos de um conjunto de dados genéricos.....	15
Tabela 2 - Condições de irradiação específicas para cada matriz.....	43
Tabela 3.1 - Faixas de concentrações dos elementos estudados neste trabalho.....	43
Tabela 3.2 - Concentrações das amostras de aço.....	44
Tabela 3.3 - Concentrações das amostras de sais de chumbo e enxofre.....	45
Tabela 3.4 - Concentrações das amostras de solos.....	46
Tabela 3.5 - Concentrações das amostras de plantas	47
Tabela 3.6 - Concentrações das amostras de terras raras.....	48
Tabela 4 - Erros RMS calculados para os modelos, para misturas de sais de chumbo e enxofre.....	58
Tabela 5 - Erros relativos para a rede BP com integrais.....	59
Tabela 6 - Erros RMS calculados para extrapolação dos modelos, para misturas de sais de chumbo e enxofre.....	61
Tabela 7 - Erros relativos para a extrapolação da rede LM.....	62
Tabela 8 - Erros RMS calculados para os modelos, para amostras de solos(1)	65
Tabela 9 - Erros RMS calculados para os modelos, para amostras de solos(2)	66
Tabela 10 - Erros relativos obtidos para a rede LM.....	67
Tabela 11 - Erros RMS calculados para os modelos, para amostras de plantas.....	72
Tabela 12 - Erros relativos para a rede RBF com integrais.....	73
Tabela 13 - Erros RMS calculados para a extrapolação dos modelos, para amostras de plantas.....	75
Tabela 14 - Erros RMS calculados para os modelos, para amostras de aço.....	77
Tabela 15 - Erros relativos para a rede BP, para o GRUPO 1.....	78
Tabela 16 - Erros relativos para a rede LM, para o GRUPO 2.....	78
Tabela 17 - Erros RMS para os espectros reescalados (amostras de aço)...	82
Tabela 18 - Erros relativos para a rede BP com integrais para o GRUPO 1.....	84

Tabela 19 - Erros relativos para a rede BP com integrais para o GRUPO 2.....	84
Tabela 20 - Erros RMS calculados para os modelos, para amostras de terras raras.....	89
Tabela 21 - Erros relativos para a rede neural BP.....	90
Tabela 22 - Erros RMS calculados para amostras de plantas (espectros comuns).....	93
Tabela 23 - Erros relativos para os espectros comuns, para a rede BP.....	94
Tabela 24 - Erros RMS calculados para amostras de plantas (espectros reescalados).....	97
Tabela 25 - Erros relativos para a rede BP, utilizando integrais, para o processo de reescalonamento.....	98
Tabela 26 - Erros RMS calculados para a rede neural BP (elemento único), com "scores" como dados de entrada.....	103
Tabela 27 - Erros RMS calculados para a rede neural BP (elemento único), integrais como dados de entrada, para os espectros reescalados.....	105
Tabela 28 - Erros RMS calculados para a rede neural BP (elemento único), com "scores" como dados de entrada, para os espectros reescalados.....	106
Tabela 29 - Erros relativos para a rede BP com "scores" (espectros comuns), correspondente aos melhores resultados da Tabela 26.....	113
Tabela 30 - Erros relativos para a rede BP com integrais (espectros reescalados), correspondente aos melhores resultados da Tabela 27.....	114
Tabela 31 - Erros relativos para a rede BP com "scores" (espectros reescalados), correspondente aos melhores resultados da Tabela 28.....	115

1. Introdução

A determinação qualitativa e quantitativa elementar em uma grande variedade de amostras é importante em muitos campos da ciência, tecnologia e processos industriais. Quando a amostra é um líquido homogêneo, de baixo teor de eletrólitos e de baixa viscosidade, os métodos mais comumente utilizados são Espectrometria de Absorção Atômica (AA) e Espectrometria de Emissão em Plasma Indutivamente Acoplado (ICP). Se o material é um sólido, pó, pasta ou líquido de alta viscosidade, estes métodos não são viáveis sem uma intensa preparação da amostra. Além disso, muitas aplicações industriais requerem análises rápidas qualitativas ou mesmo semi-quantitativas. Quando existe pouca informação sobre a amostra, os métodos AA e ICP são limitados se o interesse é qualitativo. Para determinações quantitativas, a escolha do tamanho da amostra e fatores de diluição requerem métodos de “tentativa e erro”. Estas circunstâncias, juntamente com muitos tipos de análises de rotina, são situações facilmente contornadas pela Espectrometria de Fluorescência de Raios-X de Energia Dispersiva (EDXRF) [1, 2].

1.1. Absorção e Emissão de Raios-X

A radiação-X útil analiticamente é gerada por uma variedade de métodos de bombardeamento da amostra :

- a) com elétrons na faixa de 5-100 keV;
- b) com fótons de raios-X numa faixa de energia similar, ou
- c) com partículas carregadas, como prótons ou partículas alfa, na faixa de 2-8 MeV

O uso de partículas carregadas de um acelerador para excitação deu origem à técnica PIXE (do inglês, Particle Induced X-Ray Emission). Esta técnica tem sido útil em alguns casos especiais, mas já que requer um

acelerador de partículas, não é muito usada [1, 2].

O bombardeamento eletrônico da amostra é usado em microscopia eletrônica de varredura (SEM) e em análise por microsonda eletrônica (EMPA). Embora a técnica de espectroscopia de energia dispersiva (EDS) associada com SEM seja um método excepcionalmente útil, ele não é analiticamente comparável com EDXRF. Tanto PIXE como EDS têm várias desvantagens em análises de rotina. Por exemplo, a amostra deve ser colocada em vácuo e deve ser condutora, ou tratada com uma cobertura que torne sua superfície condutora. Esta condutividade é necessária para evitar problemas de carregamento da superfície no momento do bombardeamento com partículas carregadas. Além disso, o bombardeamento com partículas, principalmente com elétrons, penetra a amostra somente superficialmente. Portanto, a emissão de raios-X resultante falha em representar a amostra em sua composição total. As restrições em relação à amostra, os complicados métodos de preparação e de uso do hardware fazem com que o método analítico não tenha a rapidez e simplicidade necessárias para análises de rotina [1, 2].

A emissão de raios-X gerada por irradiação de uma amostra com fótons de raios-X, como ocorre em EDXRF, é mais prática. A fonte dos fótons pode ser um tubo de raios-X, um alvo secundário irradiado com fótons oriundos do tubo, ou emissão de uma fonte radioisotópica. Os radioisótopos podem ser caros, e seu uso é limitado normalmente a sistemas que precisam ser portáteis. Como é necessário uma grande variedade de isótopos para o fornecimento da faixa de energia necessária e o fluxo de raios-X emitido por eles é muito baixo, os isótopos não são práticos na maioria dos laboratórios modernos. Tanto fontes isotópicas como alvos secundários emitem somente linhas características [1, 2].

Quando fótons de raios-X de energia suficiente incidem sobre um átomo, um elétron de uma camada mais interna pode ser fotoejetado do átomo. A energia necessária para iniciar a fotoejeção é denominada energia do vértice de absorção e sua intensidade depende do elemento e do elétron ejetado. A probabilidade de um raio-X ser absorvido por uma amostra é representada em

espectroscopia de raios-X pelo coeficiente de absorção de massa (μ_m). A Figura 1 mostra um gráfico do coeficiente de absorção de massa (com unidades de cm^2/g) em função do comprimento de onda do raio-X (em \AA) para alguns elementos químicos.

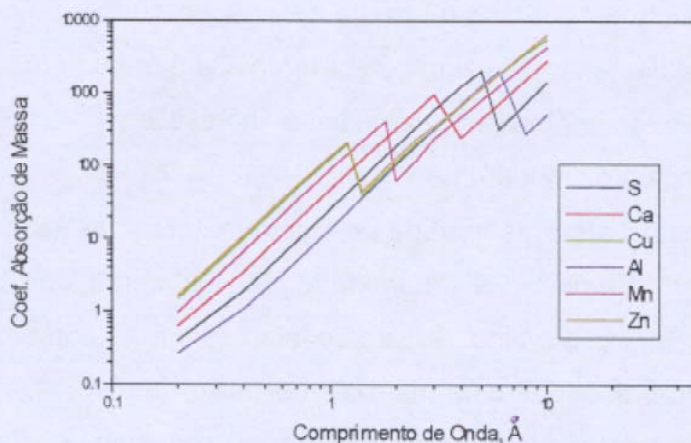


Figura 1 - Gráfico em escala logarítmica do coeficiente de absorção de massa (em unidades de cm^2/g) em função do comprimento de onda do raio-X para alguns elementos químicos [2].

É importante estabelecer de maneira clara a afirmação de que os fótons de raios-X *podem* ser emitidos quando ocorre uma transição, já que há também um outro processo competindo com a emissão de raios-X (formação de elétrons Auger). A energia liberada pode ser internamente convertida no átomo, causando a ejeção de um elétron secundário, ou Auger. Se isto acontece, nenhum raio-X é emitido. A fração de átomos excitados que emitem raios-X é chamada de rendimento de fluorescência, que é dependente do número atômico do elemento e da transição envolvida. Devido ao baixo rendimento de fluorescência dos elementos de baixo número atômico e os baixos valores de energia que eles emitem, eles são mais difíceis de serem determinados por fluorescência de raios-X [1, 2].

1.2. Espalhamento de Raios-X

Quando fótons de raios-X atingem elétrons nos átomos alvo, os fótons podem ser espalhados pelos elétrons. O espalhamento de raios-X é causado por elétrons mais externos, fracamente ligados ao átomo. Se as colisões são elásticas, o espalhamento ocorre sem perda de energia do fóton. O processo é conhecido por espalhamento Rayleigh ou coerente. Se o fóton perde energia no processo de ejeção do elétron, o processo é inelástico e é conhecido por espalhamento Compton ou incoerente [1, 2].

O espalhamento afeta as medidas espectrométricas de raios-X de duas formas importantes. Primeiro, a quantidade total de radiação espalhada aumenta com o número atômico do átomo alvo, dado o número maior de elétrons. No entanto, amostras com matrizes de baixo número atômico exibem um maior *background* devido à reduzida absorção da radiação espalhada dentro do alvo. Em segundo lugar, a razão das intensidades de espalhamento Compton/Rayleigh aumenta conforme o número atômico da matriz diminui. Esta última observação tem sido útil na estimativa da composição global de matrizes [1, 2].

1.3. Instrumentação para EDXRF

O tubo de raios-X fornece normalmente cerca de 50 W de potência e pode ser operado a um máximo de 50 kV e 0,1 mA. Já que o tubo opera a baixa potência, ele pode ser resfriado com ar. Normalmente, o material do ânodo que é escolhido para aplicações específicas, é de Rh, Ag ou Mo. Uma bandeja, com geralmente 6 filtros, fica localizada entre o tubo e a amostra. As bandejas de amostra e de filtros são as únicas partes móveis num espectrômetro EDXRF [1, 2].

Todos os espectrômetros EDXRF possuem um detector de estado sólido de Si(Li). O detector pode ser considerado como uma estrutura em camadas, na

qual uma região ativa dopada com lítio separa um lado tipo-p de outro tipo-n. Sob voltagem reversa de aproximadamente 800-1000 V, a região ativa serve como um isolante, com um gradiente de campo elétrico através de seu volume. Quando um fóton de raios-X entra na região ativa do detector, ocorre a fotoionização com um par elétron-lacuna criado para cada 3,8 eV de energia do fóton. Idealmente, o detector deveria coletar completamente a carga criada por cada entrada de fóton e resultar numa resposta para somente aquela energia. Na realidade, entretanto, algumas contagens de *background* aparecem a energias abaixo da energia do fóton, em virtude das perdas ocorridas no detector [1, 2].

A carga coletada a cada vez que um fóton de raios-X entra no detector é convertida em um valor digital que representa a energia do fóton, a qual é interpretada como um endereço na memória por um computador. O conteúdo de cada endereço é incrementado, significando portanto a detecção de um fóton correspondente àquela energia. Um circuito eletrônico, chamado frequentemente de circuito de rejeição de empilhamento de pulsos, impede a coleção de carga dos vários fótons que entram rápida e sucessivamente no detector, rejeitando a conversão em progresso. Este processo permite o acúmulo de um espectro de emissão na memória e sua subsequente imagem na tela de vídeo. Um espectro típico é mostrado na Figura 2.

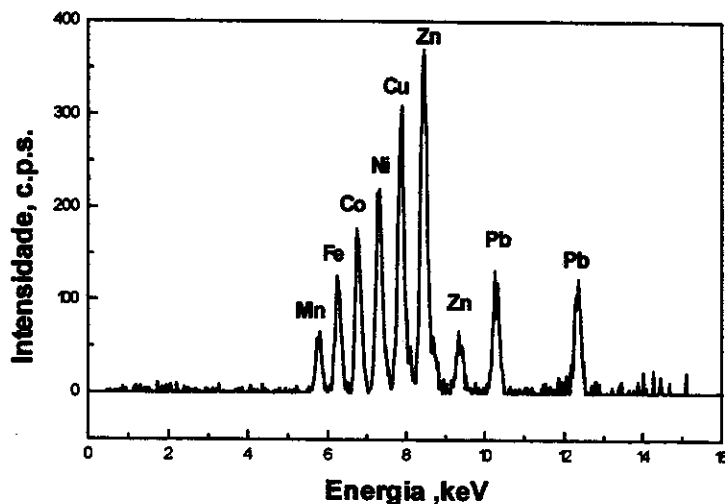


Figura 2 - Espectro de EDXRF típico para vários metais pré-concentrados em resina de troca catiônica.

1.4. Efeitos de Matriz

Muitas amostras que aparentemente seriam adequadas para análise por EDXRF, apresentam problemas que são freqüentemente atribuídos a *efeitos de matriz*. Exemplos destas “amostras problemas” são ligas, amostras geológicas e materiais biológicos. Os efeitos de matriz surgem da interação dos elementos que compõem a amostra, de modo que afetam a intensidade de emissão de uma forma não linear [1, 2]. O grau de absorção de radiação de um elemento depende de seu coeficiente de absorção de massa e da energia da radiação. O coeficiente de absorção de massa de uma amostra complexa é a média ponderada dos coeficientes dos elementos individuais constituintes da amostra. Se a composição da matriz da amostra varia, a intensidade de emissão dos elementos presentes dependerá da variação dos coeficientes de absorção, assim como da concentração do analito. Estes efeitos são especialmente importantes quando o vértice de absorção de um elemento está logo abaixo da linha de emissão de um outro. Como já foi mencionado, a absorção líquida será resultante tanto do espalhamento da radiação por elétrons mais externos como da ejeção fotoelétrica de elétrons internos. O último processo resulta em excitação adicional do elemento absorvedor, o que causa um acréscimo da intensidade de emissão daquele elemento [1, 2]. Estes efeitos combinados são conhecidos como *efeitos de matriz* e como *efeitos de acréscimo de absorção*. Estes efeitos estão ilustrados na Figura 3.

Existem métodos para correção destes efeitos, sendo uma classe deles os métodos empíricos conhecidos como “métodos do coeficiente alfa”, baseado no ajuste das concentrações e das intensidades usando padrões e regressões múltiplas não lineares. Neste caso, pode haver necessidade de um número muito grande de padrões, o que pode ser muito inconveniente [1, 2].

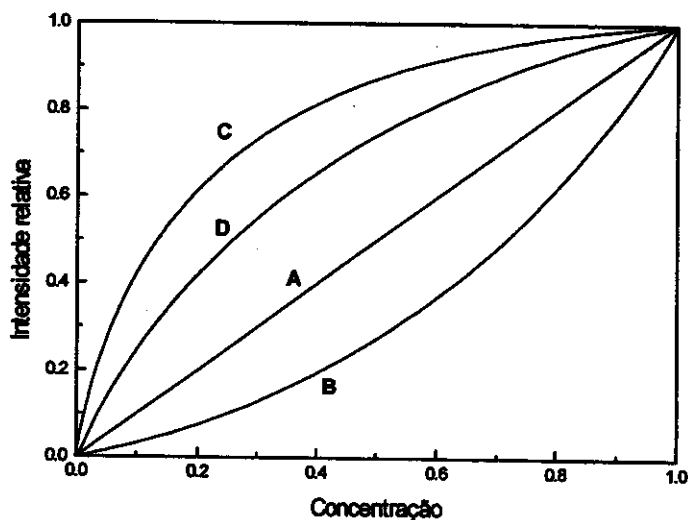


Figura 3 - Tipos de efeitos interelementares em espectroscopia de fluorescência de raios-X: (A) curva de calibração linear; (B) absorção preferencial pela matriz; (C) absorção preferencial pelo analito; (D) fluorescência secundária.

Como exemplo, vamos mostrar um dos algoritmos mais utilizados, conhecido como modelo de Lucas-Price (eq.1). A concentração de um elemento (C_i) é corrigida por coeficientes, obtidos com padrões:

$$C_i = B_i + K_i R_i + \sum_{j=1}^n \alpha_{ij} R_i R_j \quad (\text{eq. 1})$$

onde C_i é a concentração do elemento i ; R é a intensidade relativa do elemento; α é uma constante que representa o efeito do elemento j na intensidade do elemento i com n possíveis elementos interferentes; B_i e K_i são coeficientes de calibração do elemento i em função de sua intensidade.

Neste caso quanto mais elementos houver na amostra, mais termos será necessário adicionar à equação 1.

1.5. Utilização de EDXRF

EDXRF pode ser usada para uma variedade muito grande de amostras,

inclusive líquidos ou soluções homogêneas. Antes da análise, suspensões e dispersões devem ser estabilizadas para evitar sedimentação durante a irradiação. Íons em solução podem ser precipitados e filtrados ou coletados em resinas de troca iônica. Celas de fluxo para líquidos são disponíveis e usadas principalmente para controle de processos industriais [2].

Já que EDXRF pode ser usada para analisar sólidos, a grande maioria de suas aplicações envolvem ligas, minerais e outros materiais sólidos. Além dos requisitos básicos de que a amostra deve estar homogênea e ter uma superfície plana, um fator essencial no uso da técnica é o de bem preparar a amostra, segundo métodos bem estabelecidos e recomendados em manuais [2].

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. Não considerando 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 (Figura 4). Neste exemplo, 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 então que o espectro desta amostra terá a linha espectral de B bem maior do que a de A.

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.

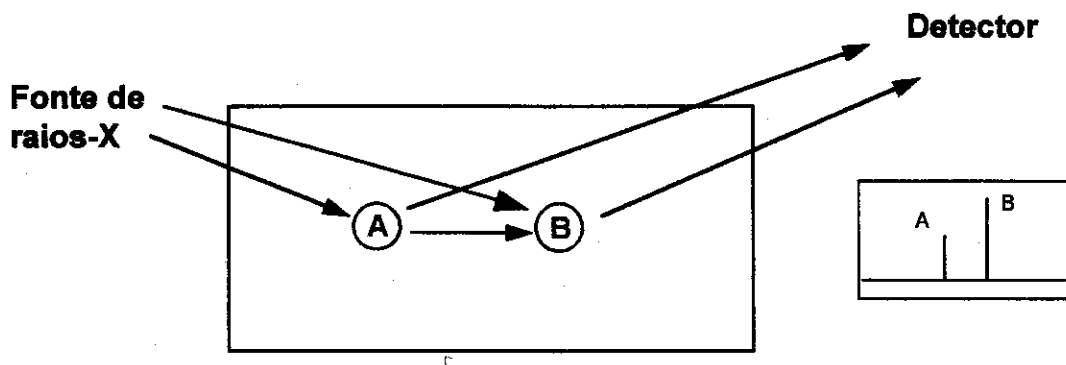


Figura 4 - Representação simplificada da fluorescência secundária.

1.6. Modelos Quimiométricos

Devido ao desenvolvimento dos computadores, a diminuição do seu custo e a automação dos métodos de análise, a Química Analítica tem crescido significativamente nos últimos anos. Aliado a isto, pode-se também destacar o aumento na demanda por análises que, associadas principalmente à automação de instrumentos, têm contribuído para o aumento do número de dados.

Devido a isso, a utilização de métodos matemáticos e estatísticos, modernos e eficientes, se faz necessária, dentro da quimiometria, para se obter o máximo de informação do conjunto de dados, resultantes de uma análise [3].

Neste trabalho, serão considerados os modelos PLS, PCR, Redes Neurais (BP, LM e RBF) e Modelagem Fuzzy, todos aplicados ao tratamento de dados de Fluorescência de Raios-X de Energia Dispersiva.

1.6.1. Modelos PLS e PCR

Os modelos PCR ("Principal Components Regression") e PLS ("Partial Least Squares") e a PCA (Análise de Componentes Principais) são modelos baseados no redimensionamento das variáveis através da geometria vetorial e utilizam conceitos como autovalores, autovetores e espaço de variáveis.

Antes de explicar estes modelos é importante entender alguns destes conceitos. Consideremos inicialmente duas matrizes de dados, como mostra a Figura 5: (1) $R_{n \times j}$ que é a matriz de respostas instrumentais, também conhecida como matriz de *variáveis independentes*, e pode conter valores de intensidades de espectros de IR, UV, fluorescência de raios-X, etc, ou pode conter ainda as respostas de um cromatograma ou de um outro instrumento analítico; n é número de objetos (espectros ou cromatogramas) e j é o número de variáveis (comprimentos de onda, tempos de retenção, etc). (2) $C_{n \times k}$ é a matriz que contém as concentrações, também conhecida como matriz de *variáveis dependentes*, onde n é o mesmo número de objetos e k representa o conteúdo dos objetos que se quer determinar, por exemplo, $k = 2$ pode significar a concentração de dois ácidos diferentes em n espectros de IV em j comprimentos de onda diferentes.

Os modelos PCR, PLS, e PCA fazem a decomposição da matriz R em produto de dois vetores chamados de *scores* e *loadings*, sendo que estes vetores (expressos como "scores") plotados em um gráfico cartesiano permitem visualizar a distribuição das variáveis, como mostrado na Figura 6.

Na parte A da Figura 6 é representado o gráfico de pontos de uma matriz $R_{15 \times 2}$ em um espaço de $j = 2$ variáveis. Cada ponto pode representar uma amostra que possui intensidades de absorção nos comprimentos de onda indicados pelos eixos j_1 e j_2 . CP1 é a reta que melhor se ajusta aos pontos, ou seja, esta reta obtida por mínimos quadrados é aquela, entre todas as possíveis, que contém o máximo de variância dos dados, ou seja, ela exhibe o máximo de informação estatística. Esta reta corresponde ao primeiro *componente principal*.

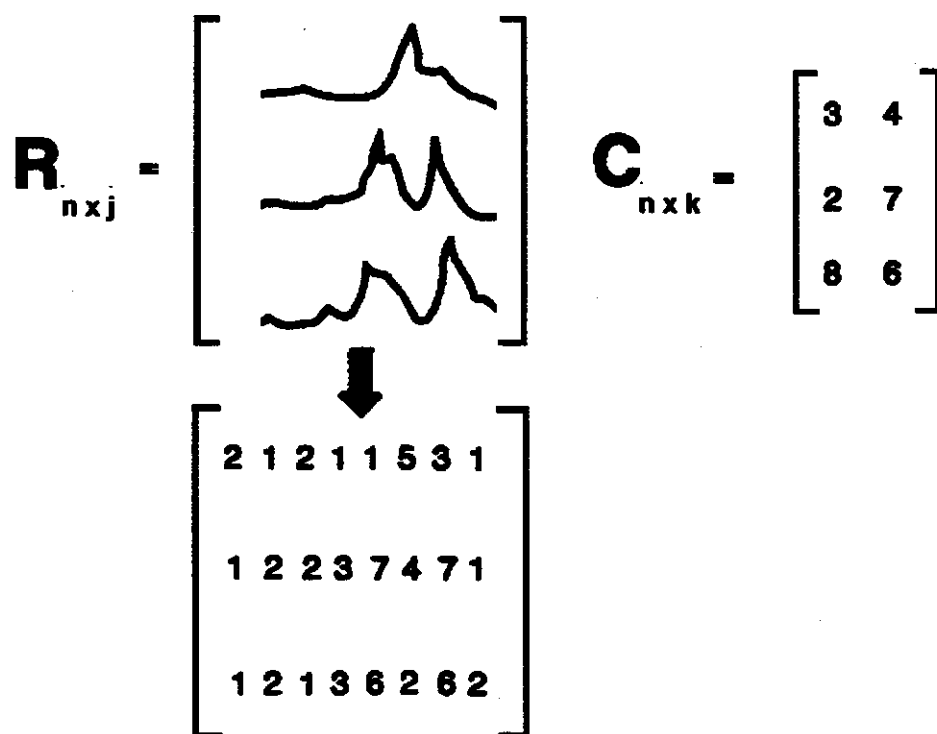


Figura 5 - Representação das matrizes R e C.

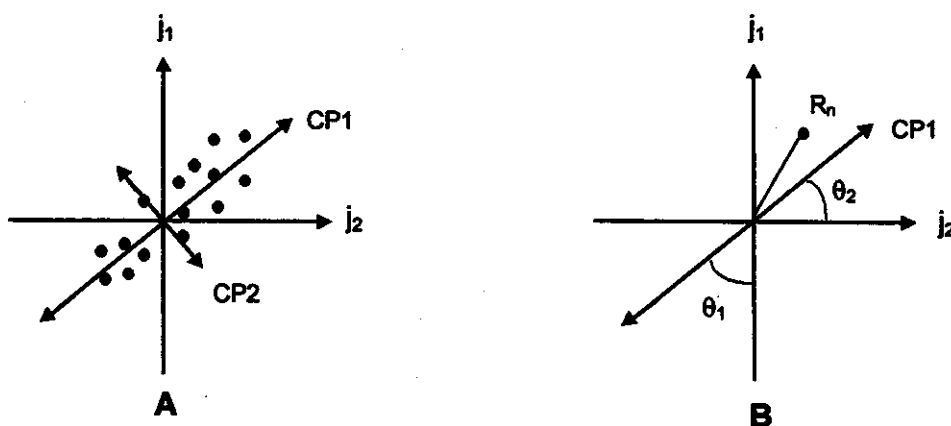


Figura 6 - Representações gráficas das matrizes de scores e de loadings.

A reta CP2, correspondente ao segundo componente principal e explica a maior quantidade de informação não descrita pelo primeiro componente. Esta reta CP2 deve ser ortogonal a CP1, ou seja, possuir um ângulo de 90 graus entre as duas retas. Na parte B da Figura 6 podemos observar um objeto

genérico R_n , em que o *score* deste ponto corresponde a sua projeção sobre a reta do componente principal CP1. Os *loadings* são os cossenos dos ângulos θ_1 e θ_2 entre a reta do componente principal e cada variável j .

Resumindo, a Análise de Componentes Principais (PCA) [4, 5] tem como objetivo encontrar combinações lineares das variáveis independentes, de modo a reduzir suas dimensões a um conjunto menor de parâmetros que ainda descreva suas características; por isso ela também é conhecida como uma técnica de *compressão de dados*. Na análise PCA, a matriz de variáveis independentes é decomposta em uma soma de um produto de vetores, que não podem mais serem reduzidos, além de uma matriz de erros, como mostra a equação 2.

$$R = t_1 p_1^T + t_2 p_2^T + \dots + t_k p_k^T + E \quad (\text{eq. 2})$$

onde E é uma matriz de erros, t_k são os *scores* e p_k são os *loadings* da correspondente matriz de variáveis instrumentais R . Esta equação 2 é a representação matemática dos parâmetros *scores* e *loadings* e na Figura 6 é possível observar sua representação gráfica.

As funções do programa “PLS Toolbox” para Matlab™ utilizadas neste trabalho geram o modelo PCA para um conjunto de dados e aplicam os testes estatísticos Q (medida do resíduo entre cada amostra e sua projeção para cada componente principal obtida pelo modelo) e T^2 (medida da variação de cada amostra em relação ao modelo PCA) para detecção dos “outliers” (ou amostras anômalas), considerando que todos os dados da matriz R estejam centrados na média [5].

O modelo PCR (e também o PCA) parte da representação da matriz de dados R através de duas matrizes, a matriz de “scores” T e a matriz de “loadings” P :

$$R P = T \quad (\text{eq. 3})$$

A matriz T possui um número menor de variáveis e deverá conter toda a informação relevante sobre R . Em seguida, este modelo estabelece uma relação

entre o bloco de variáveis dependentes (matriz C de concentrações) e a matriz de “scores” T :

$$C = T S \quad (\text{eq. 4})$$

S é a matriz dos coeficientes de regressão e é obtida do seguinte modo:

$$S = (T^t T)^{-1} T^t C \quad (\text{eq. 5})$$

A determinação de P , T e S é feita através de um conjunto de calibração (R e C conhecidos), permitindo que se possa encontrar uma matriz C de concentrações desconhecidas através da matriz instrumental R . A etapa mais importante no modelo PCR é a escolha do número de componentes principais que serão utilizados, pois este número representa a quantidade de informação sobre o sistema.

O modelo PLS estima matrizes de “scores” e “loadings” tanto para R como para C :

$$R = T P + E \quad (\text{eq. 6})$$

$$C = U Q + F \quad (\text{eq. 7})$$

Os elementos de T e U das equações 6 e 7 são chamados *scores*; os elementos de P e Q são chamados de *loadings*; E e F são as matrizes de resíduos não explicados pelo modelo, de R e C , respectivamente.

As matrizes R e C são usadas indiretamente para estimar os componentes principais. Pode-se então utilizar a relação:

$$U = b T + G \quad (\text{eq. 8})$$

b é chamado de termo de relação interna e G é o erro experimental. Através de um conjunto de calibração, determinam-se todas as matrizes de *scores* e *loadings* e aplica-se o modelo através da relação:

$$C = b T Q + G \quad (\text{eq. 9})$$

C é a matriz de concentração desconhecida e G é o erro experimental. O modelo PLS estima os fatores para R e C usando toda a informação disponível de ambos.

1.6.1.1. O método “Cross Validation”

A escolha do número de componentes principais que determinará o número de “scores” é feita neste trabalho por análise PCA através do algoritmo “Cross Validation” (CV) ou Validação Cruzada, que consiste na determinação do menor erro de previsão conhecido como PRESS (eq. 10), em função do número de componentes principais [6],

$$\text{PRESS} = \sum_{i=1}^n (C_i - \hat{C}_i)^2 \quad (\text{eq. 10})$$

onde,

C_i é a concentração real das amostras de calibração;

\hat{C}_i é a concentração calculada para as amostras de calibração;

n é o número de amostras de calibração.

O procedimento consiste em calcular o erro para um determinado número de componentes principais:

- Remove-se uma ou mais amostras do conjunto de calibração e constrói-se o modelo;
- Utiliza-se então este modelo para prever os dados removidos;
- Calcula-se o erro de previsão ($C_i - \hat{C}_i$);
- Calcula-se a soma dos quadrados dos erros de previsão (PRESS);
- Repete-se todo o procedimento para outras amostras.

O processo é repetido para modelos com uma, duas, três, e assim por

diante, componentes principais. Para cada sistema em estudo, o número adequado de componentes será correspondente ao menor valor de PRESS.

Além do método CV, existem outros métodos para a determinação do número de componentes principais em uma análise PCA [7]. São eles: a porcentagem de variância explicada, o critério de autovalor e o "Scree-test".

Para entender a diferença entre eles, vamos utilizar a seguinte tabela para um conjunto de dados genéricos:

Tabela 1 - Dados estatísticos obtidos de um conjunto de dados genéricos.

Componente	Autovalor (u.a.)	Variância explicada (%)	Variância acumul. (%)
1	3,352	67,05	67,05
2	1,182	23,65	90,70
3	0,285	5,70	96,40
4	0,135	2,70	99,10
5	0,045	0,90	100,00

u.a. = unidades arbitrárias, acumul. = acumulada

O critério da porcentagem da variância explicada é obtido através do cálculo do desvio padrão dos valores individuais dos dados, dentro do conjunto. Deste modo, se todos os possíveis componentes principais forem usados no modelo, isto corresponderia a uma variância de 100%. Normalmente uma variância explicada em torno de 90% é suficiente para determinar o número de componentes. De acordo com a Tabela 1, 90,7% da variância pode ser explicada pelo uso de duas componentes principais.

O critério de autovalor é baseado no fato que o valor médio dos autovalores para dados autoescalados geralmente é 1. Somente componentes com autovalores maiores que 1 são considerados importantes. Por este critério, para os dados da Tabela 1, apenas duas componentes principais são importantes.

Finalmente, o método do "Scree-test" é baseado nos níveis de variância residual. O número apropriado de componentes é obtido analisando-se o

primeiro menor resíduo entre as variâncias associadas as componentes, algo em torno de 5%. Para a Tabela 1, este número estaria entre a 2ª e a 3ª componente principal.

A determinação do número de componentes principais após a decomposição da matriz de dados é fundamental para separar quais fatores estão associadas aos componentes puros e quais estão relacionadas ao ruído e/ou componentes menos importantes.

1.6.2. 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 [8]. 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 [8, 9]. 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 representado na Figura 7.

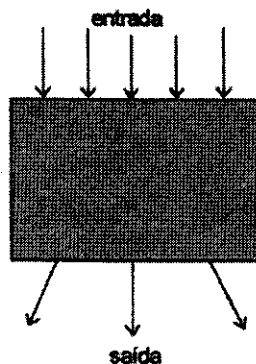


Figura 7 - Representação da rede neural como uma “caixa de processamento”, mostrando a entrada e a saída de dados.

Como as redes neurais podem ser utilizadas para diversas aplicações, os dados de entrada da figura acima 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 [8].

Dentro desta “caixa de processamento”, a rede neural transforma uma entrada de m variáveis em uma saída de n variáveis, através das unidades básicas de processamento ou neurônios como mostrado na Figura 8.

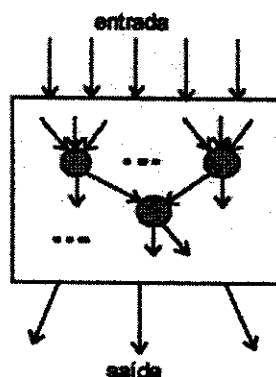


Figura 8 - Representação dos neurônios, os quais são as unidades básicas de processamento em uma rede neural artificial.

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, como é mostrado na Figura 9.

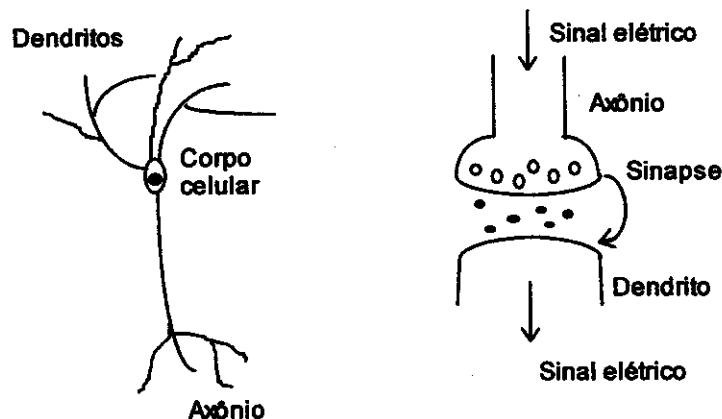


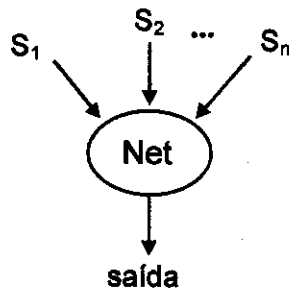
Figura 9 - Representações simplificadas de um neurônio biológico e de uma sinapse.

As 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 . A Figura 10 mostra uma representação esquemática de uma sinapse em um neurônio artificial, onde p é o sinal de saída.



Figura 10 - Representação da sinapse de um neurônio matemático.

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_i , que entra em um neurônio, será dado pela soma de todos os produtos de s_i com seus respectivos pesos w_i :



$$\text{Net} = w_1s_1 + w_2s_2 + \dots + w_is_i + \dots + w_ns_n \quad (\text{eq. 11})$$

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

$$\text{saída} = f(\text{Net})$$

Estas funções produzem uma transformação não linear no sinal total de entrada, ou seja, o sinal de entrada é transformado de acordo com a característica da função. Assim como nos neurônios biológicos, o sinal de entrada é diferente do sinal de saída [8]. Cada neurônio (ou um conjunto deles) pode ter um tipo diferente de função de transferência. Existem muitos tipos destas funções e as três mais utilizadas, mostradas na Figura 11, são: a “hard limit”, a “threshold logic” e a sigmoidal. O sinal total de entrada é modificado gerando uma saída com intervalo entre 0 e 1.

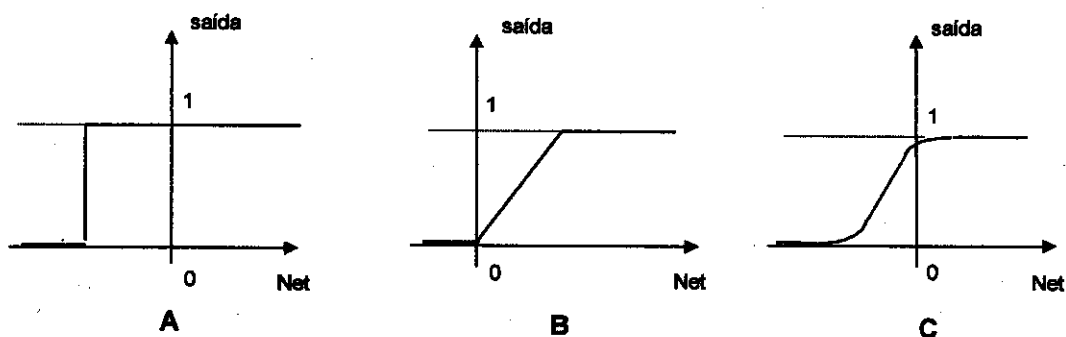


Figura 11 - Funções de transferência usadas em redes neurais: (A) “hard limit”, (B) “threshold logic” e (C) sigmoidal.

A rede neural artificial pode ser formada por várias camadas de neurônios, sendo que a distribuição deles deve ser otimizada em função da minimização dos erros. Esta distribuição varia também de acordo com o tipo de rede neural. As camadas de neurônios intermediárias entre a primeira e a última, por convenção, são ditas *camadas escondidas*. Um exemplo de arquitetura de rede neural está representada na Figura 12.

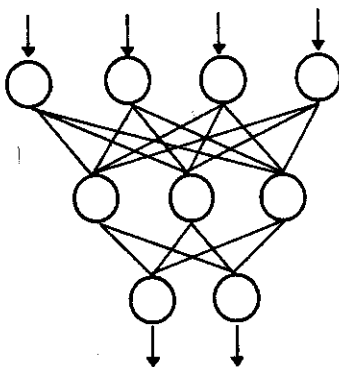


Figura 12 - Representação de uma rede neural 4-3-2.

A rede neural mostrada na Figura 12 possui uma camada de entrada com 4 neurônios, uma camada escondida com 3 neurônios e uma camada de saída com 2 neurônios.

1.6.2.1. Treinamento da rede neural

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 razoável. Este erro também é conhecido como SSE ("sum squared error") ou *soma dos erros quadráticos*:

$$SSE = \sum_{i=1}^n \left(y_i - \hat{y}_i \right)^2 \quad (\text{eq. 12})$$

onde y é o valor real e \hat{y} é o valor calculado pela rede.

O erro máximo estipulado, o número de interações 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 se encerra quando o erro estipulado atingir o critério de convergência previamente estabelecido.

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, “Radial Basis Functions” (RBF), que é um outro tipo de rede neural.

1.6.2.2. 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 [8]. Uma das principais aplicações do BP é o modelamento de dados não lineares, onde o comportamento matemático entre as variáveis não são representadas por funções de ordem definida. A equação de correção de pesos pode ser escrita resumidamente como:

$$\Delta w^k = \underbrace{\eta \cdot D}_{1^{\text{º termo}}} + \underbrace{\mu \cdot \Delta w^{k-1}}_{2^{\text{º termo}}} \quad (\text{eq. 13})$$

onde:

Δw^k representa a diferença entre os pesos inicial e final para a iteração k ;

η 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 foram corrigidos 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;

μ é chamado de "momentum", uma constante que é multiplicada pela variação de pesos da iteração anterior (Δw^{k-1}), criando uma "relação matemática" com a variação de pesos da iteração corrente (2º termo da equação 13), promovendo assim um acréscimo no valor numérico de Δw obtido no 1º termo da equação 13. A adição deste 2º termo à equação 12 evita que ocorram convergências para mínimos locais, pois o mínimo obtido pelo primeiro termo da equação 13 era realmente global. Caso contrário, o mínimo era local e o processo de busca pelo menor erro prossegue (Figura 13). Resumindo, esta constante evita a possibilidade de se achar um mínimo local;

Δw^{k-1} representa a diferença entre os pesos inicial e final para a iteração $k-1$.

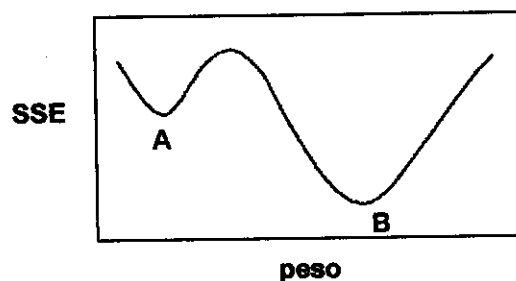


Figura 13 - Representação gráfica da busca do mínimo global. A variação do erro SSE em função dos valores dos pesos gera um gráfico com um comportamento de "altos" e "baixos", como se fossem montanhas e vales; o mínimo é então representado por um "vale". A posição A na figura acima representa um *mínimo local*, e a posição B representa um *mínimo global*.

1.6.2.3. Rede neural LM

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

$$\Delta w = (J^T \cdot J + \alpha \cdot I)^{-1} \cdot J^T \cdot \varepsilon \quad (\text{eq. 14})$$

onde

Δw representa a diferença entre os pesos inicial e final;

α 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 α , pode-se gerar uma nova matriz contendo como elementos apenas valores de α ;

ε é um vetor de erros calculados.

1.6.2.4. Rede neural RBF

A rede RBF tem esse nome pois 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 [10]. O algoritmo desta rede utiliza neurônios com a função RBF:

$$f(n, b) = e^{-(b-n)^2} \quad (\text{eq. 15})$$

sendo

$$n = \left[\sqrt{(p-w)^2} \right] \cdot b \quad (\text{eq. 16})$$

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 14). O ajuste de r é feito através de um parâmetro escolhido pelo usuário (chamado pelo algoritmo de “spread”) em função do conjunto de calibração escolhido.

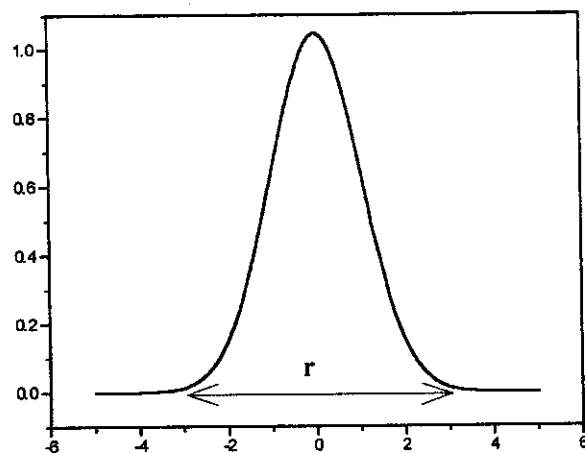


Figura 14 - Representação do raio de influência r de uma função gaussiana RBF.

Este termo “spread” 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 [11].

A otimização da arquitetura (busca do menor erro em função da adição de neurônios à camada intermediária) neste algoritmo é feito de uma maneira diferente, pois o algoritmo adiciona estes neurônios até encontrar o menor erro pré-estabelecido. Isto pode ser considerado como uma vantagem para a utilização deste tipo de rede, uma vez que para os outros dois algoritmos é necessário este procedimento, além do mesmo ser feito manualmente.

1.6.2.5. “Overfitting” e “Underfitting”

O processo de otimização da arquitetura de uma rede neural, consiste (como mencionado anteriormente) na busca do menor erro de treinamento ou SSE, (vide eq. 12) em função da adição de neurônios à camada intermediária. A escolha inadequada deste número de neurônios pode levar a rede neural aos estados conhecidos como “overfitting” (ou super ajuste) e “underfitting” (ou sub-ajuste), gerando altos erros e comprometendo assim todo o processo de treinamento.

Para tornar o entendimento deste problema mais claro, vamos mostrar um exemplo de utilização de uma rede neural genérica para a aproximação de uma função de ordem complexa [11]. Na Figura 15 é mostrada a representação, na forma de pontos, de uma função a qual se deseja ser aproximada. Este processo determina que seja feito o ajuste de uma linha, unindo os pontos de modo que todos eles (ou a maioria) estejam contidos nesta linha.

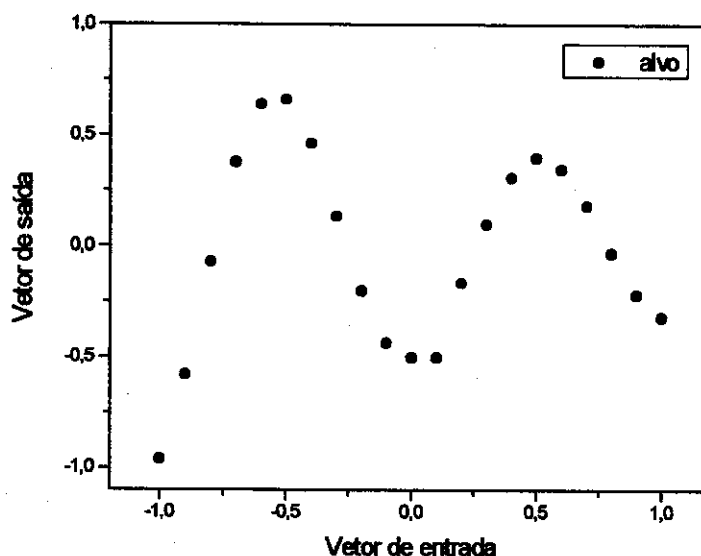


Figura 15 - Representação dos pontos de uma função de ordem complexa.

Na Figura 16 abaixo é mostrado a representação do melhor ajuste

possível, o qual representa a função de ordem complexa.

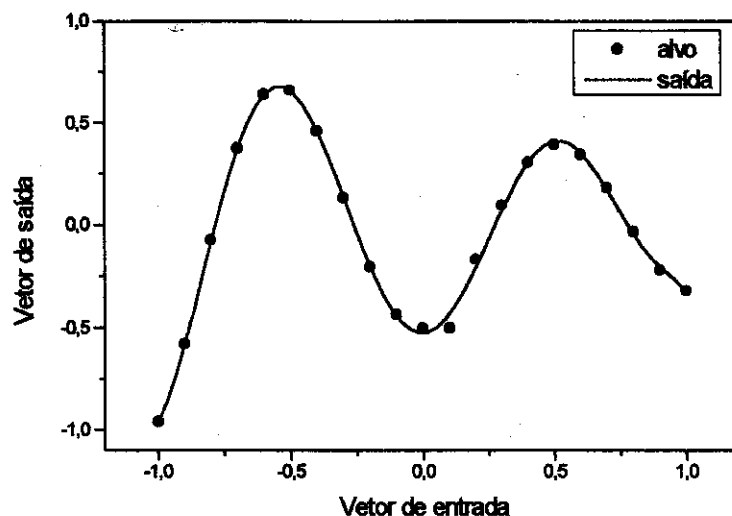


Figura 16 - Representação dos pontos de uma função de ordem complexa, e do perfeito ajuste representado pela linha azul.

A Figura 17 mostra a representação do estado de “underfitting”, ou seja, não foi possível realizar o ajuste da função, significando que é necessário aumentar o número de neurônios à camada intermediária desta rede neural.

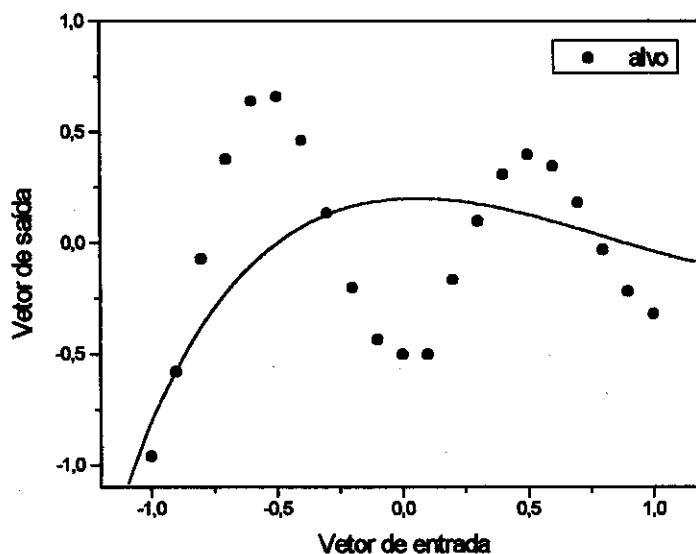


Figura 17 - Representação do estado de “underfitting” de uma rede neural, ou seja, não foi possível realizar o ajuste da função. A linha azul representa a saída da rede.

A Figura 18 mostra a representação do estado de “overfitting”, ou seja, o

ajuste da função foi feito de maneira exagerado, devido ao número excessivo de neurônios na camada intermediária.

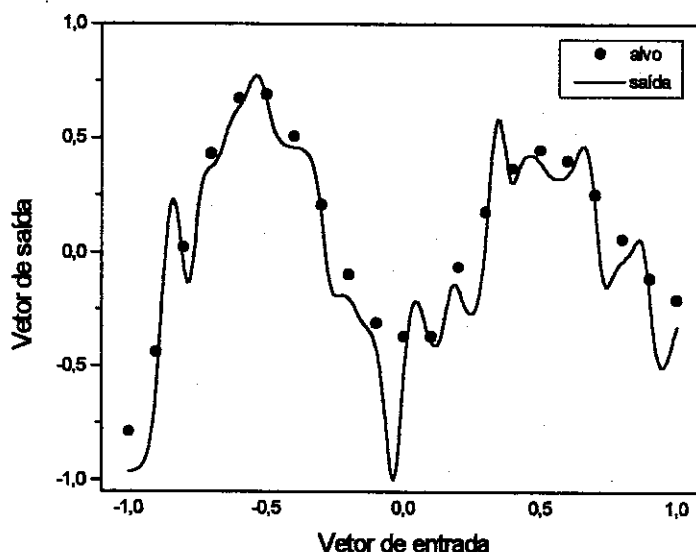


Figura 18 - Representação do estado de "overfitting" de uma rede neural, ou seja, o ajuste da aproximação da função foi exagerado. A linha azul representa a saída da rede.

1.6.2.6. Uso de redes neurais em fluorescência de raios-X

O emprego de redes neurais artificiais como ferramenta para tratamento de dados é ainda recente e muitos estudos estão sendo feitos para seu aprimoramento. Sua aplicação no tratamento de dados de fluorescência de raios-X (FRX) é ainda pouco explorada e é pequeno o número de trabalhos descritos na literatura que tratam deste tema [12- 17].

Bos e colaboradores [12] estudaram a fluorescência secundária em diversas ligas de Fe-Ni-Cr. Eles aplicaram aos espectros de FRX, redes neurais do tipo BP e compararam com outros três modelos matemáticos, a equação de Rasberry-Heinrich (RH), um modelo semi-linear e um linear. Neste tipo de matriz ocorrem quatro situações em espectros de FRX: (1) a linha $K\alpha$ do Fe é absorvida pelo cromo, (2) fluorescência secundária é induzida ao ferro pelo níquel, (3) a linha $K\alpha$ do Ni é absorvida pelo cromo e pelo ferro e, (4) ferro e níquel induzem ambos fluorescência secundária ao cromo. Foram utilizadas 130 amostras diferentes de aço. Os erros relativos encontrados para os modelos na

previsão das concentrações para cada metal foram de 0,68%, 29,0% e 20,6% respectivamente, utilizando redes neurais; 2,99%, 30,9% e 7,9%, respectivamente, utilizando a equação RH; 0,82%, 49,4% e 11,2%, respectivamente para o modelo semi-linear e 4,5%, 128% e 136%, para o modelo linear. Eles concluíram que as redes neurais foram mais eficientes.

Bos e Weber [13] estudaram duas aplicações experimentais através de FRX. Primeiro, eles utilizaram valores teóricos provenientes de ligas Fe-Ni-Cr. Eles utilizaram, no tratamento de dados, redes neurais BP, a equação RH, algoritmos genéticos e o algoritmo Ho-Kashyap (HK). Os dados de entrada para os modelos foram as intensidades relativas individuais para cada metal, e os dados de saída, as concentrações. Eles utilizaram um conjunto de calibração com 9 amostras e um de validação com 5 amostras. Os erros padrão médios de previsão obtidos foram de 19,52% para o algoritmo genético, 3,66% para a equação RH, 40,64% para o algoritmo HK e 12,71% para a rede neural BP. Segundo os autores, em função do pequeno conjunto de calibração usado, o erro para as redes neurais não foi melhor.

Na segunda aplicação experimental, os autores utilizaram 36 filmes finos contendo Fe e Ni em diferentes espessuras. Os dados foram tratados apenas com redes neurais. A entrada de dados foram as linhas $K\alpha$ do Ni e do Fe, e a saída as concentrações de ambos, além da espessura do filme. Foram utilizados dois conjuntos de validação diferentes, um com 7 amostras e outro com 5 amostras, o restante foi utilizado para calibração. Os autores testaram diversas arquiteturas de rede e obtiveram erros de previsão bem baixos.

Luo e colaboradores apresentaram, em 1997 [14], um estudo sobre a influência dos parâmetros matemáticos da rede neural em relação a velocidade de treinamento e a busca do menor erro. Foram utilizados espectros de FRX de amostras geológicas certificadas, com interesse nos metais Mg, Al, Si, K e Fe. A entrada de dados foram as intensidades das linhas $K\alpha$ destes metais e a saída, as suas concentrações. Utilizando uma rede BP, eles estudaram a variação do erro de previsão contra o número de iterações (de 10000 a 50000) para os

parâmetros "learning rate", "momentum" e uma constante de multiplicidade. Este último significa um fator que deve ser multiplicado para o escalonamento dos dados de entrada (geralmente é feito entre 0 e 1). Os melhores valores encontrados foram de 0,75, tanto para o "learning rate" como para o "momentum", e 0,636 como a constante de multiplicidade. Uma vez escolhidos estes valores, os autores aplicaram a rede neural BP e compararam com PLS e com o Método dos Parâmetros Fundamentais (MPF), que consiste em uma equação teórica baseada em parâmetros físicos e necessita uma série de padrões para cálculo das intensidades de cada elemento presente na amostra. Os erros encontrados para as redes neurais são bem menores do que os encontrados para os outros dois modelos (PLS e MPF).

Luo e colaboradores apresentaram, em 1998, um outro trabalho [15] comparando redes neurais BP em duas variantes chamadas simples componente (SC) e multi componente (MC). Utilizando espectros de FRX de amostras de ligas de cobre, os autores determinaram as concentrações de Cu, Zn, Pb e Fe de maneira univariada (para SC) e multivariada (para MC), a partir das linhas $K\alpha$ dos metais, exceto para o Pb, o qual utilizou a linha $L\alpha$. Os autores variaram os valores dos parâmetros "learning rate" e "momentum", e também o número de padrões no treinamento das redes. As redes SC obtiveram menores erros de previsão, de maneira geral para os metais, em relação as redes MC.

Boger e Karpas [16] obtiveram espectros de fluorescência de raios-X de energia dispersiva de misturas de sais de Re, Os Ir e Pt, em matriz de polietileno prensado. Estes elementos são vizinhos na Tabela Periódica e portanto, têm energias de emissão muito próximas. Há portanto sobreposição de linhas nos espectros de suas misturas. Foi utilizada uma rede neural do tipo BP com 100 neurônios na camada de entrada, correspondente a 100 canais do espectro na região de sobreposição das linhas dos 4 elementos (entre 8,5 e 9,5 KeV), 4 neurônios na camada intermediária e 4 na camada de saída (as concentrações dos 4 sais), ou seja, uma arquitetura 100-4-4. A função de transferência utilizada entre as camadas foi apenas a sigmoidal. Foram

preparadas 8 misturas dos sais dos 4 elementos com concentração entre 10 e 40 ppm. A rede foi treinada utilizando-se o método "Cross Validation" (variante "leaving one out"), ou seja, repetindo 7 amostras para calibração e uma para validação e foram encontrados erros quadráticos médios de 3,2% para o Re; 2,9% para o Os; 2,7% para o Ir e 5,5% para a Pt, no cálculo das concentrações. Os autores concluem que as redes neurais foram melhores para estimar a contribuição individual de cada metal, em relação aos métodos convencionais univariados.

Facchin e colaboradores [17] utilizaram espectros de EDXRF de 38 misturas sintéticas de sais de chumbo e enxofre em matriz de sílica, onde ocorre sobreposição das linhas de ambos. Os autores compararam uma rede neural LM com outros métodos de calibração multivariada, como o PLS-rede neural, o PLS linear, o PLS polinomial, o método de correção de intensidade e a regressão linear. Os autores utilizaram 58 pontos como entrada de dados da rede, correspondente às regiões de sobreposição das linhas $K\alpha$ do S e $M\alpha$ do Pb, além da linha $L\alpha$ do Pb. As redes neurais apresentaram os menores erros de previsão em relação aos outros modelos.

1.6.3. Modelagem Fuzzy

Matematicamente, lógica fuzzy pode ser definida como um superconjunto da lógica Booleana convencional aplicado ao conceito de verdades parciais (quando as variáveis assumem apenas os valores *verdadeiro* e *falso*), podendo ser aplicada para valores compreendidos entre "completamente verdadeiro" e "completamente falso"; estes valores são conhecidos como *variáveis linguísticas*. A lógica fuzzy foi desenvolvida por Lotfi Zadeh na Universidade da Califórnia na década de 60 com a finalidade de modelar a incerteza da linguagem natural aplicada a matemática, sendo possível então a utilização de variáveis linguísticas [18].

Conjuntos fuzzy e lógica fuzzy têm auxiliado no desenvolvimento de tecnologias computacionais contemporâneas e no processamento de informação. Sua aplicação vai desde sistemas de controle e reconhecimento de padrões até inteligência artificial [18].

A aplicação de tratamento fuzzy a um sistema permite interpretar valores em um conjunto de entrada e estabelecer correlações com um conjunto de saída, através de um conjunto de regras condicionais [19, 20]. O tratamento fuzzy de um sistema deve ser feito através de algumas etapas, são elas: definição dos conjuntos fuzzy para os dados de entrada e de saída; definição dos operadores fuzzy; definição das regras condicionais fuzzy; e a defuzzificação.

1.6.3.1. Definição de conjuntos fuzzy

Para mostrar os conceitos utilizados na aplicação de um tratamento fuzzy a um sistema, como exemplo vamos estudar a altura de pessoas utilizando conjuntos fuzzy [20]. Neste caso, vamos chamar de S o conjunto universo de pessoas, também conhecido como “universo de discurso”, que poderia ser a população de uma cidade, os alunos de uma escola, etc. Vamos definir o subconjunto *alto* como pertencente a S . Este conjunto responderá a questão “qual o grau que uma pessoa x é considerada alta?”. Neste caso, *alta* é uma variável lingüística relacionada à categoria *altura*. Para cada pessoa no universo de discurso S pode-se fazer uma correspondência através do “grau de pertinência” (do inglês, *degree of membership*) para o subconjunto fuzzy *altura*. Isto é feito através da “função de pertinência” baseada na altura das pessoas, através das relações fuzzy:

$$\text{alto}(x) = 0 \quad \text{if } \text{altura}(x) < 1,60 \text{ m};$$

$$\text{alto}(x) = [\text{altura}(x)-1,6]/0,15 \quad \text{if } 1,60 \text{ m} \leq \text{altura}(x) \leq 1,75 \text{ m};$$

$$\text{alto}(x) = 1 \quad \text{if } \text{altura}(x) > 1,75 \text{ m}.$$

A representação gráfica desta função de pertinência é mostrada na Figura 19:

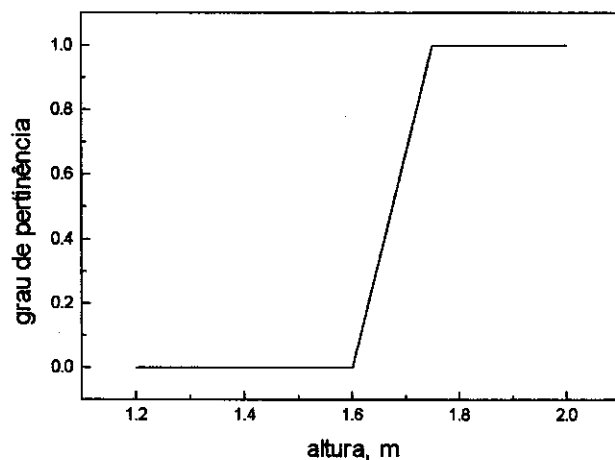


Figura 19 - Função de pertinência para a variável *alta(x)*.

Aplicando este sistema para alguns exemplos temos:

Pessoa	Altura	Grau de altura
Billy	1,55 m	0,00
Maria	1,65 m	0,33
João	1,70 m	0,67
Erik	1,73 m	0,87
Marco	1,76 m	1,00
Carlos	1,85 m	1,00

A relação entre a variável lingüística X e o elemento A pertencente ao conjunto universo S é feita através de expressões do tipo " A is X ", neste caso, " $João$ is *alto*" = 0,67. Traduzindo, o grau de pertinência da variável *altura* para o elemento *João* é igual a 0,67.

1.6.3.2. Definição de operadores fuzzy

As funções de pertinência utilizadas na maioria das aplicações normalmente não são tão simples em relação a forma, como foi mostrado no exemplo anterior. Elas podem ser triangulares, trapezoidais ou gaussianas. Da mesma forma, a caracterização das funções de pertinência nem sempre possuem relações fuzzy tão simples como aquelas mostradas anteriormente.

Para um sistema multivariado por exemplo, a função de pertinência dependerá de relações fuzzy mais complexas entre as variáveis. Neste sistema é necessário também a utilização de *operadores lógicos*, os quais são rotineiramente usados nas operações entre conjuntos e são fundamentais nas definições das relações fuzzy.

Para exemplificar, vamos utilizar novamente o sistema anterior adicionando mais uma variável lingüística, que será chamada *velho*, para o subconjunto fuzzy *idade* das pessoas [20].

As relações fuzzy para a altura das pessoas são as mesmas usadas anteriormente, e a determinação da função de pertinência para a idade será definida por:

$$\begin{aligned}\text{velho}(x) &= 0 && \text{if idade}(x) < 18 \text{ anos;} \\ \text{velho}(x) &= [\text{idade}(x)-18]/42 && \text{if } 18 \leq \text{idade}(x) \leq 60; \\ \text{velho}(x) &= 1 && \text{if idade}(x) > 60 \text{ anos.}\end{aligned}$$

A função de pertinência neste caso será idêntica a da Figura 19. Os operadores lógicos que podem ser utilizados neste caso são os seguintes:

i. Operador AND:

$a = X \text{ is } \textit{alto} \text{ and } X \text{ is } \textit{velho}$, significa que a variável a assume o menor valor entre as funções de pertinência *alto* e *velho*:

$$x \text{ AND } y = \min[\text{alto}(x), \text{velho}(x)] \quad (\text{eq. 17})$$

ii. Operador OR:

$b = X \text{ is } \textit{alto} \text{ or } X \text{ is } \textit{velho}$, significa que a variável b recebe o maior valor entre as funções de pertinência *alto* e *velho*:

$$x \text{ OR } y = \max[\text{alto}(x), \text{velho}(x)] \quad (\text{eq. 18})$$

iii. Operador NOT:

$c = \text{not}(X \text{ is } \textit{alto})$, significa que a variável c assume o valor complementar da função de pertinência *alto*:

$$\text{NOT } x = 1,0 - \text{alto}(x) \quad (\text{eq. 19})$$

Podemos então utilizar os seguintes valores:

altura	idade	X is <i>alto</i>	X is <i>velho</i>	a	b	c
1,60	65	0,00	1,00	0,00	1,00	1,00
1,63	30	0,21	0,29	0,21	0,29	0,79
1,67	27	0,38	0,21	0,21	0,38	0,62
1,66	32	0,42	0,33	0,33	0,42	0,58
1,68	31	0,54	0,31	0,31	0,54	0,46
1,80	45	1,00	0,64	0,64	1,00	0,00
1,00	4	0,00	0,00	0,00	0,00	1,00

A utilização dos resultados obtidos para as variáveis a , b , e c associados com um conjunto de regras permite que seja possível construir um modelo fuzzy.

1.6.3.3. Definição de regras fuzzy

O sistema de regras fuzzy é todo baseado em estruturas condicionais do tipo *If-then* (se-então). Uma regra simples pode ter a forma:

If x is A then y is B

Onde A e B são variáveis lingüísticas definidas por conjuntos fuzzy dentro dos domínios de x e de y, respectivamente. A parte *if* da regra "x is A" é chamada de *premissa* ou *antecedente*, enquanto que a parte *then* da regra "y is B" é chamada *conseqüente* ou *conclusão*. Para o exemplo:

If temperatura is alta then vazão is média

A variável *alta* pode ser representada por um número entre 0 e 1, (por exemplo, *alta* = 1 e *baixa* = 0), enquanto que *média* é representada por um conjunto fuzzy. O antecedente descreve qual o grau da regra aplicada, enquanto que o conseqüente relaciona a função de pertinência com uma ou mais variáveis de saída. O conjunto de regras em um sistema fuzzy é conhecido com *base de regras*.

1.6.3.4. Defuzzificação

A saída de um sistema fuzzy pode ser composta por um ou mais números, assim como uma ou mais funções. Para este último caso é necessário se fazer a *defuzzificação*, que consiste em fazer uma associação entre uma função de pertinência e um número. Existe uma centena de métodos para isto, mas os mais utilizados são o método da *centróide* e o da *média ponderada*. O primeiro método encontra a posição central (ou o centro de gravidade) de uma função ou de um conjunto delas, e o segundo atribui pesos para determinadas variáveis associadas a uma função e então calcula-se uma média ponderada.

1.6.3.5. Modelo fuzzy para o “problema da gorjeta”

Para tornar mais clara sua aplicação, será mostrado um exemplo *não químico* bem simples em que se deseja fazer uma modelagem entre dados de entrada e saída. Este sistema possui duas variáveis de entrada e uma de saída, sendo denominado “o problema da gorjeta” [19], e consiste em saber qual gorjeta será dada ao garçom ao fim de uma refeição servida em um restaurante, através da avaliação das qualidades do serviço e da comida. Vamos considerar o sistema norte-americano em que as gorjetas dadas variam entre 5 e 25% sobre o total da conta, sendo que 15% é o valor mais utilizado; serão dadas notas para avaliação do *serviço* e da *comida*, variando entre 0 e 10.

A base de regras para o problema da gorjeta será o seguinte [19]:

1. If *serviço* is *pobre* or *comida* is *rançosa*, then *gorjeta* is *pequena*.
2. If *serviço* is *bom*, then *gorjeta* is *média*.
3. If *serviço* is *excelente* or *comida* is *deliciosa*, then *gorjeta* is *generosa*.

Neste caso a escala da variável **serviço** vai de *pobre* até *excelente* (escala de 0 a 10); a variável **comida** vai de *rançosa* até *deliciosa* (de 0 a 10), e a variável **gorjeta** vai de *pequena* até *generosa* (de 5 a 25%). A aplicação do tratamento fuzzy está mostrada na Figura 20.

Pode-se observar nesta que para um exemplo em que o *serviço* tenha uma nota 7,3 e a *comida* tenha nota 7,5, a *gorjeta* será de 19,7%. O método de defuzzificação utilizado neste problema é o da *centróide*, por isto é possível observar na Figura 20 dois trapézios atravessados por uma barra rosa indicando o centro de gravidade destas duas figuras. Os triângulos azuis a direita têm a altura preenchida e definida como sendo o valor máximo da aplicação do operador OR para as funções de pertinência gaussianas e trapezoidais usadas, associadas respectivamente às variáveis *serviço* e *comida*. Os triângulos azuis representam a função de pertinência para a variável *gorjeta*.

Desta forma é possível estabelecer relações entre as funções de entrada e as de saída, para o problema da gorjeta. O sistema fuzzy utilizado neste exemplo também é conhecido como sistema de *Mandani* e é o mais empregado na modelagem de sistemas [19].

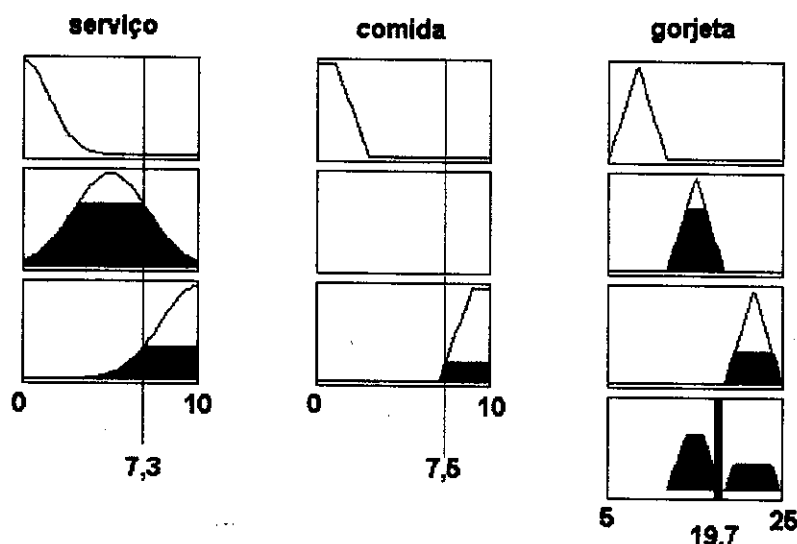


Figura 20- Diagrama mostrando o tratamento fuzzy para o “problema da gorjeta”, com as funções de pertinência utilizadas para as três variáveis, bem como os valores numéricos utilizados no exemplo.

1.6.3.6. O algoritmo “cluster estimation”

A definição da base de regras é a parte mais difícil do tratamento fuzzy de um sistema, mas existem algoritmos para auxiliar esta decisão. Stephen Chiu [21] desenvolveu um algoritmo baseado em *estimação de grupos* (“cluster estimation”), que permite construir uma base de regras fuzzy para um sistema através da solução de uma equação matricial algébrica, a qual relaciona os conjuntos de dados de entrada e saída. Este algoritmo gera um sistema fuzzy conhecido como sistema de *Tagaki-Sugeno* [22]. Ele é muito semelhante ao sistema de Mandani e difere apenas em relação a escolha das funções de saída, as quais são funções conhecidas como lineares. Por exemplo, um

sistema com duas entradas e uma saída poderia ser representado da seguinte maneira:

$$\text{If } x \text{ is } A \text{ and } y \text{ is } B \text{ then } z = p.x + q.y + r$$

onde A e B são conjuntos fuzzy enquanto p , q e r são constantes. Neste caso existe uma dependência linear entre z e x , e entre z e y . Como característica principal, o algoritmo “cluster estimation” calcula os melhores valores das constantes p , q e r e utiliza funções gaussianas para representar os conjuntos A e B .

O algoritmo “cluster estimation” faz uma procura por grupos de dados (também conhecidos como *clusters*) representados por um ponto em um espaço M dimensional [21]. Considerando-se uma coleção de n pontos $\{x_1, x_2, \dots, x_n\}$, podemos escrever uma equação:

$$P_i = \sum_{j=1}^n e^{-\alpha \|x_i - x_j\|^2} \quad (\text{eq. 20})$$

onde

$$\alpha = \frac{4}{r_a^2} \quad (\text{eq. 21})$$

e r_a é uma constante positiva chamada de “raio de influência” e é escolhida pelo usuário, sendo um valor entre 0 e 1; P_i é chamado de potencial do ponto x_i em relação ao ponto x_j . A medida do potencial de um ponto é uma função das suas distâncias em relação a todos os outros pontos. Um ponto com muitos pontos vizinhos terá um alto potencial de ponto. A próxima etapa deste algoritmo é fazer uma comparação entre os potenciais dos pontos e estabelecer quais os valores mais altos, sendo então definido o número de grupos. Cada regra fuzzy que descreve o comportamento de um sistema está associada a um grupo através de uma função relacionada com a coordenada do centro do grupo. Dado uma coleção de m centros de grupos $\{x_1^*, x_2^*, \dots, x_m^*\}$, considera-se que cada centro x_m^* possa ser decomposto em dois novos vetores: y_m^* (coordenadas dos

centros de grupos para o espaço de entrada) e z_m^* (coordenadas dos centros de grupos para o espaço de saída). Dado um vetor de entrada y , a função associada com a regra i é definida por:

$$\mu_i = e^{-\alpha \|y - y_i^*\|^2} \quad (\text{eq. 22})$$

onde α é o mesmo definido na equação 21. O vetor de saída z é dado por:

$$z = \frac{\sum_{i=1}^m \mu_i z_i^*}{\sum_{i=1}^m \mu_i} \quad (\text{eq. 23})$$

Considerando-se agora um sistema que tenha duas entradas e duas saídas, como por exemplo, duas integrais de linhas espectrais distintas relacionadas a duas concentrações de elementos diferentes, a estrutura condicional pode ser escrita como:

If Y_1 is A_1 and Y_2 is A_2 then Z_1 is B_1 and Z_2 is B_2

onde os Y_j são as variáveis de entrada, Z_j são as variáveis de saída, A_j são funções de pertinência do tipo gaussiana e B_j são funções de pertinência do tipo linear. O vetor z_j^* está relacionado diretamente às funções B_j . A relação entre os dados de entrada e saída neste algoritmo é representado por:

$$z_i^* = G_i y_i + h_i \quad (\text{eq. 24})$$

onde o vetor z_i^* é expresso como uma função linear da entrada do vetor y_i ; G_i é uma matriz de constantes, h_i é um vetor coluna de constantes também. Esta equação 24 está relacionada com a equação 23 e pode ser reescrita como:

$$z = \sum_{i=1}^m \rho_i z_i^* = \sum_{i=1}^m \rho_i (G_i y_i + h_i) \quad (\text{eq. 25})$$

com

$$\rho_i = \frac{\mu_i}{\sum_{j=1}^m \mu_j} \quad (\text{eq. 26})$$

Para este sistema com duas entradas e duas saídas, a equação 25 pode ser escrita na forma matricial do seguinte modo:

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \rho_{1,1} y_1 & \rho_{1,1} & \rho_{2,1} y_1 & \rho_{2,1} \\ \rho_{1,2} y_2 & \rho_{1,2} & \rho_{2,2} y_2 & \rho_{2,2} \end{bmatrix} \begin{bmatrix} G_1 \\ h_1 \\ G_2 \\ h_2 \end{bmatrix} \quad (\text{eq. 27})$$

Z **P** **G**

A resolução do sistema linear matricial acima (eq. 27) é feita através da estimativa de mínimos quadrados. Esta equação 27 tem a forma $PG = Z$, onde Z é a matriz dos valores de saída, P é a matriz de constantes e G é a matriz dos parâmetros a serem estimados. Nesta etapa o algoritmo tenta encontrar a solução:

$$G = (P^T P)^{-1} P^T Z \quad (\text{eq. 28})$$

Caso a equação 28 não possa ser resolvida, pois $(P^T P)$ não é inversível, o algoritmo fará uso de um procedimento conhecido como “estimativa de mínimos quadrados recursiva”.

Vamos considerar agora que no nosso sistema com duas entradas e duas saídas o algoritmo de estimação de grupos tenha encontrado duas regras (2 grupos), a matriz de parâmetros G da equação 27 terá a seguinte forma:

$$G = \begin{bmatrix} b_{1,1,1} & b_{1,1,2} \\ b_{2,1,1} & b_{2,1,2} \\ h_{1,1} & h_{1,2} \\ b_{1,2,1} & b_{1,2,2} \\ b_{2,2,1} & b_{2,2,2} \\ h_{2,1} & h_{2,2} \end{bmatrix} \quad (\text{eq. 29})$$

Assim, a base de regras para este sistema de duas entradas/duas saídas e as funções lineares de saída podem ser escritos como:

regra 1:

If $z_{1,1}$ is $A_{1,1}$ and $z_{1,2}$ is $A_{1,2}$

then $B_{1,1} = b_{1,1,1} y_1 + b_{2,2,2} y_2 + h_{1,1}$ and $B_{1,2} = b_{1,1,1} y_1 + b_{2,1,2} y_2 + h_{1,2}$

regra 2:

If $z_{2,1}$ is $A_{2,1}$ and $z_{2,2}$ is $A_{2,2}$

then $B_{2,1} = b_{1,2,1} y_1 + b_{2,2,1} y_2 + h_{2,1}$ and $B_{2,2} = b_{1,2,2} y_1 + b_{2,2,2} y_2 + h_{2,2}$

Os b 's e os h 's são as constantes otimizadas pelo algoritmo e são oriundas da matriz G_i , sendo então associadas às funções lineares de saída B_i do sistema.

Este algoritmo exige apenas que o usuário defina o valor do parâmetro r_s (equação 21). Matematicamente pode-se observar que para grandes valores de r_s (próximo de 1,0) menor será a quantidade de *grupos* ("clusters") encontrados no modelo, enquanto que, para pequenos valores de r_s (próximo de 0), maior será o número de *grupos* encontrados.

Finalmente, na literatura existem alguns trabalhos com aplicação de teoria fuzzy em química analítica [23-26], com destaque especial aos trabalhos desenvolvidos pelo prof. Matthias Otto da Universidade de Fraiburg na Alemanha, principalmente na área de calibração, reconhecimento de padrões e sistemas, auto-identificação e interpretação de dados químicos.

2. Objetivos

O objetivo deste trabalho foi o de utilizar alguns modelos quimiométricos, tais como redes neurais (com três algoritmos diferentes) e modelagem fuzzy, para tratar espectros de fluorescência de raios-X de energia dispersiva, a fim de se realizar a determinação simultânea de elementos químicos em diferentes tipos de amostras (misturas sintéticas e amostras naturais), principalmente naquelas onde ocorrem fluorescência secundária e sobreposição de linhas espectrais, problemas comuns relacionados com o emprego desta técnica.

Um segundo objetivo deste trabalho foi fazer um estudo comparativo entre os modelos, na tentativa de buscar um modelo mais generalizado para resolução da maioria dos problemas e também tentar identificar a forma do melhor tipo de conjunto de dados de entrada para os modelos, entre os dois tipos estudados: "scores" e integrais.

3. Parte experimental

Neste trabalho foi utilizado o espectrômetro de Fluorescência de Raios-X de Energia Dispersiva Spectrace 5000, instalado no Instituto de Química da Unicamp, em condições de irradiação específicas para cada amostra, conforme mostrado na Tabela 2.

Tabela 2 - Condições de irradiação específicas para cada matriz.

Matrizes	Voltagem do tubo (KV)	Corrente do tubo (mA)	Tempo de Irradiação (s)	Tipo de filtro	Atmosfera
aços	30	0,25	200	thin	vácuo
t. raras	20	0,07	100	alumínio	ar
mist. Pb/S	15	0,02	50	nenhum	vácuo
solos	30	0,07	200	thin	vácuo
plantas	15	0,03	150	nenhum	vácuo

t. raras: óxidos de terras raras; mist. Pb/S: misturas de sais de chumbo e enxofre.

Foram utilizados padrões certificados de várias partes do mundo fornecidos pelo Instituto de Geociências da Unicamp; padrões de diversos tipos de plantas (a maior parte de comestíveis) fornecidos pelo Instituto Agrônômico de Campinas; amostras de diferentes ligas de aço, misturas sintéticas de sais de chumbo e enxofre em sílica, misturas sintéticas de óxidos de Pr, Nd e Sm em sílica. As faixas de concentração dos elementos estudados nestas amostras estão mostradas na Tabela 3.1.

Tabela 3.1 - Faixas de concentrações dos elementos estudados neste trabalho em função do tipo de matriz, com a suas respectivas unidades.

Matriz	Elemento	Nível inferior	Nível superior
solos	Rb	3,6 µg/g	466 µg/g
	Sr	3 µg/g	790 µg/g
	Y	5 µg/g	184 µg/g
	Zr	30 µg/g	780 µg/g

mist. Pb/S	S	2 % (m/m)	17 % (m/m)
	Pb	2 % (m/m)	17 % (m/m)
t. raras	Pr	5 µg/g	20 µg/g
	Nd	5 µg/g	21 µg/g
	Sm	5 µg/g	21 µg/g
aços	Mn	0,2 %	0,9 %
	Cr	0,02 %	0,031 %
	Fe	97,57 %	98,40 %
plantas	Si	145 mg/Kg	37500 mg/Kg
	P	2,19 mmol/Kg	224 mmol/Kg
	S	11,0 mmol/Kg	451 mmol/Kg
	K	11,3 mmol/Kg	1407 mmol/Kg
	Ca	12,6 mmol/Kg	2092 mmol/Kg

Os valores completos das concentrações de todas as amostras utilizadas estão mostrados nas Tabelas 3.2 a 3.6, para cada tipo de matriz.

Tabela 3.2 - Concentrações das amostras de aço. Todas as unidades estão em %.

	Cr	Mn	Fe
1	0,021	0,57	98,22
2	0,024	0,531	97,62
3	0,023	0,608	98,12
4	0,028	0,451	97,70
5	0,022	0,255	98,32
6	0,024	0,277	97,64
7	0,023	0,232	97,98
8	0,022	0,23	97,75
9	0,02	0,214	98,41
10	0,025	0,539	97,78
11	0,023	0,618	98,33
12	0,021	0,643	98,40
13	0,031	0,939	97,57

Parte experimental

14	0,025	0,41	98,04
15	0,025	0,918	97,70
16	0,026	0,606	98,33

Tabela 3.3 - Concentrações das amostras de sais de chumbo e enxofre. Todas as unidades estão em %.

	S	Pb
1	2,0	14,01
2	2,0	16,99
3	2,0	1,99
4	5,0	4,97
5	3,0	2,0
6	16,0	2,01
7	2,0	3,0
8	7,0	8,0
9	14,01	7,99
10	8,0	7,0
11	8,0	11,0
12	9,0	3,99
13	11,0	6,01
14	4,0	9,0
15	3,0	11,99
16	2,0	15,0
17	15,0	2,01
18	6,0	3,99
19	8,0	2,01
20	10,99	1,99
21	2,01	5,02
22	14,0	5,0
23	16,99	4,01
24	5,0	7,97
25	10,99	7,98
26	16,0	7,99
27	12,01	2,99
28	4,0	11,0

29	5,0	1,96
30	14,01	1,98
31	16,99	2,0
32	2,0	7,99
33	2,0	11,0
34	8,0	5,0
35	11,0	5,01
36	4,99	11,0
37	16,0	3,0
38	11,0	3,99

Tabela 3.4 - Concentrações das amostras de solos. Todas as unidades estão em µg/g.

	Rb	Sr	Y	Zr
1	11	403	27,6	179
2	74	336	24	185
3	149	108	25	219
4	127	183	40	290
5	4,5	145	18	41
6	20	194	24	94
7	90	35	40,6	110
8	8,5	266	14	108
9	77	331	40	760
10	39	222	16,4	156
11	69	168	22	270
12	73	400	26	125
13	3,6	110	114	550
14	185	570	16	235
15	152	3	184	780
16	180	185	31,6	121
17	297	16	88,2	101
18	66	372	17,2	143
19	6,2	178	24,9	51,4
20	13	395	27	98,3
21	257	30	45,4	101

Parte experimental

22	297	8	51,3	97,2
23	68	252	18,1	119
24	36	294	21,3	119
25	4	321	10,75	33,5
26	466	106	62	167
27	37,6	790	9,3	99
28	140	155	25	245
29	88	187	21,7	219
30	85	380	15	246
31	75	77	39	500
32	17	41,5	21	272
33	237	39	18,8	220
34	15,8	26	26,6	318
35	96	236	26	229
36	5	32	5	30
37	24	250	19	134
38	85	220	44	254
39	78	240	30	178
40	28	110	23	105
41	30	170	42	218
42	104	400	37	185
43	68	230	36	196
44	39	350	24	190

Tabela 3.5 - Concentrações das amostras de plantas. Todas as unidades estão em mmol/Kg, com exceção do elemento Si que está em mg/Kg.

	S	K	Ca	P	Si
1	123,1	1407	321	123	1002
2	80,2	647	125	105	981
3	226,8	576	409	108	695
4	58,2	291	207	79,1	1968
5	49,9	284	163	81,4	877
6	40,3	221	123	62,3	551
7	25,8	237	59,1	28,6	514
8	59,88	117	11,4	106	377

9	65,84	250	201	53,2	18494
10	75,7	399	75,6	27,1	13888
11	158,2	555	184	123	27674
12	67,7	621	80,2	79,4	886
13	130,7	1168	484	224	5233
14	51,4	734	74,5	102	2156
15	61,6	602	291	65,7	1872
16	164,6	669	137	125	6791
17	42	240	157	60,8	1661
18	149,1	672	567	75,2	1188
19	56,2	129	14,5	123	754
20	76,6	290	160	55,2	9475
21	224,3	267	76,1	62,6	4469
22	71	372	69,6	29,2	14014
23	54,9	635	81,9	85,8	3513
24	35,52	139	9,02	124	145
25	65,7	633	100	64,4	6860
26	49,6	354	135	83,3	2780
27	451	1101	2092	195	37509
28	117	1357	399	142	1285
29	11,02	11,3	12,6	2,19	1125
30	125,8	397	467	90,5	1207
31	132,3	1312	753	169	2862
32	322	729	196	117	671

Tabela 3.6 - Concentrações das amostras de terras raras. As unidades estão em µg/g.

	Nd	Sm	Pr
1	5,09	5,2	5,73
2	5,23	11,05	5,24
3	5,97	15,15	5,76
4	5,37	21,06	5,59
5	10,18	5,31	5,61
6	10,14	11,84	5,75
7	10,66	15,26	5,53

8	11,32	20,26	5,13
9	17,31	5,17	5,53
10	16,17	10,76	5,49
11	15,91	15,17	5,73
12	15,74	20,85	5,24
13	21,15	6,53	5,22
14	20,97	10,28	5,32
15	20,59	16,03	5,56
16	20,15	21,14	5,5
17	6,11	5,43	10,15
18	6,91	10,23	10,77
19	5,34	15,16	10,22
20	6,17	20,43	10,09
21	10,54	5,48	10,53
22	11,17	10,91	10,57
23	10,23	15,76	10,08
24	10,18	20,1	10,46
25	14,98	5,11	10,34
26	15,2	10,17	10,11
27	15,4	17,61	11,05
28	15,41	20,15	10,8
29	20,73	6,19	10,69
30	20,43	10,11	10,74
31	20,64	15,11	10,28
32	20,77	20,37	10,48
33	5,86	5,6	16,47
34	5,31	10,4	15,41
35	5,06	15,2	15,36
36	5,52	20,15	15,27
37	10,37	5,23	15,36
38	10,2	12,11	15,39
39	10,77	15,26	15,38
40	10,63	20,2	15,39
41	15,47	5,2	15,22
42	15,38	10,12	15,05
43	15,15	15,34	15,33

44	15,12	20,15	15,33
45	20,33	5,89	15,16
46	21,37	10,11	15,1
47	20,78	12,16	15,05
48	20,35	20,15	15,1
49	5,6	5,43	20,24
50	4,94	10,03	20,4
51	5,94	15,31	20,11
52	5,32	20,49	20,17
53	10,09	6,28	20,32
54	10,43	10,14	20,26
55	10,29	15,01	20,2
56	10,37	20,57	20,09
57	15,2	7,64	20,16
58	16,65	10,34	20,14
59	15,66	15,86	20,2
60	15,34	20,22	20,1
61	20,57	5,77	20,23
62	20,92	10,06	20,11
63	20,86	17,22	20,37
64	20,26	20,14	20,28

As amostras sintéticas de sais de chumbo e enxofre e de óxidos de terras raras formaram misturas homogêneas com sílica, sendo irradiadas sem nenhum pré-tratamento. As amostras certificadas de solos apresentaram-se na forma de discos prensados, conforme um procedimento de preparação específico e também foram irradiadas sem nenhum pré-tratamento. Este procedimento de preparação garante que todas as partículas das amostras possuam o mesmo tamanho e que estejam igualmente homogeneizadas em relação a todos os constituintes. As amostras de plantas apresentaram-se na forma de um pó seco, homogêneo, o qual foi prensado no porta amostra antes de ser irradiado. As amostras de aço apresentaram-se na forma de tarugos cilíndricos, os quais foram lixados com esponja de aço e polidos com solução de sílica de partículas 0,1 μm e 0,01 μm até atingir uma superfície espelhada, o que garante uma

uniformidade nas partículas.

O trabalho do tratamento de todos os dados com todos os modelos quimiométricos foi realizado em um microcomputador PC Pentium™ de 133 MHz com 16 MBytes de RAM, utilizando o software Matlab da Mathworks Inc. versão 4.2 com os pacotes "Neural Network Toolbox" versão 2.0, "Fuzzy Logic Toolbox" versão 1.0 e "PLS Toolbox" versão 1.0

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 [28]:

$$RMS = \sqrt{\frac{\sum_{i=1}^n (C_i - \hat{C}_i)^2}{n}} \quad (\text{eq. 30})$$

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 30 acima é bastante difundida em trabalhos de Quimiometria [28] e é uma representação dos erros absolutos médios para as concentrações, calculados para um conjunto de previsão. Os erros calculados por esta equação mostrados nas tabelas representam na verdade uma média entre os erros calculados para cada elemento nos modelos.

Todas as concentrações usadas pelos modelos foram normalizadas entre 0 e 1 devido basicamente a três motivos:

- 1) Características e/ou "exigências" matemáticas dos modelos;
- 2) Diferentes faixas de concentração entre os elementos em uma mesma amostra;
- 3) Possibilidade de comparação dos erros entre modelos e amostras diferentes.

O procedimento de normalização dos dados (entre 0 e 1) pode ser

explicado através da aplicação da seguinte equação:

$$C_n[i] = \frac{C[i] - C_{\min}}{C_{\max} - C_{\min}} \quad (\text{eq. 31})$$

onde,

$C_n[i]$ é o i ésimo valor de concentração normalizada ;

$C[i]$ é o i ésimo valor de concentração do conjunto de dados;

C_{\min} é a menor concentração do conjunto de dados;

C_{\max} é a maior concentração do conjunto de dados.

A aplicação dos modelos quimiométricos aos espectros de EDXRF pode ser entendida de uma maneira bastante simplificada, ou seja, como se cada espectro de cada tipo de amostra pudesse ser representado como um par ordenado de matrizes (do tipo $[x, y]$ onde x é o dado de entrada e y o dado de saída) de duas classes diferentes: ["scores", concentrações] e [integrais, concentrações].

A utilização de "scores" como dados de entrada é realizada através da aplicação do método conhecido como Análise de Componentes Principais (PCA) [19] aos espectros de EDXRF. Esta análise consiste em um tipo de técnica conhecida como "compressão de dados", ou seja, transforma-se o conjunto de dados originais em um novo conjunto menor, mas com informação estatística relevante. Matematicamente, o PCA produz um redimensionamento das variáveis em um espaço vetorial, utilizando elementos deste espaço chamados de "scores" para a representação destas variáveis.

A escolha das amostras utilizadas como conjunto de validação foi feita de maneira aleatória, para todas as aplicações.

A escolha do número de "scores", como conjunto de dados de entrada dos modelos, foi feita e determinada através da aplicação do algoritmo "Cross Validation" (CV) ou validação cruzada, variante "*leave one out*" [7, 19].

Os "outliers" são as amostras *anômalas* que devem ser retiradas do conjunto de calibração para não prejudicar a etapa de treinamento dos modelos.

O critério para determinar qual amostra pode ser considerada um “outlier” é dado por testes estatísticos presentes no “PLS Toolbox” do programa Matlab™.

A *arquitetura* (também chamada de *configuração*) das redes neurais em relação ao número de neurônios, foi definida na forma escrita X-Y-Z, onde X é o número de neurônios na camada de entrada, Y é o número de neurônios na camada intermediária e Z é o número de neurônios na camada de saída.

As redes BP, LM e RBF utilizadas neste trabalho, possuem, para efeito de simplificação, apenas uma camada escondida. As funções de transferência utilizadas entre as camadas para as duas primeiras redes são as seguintes: *sigmoidal* entre a camada inicial e a camada escondida, e *linear* entre a camada escondida e a camada de saída. As funções de transferência utilizadas para a rede RBF são as seguintes: *gaussiana* entre a camada inicial e a camada escondida, e *linear* entre a camada escondida e a camada de saída. Os valores de “learning rate” e “momentum” utilizados pela rede neural BP foram de 0,01 e 0,9 respectivamente [11]. Estes valores são sugeridos pelo “User’s Guide” do Toolbox de redes neurais [11], pois promovem os melhores desempenhos para as redes.

As redes neurais BP e LM são treinadas três vezes utilizando os conjuntos de calibração considerando os parâmetros acima e também o erro SSE máximo de 0,001 e um número de 5000 iterações. Quando um destes dois parâmetros é atingido a etapa de treinamento da rede é encerrada. O programa Matlab™ utiliza então uma função que faz uso dos pesos matemáticos encontrados na etapa de treinamento para calcular os valores de concentração do conjunto de previsão.

O teste estatístico F utilizado para comparação da precisão de resultados, com um grau de 95% de confiança, é feito através da relação entre os erros RMS calculados por dois modelos, segundo a relação [29]:

$$F = \frac{(RMS_1)^2}{(RMS_2)^2} = x \quad (\text{eq. 32})$$

em que o valor de x é comparado com um valor crítico tabelado [29] (segundo o número de amostras ou graus de liberdade); sendo ele menor que o tabelado, não há estatisticamente diferença entre o erro RMS1 e o erro RMS2. Foi adotado o critério de utilizar o maior erro RMS no numerador da equação 32 e o menor erro RMS no denominador da mesma equação.

Os *erros relativos* calculados para as concentrações dos elementos e mencionados em diversas tabelas, obedecem a seguinte equação:

$$\text{err} = \left(\frac{\text{real} - \text{calc}}{\text{real}} \right) \times 100\% \quad (\text{eq. 33})$$

onde,

real é a concentração verdadeira dos elementos;

calc é a concentração calculada dos elementos.

4. Resultados e discussão

Neste trabalho a utilização de integrais como dados de entrada foi sugerida pois, para os espectros de EDXRF, as linhas espectrais correspondentes a emissões características do 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.

Também foi realizado um estudo da extrapolação dos modelos em relação a capacidade de previsão de concentrações dos elementos fora do conjunto de calibração, ou seja, foram utilizadas amostras com valores muito menores e muito maiores que a maioria usada como conjunto de calibração. Para esta situação, foram utilizados apenas os "scores" como conjunto de dados de entrada e, as amostras "outliers" foram incluídas no conjunto de calibração de modo a dificultar o processo de modelamento.

A aplicação dos modelos fuzzy neste trabalho foi feita através da utilização de funções específicas para este fim, presentes no "Fuzzy Logic Toolbox". 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 de "raio de influência" (equação 20), sendo este último um número que varia entre 0 e 1.

Para a rede neural RBF, a busca do menor erro para o treinamento da rede é feito através variação do parâmetro chamado de "spread" que é um número que varia entre 0,01 e 1,5. Esta faixa foi 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 valores de erro de treinamento razoáveis para a estimativa das concentrações.

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 MatlabTM, através de uma função randômica. 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.

O procedimento chamado de *otimização da arquitetura da rede neural* 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 com 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 “overfitting”.

Os modelos PLS e PCR foram utilizados neste trabalho apenas para comparação dos resultados. Estes dois modelos são bastante conhecidos, possuem uma vasta gama de aplicações na literatura e são muito simples de serem utilizados.

Os testes F aplicados aos modelos geralmente mostram uma comparação entre os erros encontrados pelos modelos PLS e PCR em relação ao melhor resultado obtido por um outro modelo (rede neural ou modelo fuzzy), ou também podem mostrar a comparação entre os dois melhores resultados obtidos por dois diferentes modelos, para um determinado tipo de amostra.

As escalas das figuras (em que se apresenta a variação de algum parâmetro dos modelos contra o erro RMS) são diferentes para cada modelo e para cada aplicação, pois o objetivo da sua apresentação foi o de mostrar, com clareza, a posição do menor erro e o seu comportamento, em relação a variação dos parâmetros.

4.1. Misturas de sais de chumbo e enxofre

Uma primeira aplicação dos modelos quimiométricos para este trabalho foi feita através de 38 misturas sintéticas de sais contendo chumbo ($\text{Pb}(\text{NO}_3)_2$) e enxofre (K_2SO_4) em matriz de sílica. Este sistema é um caso típico de sobreposição de linhas espectrais correspondentes as linhas $L\alpha$ do chumbo (10,02 a 10,51 keV) e a sobreposição das linhas $K\alpha$ do enxofre e $M\alpha$ do chumbo (2,07 a 2,50 keV), como mostrado na Figura 21. Para esta amostra foram utilizados 24 espectros para calibração, 7 para validação e foram retirados 7 "outliers", para a determinação simultânea de Pb e S.

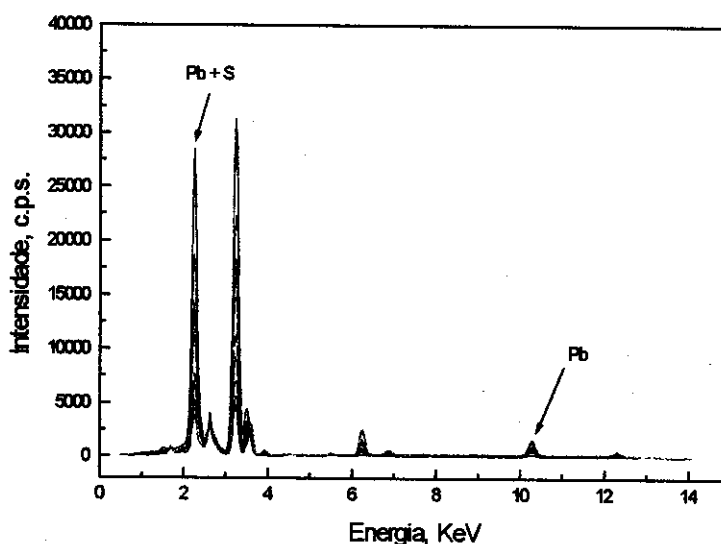


Figura 21 - Parte de espectros de EDXRF utilizados neste trabalho, mostrando as linhas espectrais dos elementos Pb e S.

Na análise CV dos espectros das amostras foi determinado um número de 6 componentes principais representando por 6 "scores", por espectro. 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 4.

Tabela 4 - 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
PCR	0,0863	0,0559
PLS	0,0863	0,0559

De acordo com a Tabela 4, os erros menores foram obtidos através da utilização de integrais para a rede BP. As arquiteturas das redes que apresentaram os resultados da Tabela 4 são as seguintes: 6-3-2 para BP e LM com "scores", 6-18-2 para RBF com "scores", 2-4-2 para BP e LM com integrais, 2-15-2 para RBF com integrais. Um teste F aplicado aos valores obtidos pelo modelo fuzzy e pelos modelos PLS e PCR em relação a rede BP (integrais), indica haver diferença significativa ao nível de 95% apenas em relação ao segundo modelo (considerando-se neste caso, $n = 38$ então $F_{crit.} = 1,71$) :

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

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, como mostrado na Tabela 5.

Tabela 5 - Erros relativos para a rede BP com integrais (RMS: 0,0229).

	S			Pb		
	real	calc.	err.	real	calc.	err.
1	14,01	14,32	-2,23	1,97	1,90	3,50
2	16,99	16,71	1,60	2,001	1,91	4,45
3	7,99	7,995	0,01	4,99	4,65	6,85
4	10,99	11,13	-1,21	5,01	5,20	-3,84
5	4,99	4,47	10,3	11,001	11,02	-0,239
6	15,99	16,37	-2,37	2,99	3,13	-4,73
7	11,00	10,43	5,16	3,99	4,73	-18,5
RMS	0,0241			0,0215		

calc.: calculado; err.: erro relativo

As figuras apresentadas a seguir (Figuras 22 a 25) mostram, em forma gráfica, os modelos desenvolvidos e os resultados mostrados na Tabela 4.

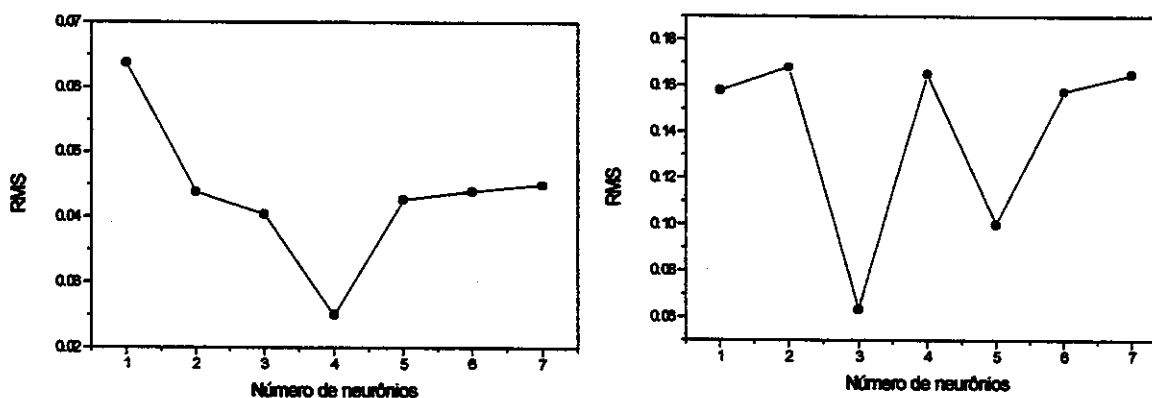


Figura 22 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP. O gráfico da esquerda representa o conjunto de entrada como integrais, e o da direita como "scores".

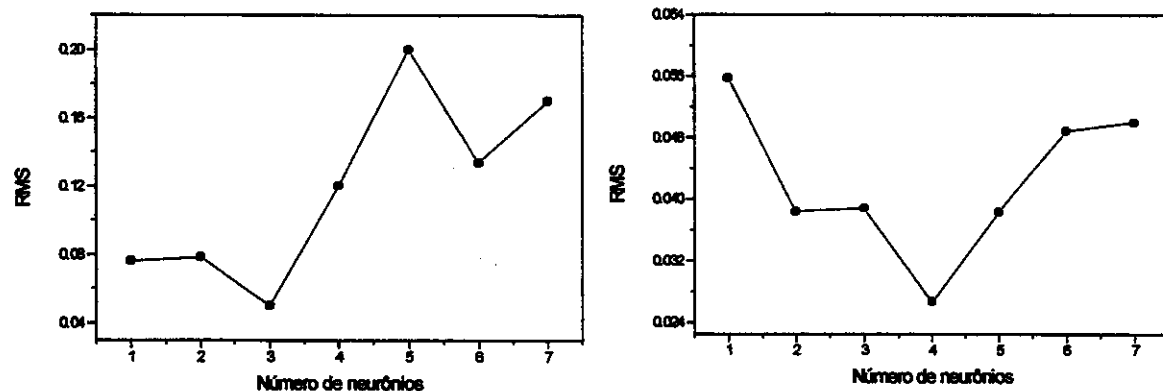


Figura 23 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural LM. O gráfico da esquerda representa o conjunto de entrada como "scores", e o da direita como integrais.

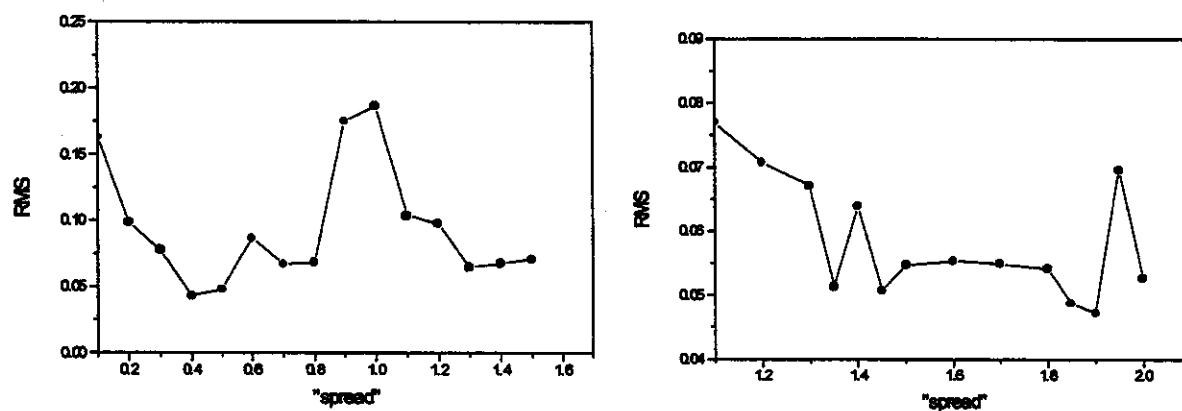


Figura 24 - Erros RMS calculados contra o parâmetro "spread", para a rede neural RBF. O gráfico da esquerda representa o conjunto de entrada como "scores", e o da direita como integrais.

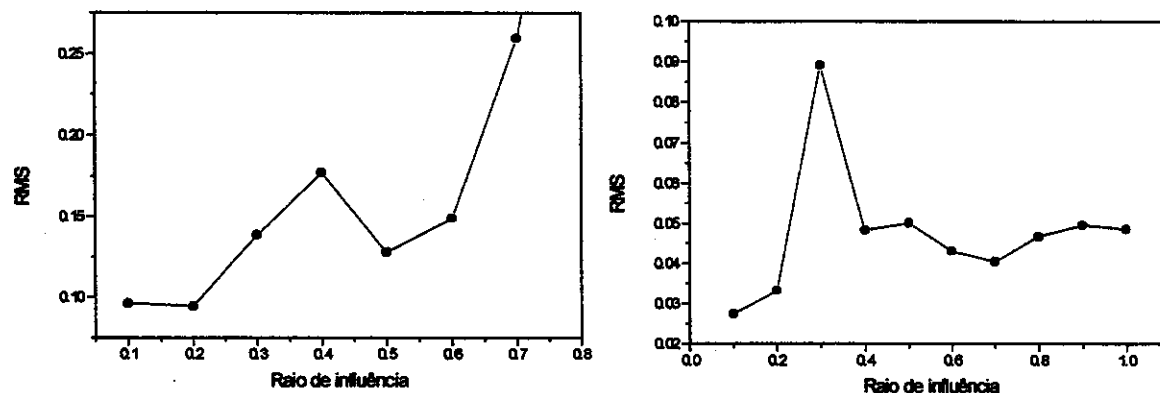


Figura 25 - Erros RMS calculados contra o parâmetro "raio de influência", para o modelo fuzzy. O gráfico da esquerda representa o conjunto de entrada como "scores", e o da direita como integrais. Na figura da esquerda é mostrado apenas a região entre 0 e 0.8, pois o erro RMS cresce além da escala do gráfico.

As figuras 22 e 23 mostram o erro RMS contra a adição de neurônios na camada intermediária das redes neurais. É possível observar que existe um tipo de comportamento que se repete: existe um número ideal em que se encontra o menor erro (os vales mais baixos), depois o erro torna a aumentar, em alguns casos até de forma brusca. Este comportamento é observado para as redes neurais BP e LM. Isto ocorre de maneira semelhante com a rede RBF e com o modelo fuzzy, mas os parâmetros são diferentes. Os erros para estes dois últimos modelos também são determinados pelos pontos mais baixos nas Figuras 24 e 25, mas a tendência do erro nas vizinhanças do "vale" não é muito bem definida.

Para este conjunto de dados foi feita também a extrapolação dos modelos, utilizando-se 4 espectros correspondentes às maiores e menores concentrações dos dois elementos e os outros 34 espectros como conjunto de calibração. Os dados de entrada utilizados são apenas os "scores". Os resultados são mostrados na Tabela 6.

Tabela 6 - Erros RMS calculados para os modelos utilizados para misturas de sais de chumbo e enxofre, para extrapolação dos modelos. Somente "scores" são usados com dados de entrada.

rede neural BP	0,115
rede neural LM	0,0694
rede neural RBF	0,124
modelo Fuzzy	0,0876
PCR	0,172
PLS	0,172

As configurações das redes utilizadas foram os seguintes: rede LM (6-3-2), rede RBF (6-26-2) e rede BP (6-5-2). A rede LM apresentou melhor desempenho, embora o modelo fuzzy esteja bem próximo. Este resultado é significativamente diferente em relação aos obtidos com os modelos PLS e PCR e a rede BP, de acordo com o teste F ($F_{crit.} = 1,71$):

$$F_1 = \frac{(\text{RMS}_{\text{BP}})^2}{(\text{RMS}_{\text{LM}})^2} = \frac{(0,115)^2}{(0,0694)^2} = 2,75 \quad F_2 = \frac{(\text{RMS}_{\text{PLS}})^2}{(\text{RMS}_{\text{LM}})^2} = \frac{(0,172)^2}{(0,0694)^2} = 6,14$$

Neste tipo de amostra a extrapolação dos modelos mostrou erros baixos pois a variação das concentrações para os dois elementos não é muito grande. Os erros relativos para as concentrações do melhor modelo da Tabela 6 estão mostrados na Tabela 7.

Tabela 7 - Erros relativos para a extrapolação da rede LM (RMS: 0,0694).

	S			Pb		
	real	calc.	err.	real	calc.	err.
1	2,00	1,69	15,1	16,99	14,58	14,2
2	1,99	2,04	-2,56	2,99	2,35	21,4
3	16,99	17,17	-1,09	4,00	2,80	30,1
4	5,00	6,33	-26,6	1,96	2,32	-18,4
RMS	0,0459			0,0929		

calc.: calculado; err.: erro relativo

As Figuras 26 e 27 mostram os resultados gráficos que originaram a Tabela 6. As redes BP e LM possuem um comportamento em que o menor erro fica situado em um “vale” sendo que o erro tende a crescer nas vizinhanças do mesmo. Os modelos fuzzy e as redes RBF apresentam um tipo de comportamento semelhante, mas a tendência do erro nas vizinhanças do “vale” não é muito bem definida.

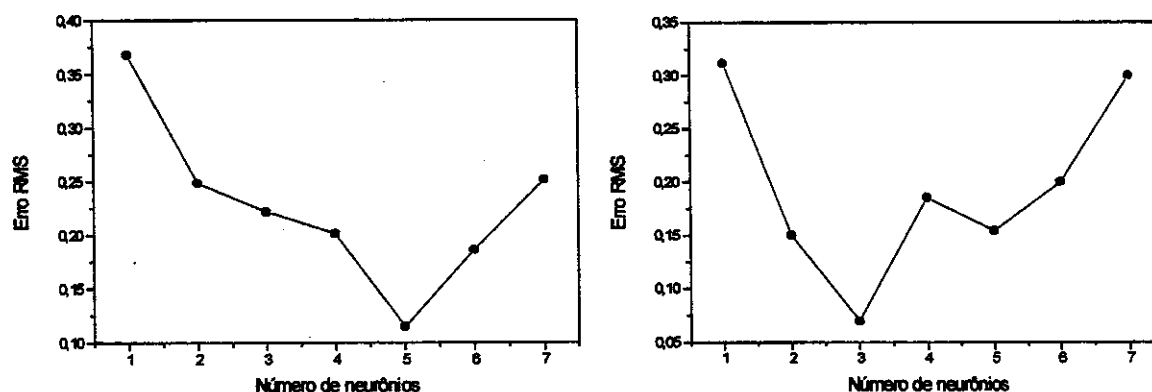


Figura 26 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP (a esquerda) e para a rede neural LM (a direita).

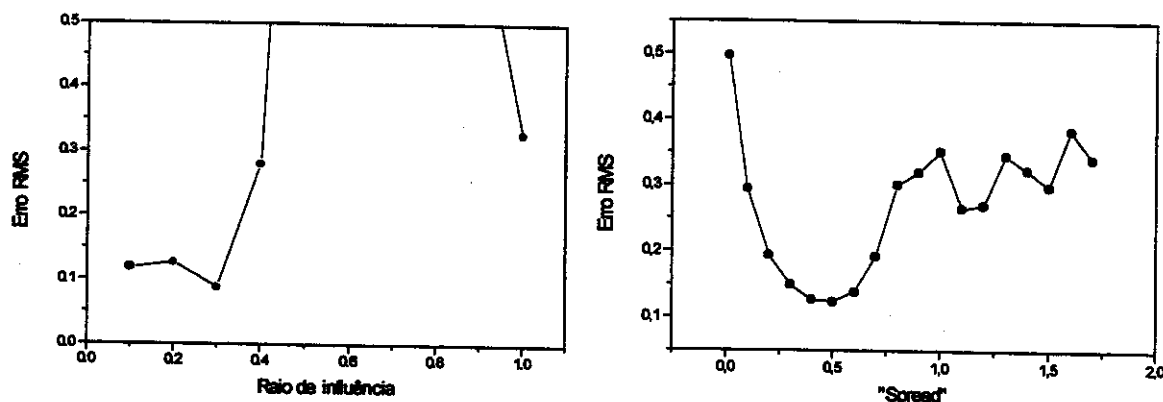


Figura 27 - Erros RMS calculados contra o parâmetro "raio de influência", para o modelo fuzzy (a esquerda), nesta figura não é mostrado a região entre 0,5 e 0,9, pois o erro RMS cresce além da escala do gráfico. Erros RMS calculados contra o parâmetro "spread", para a rede neural RBF (a direita).

4.2. Amostras de solos

Foi feito um estudo com padrões certificados de solos com o objetivo de verificar a calibração de espectros contendo elementos vizinhos da Tabela Periódica, especificamente Rb, Sr, Y e Zr.

A Figura 28 mostra um espectro típico de EDXRF com a sobreposição das linhas espectrais da região de interesse desses 4 metais. Este tipo de amostra foi escolhida 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 a linha base. Esta amostra utilizada como branco foi uma amostra de solo em que as concentrações dos quatro elementos eram muito pequenas.

Especificamente para este tipo de amostra a aplicação dos modelos foi feita considerando-se duas situações: (1) espectro completo 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.

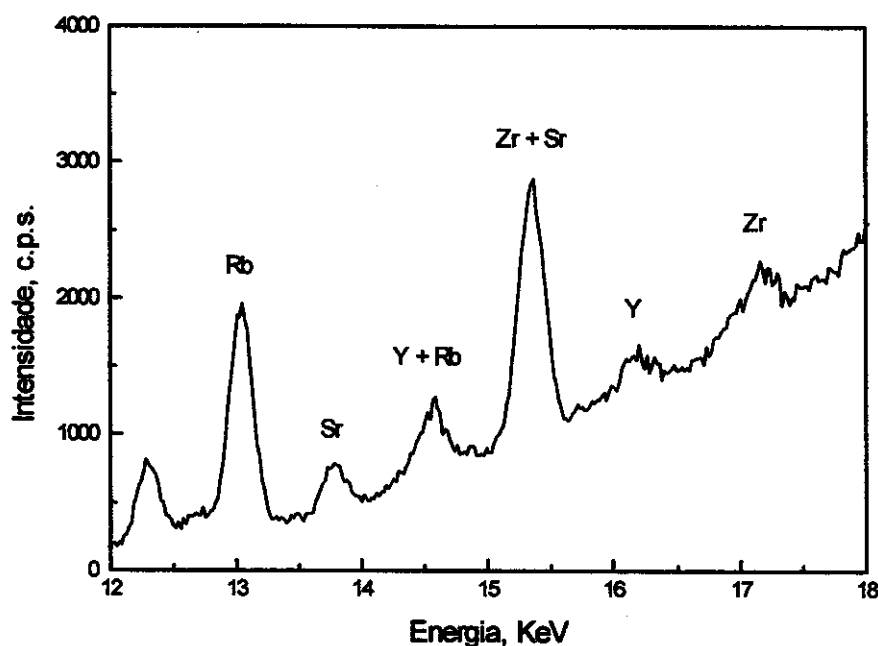


Figura 28 - Parte de um espectro de EDXRF típico, mostrando as linhas espectrais sobrepostas dos quatro elementos estudados.

Primeiramente considerou-se a região onde aparecem as principais linhas espectrais situadas entre 12,8 e 15,9 KeV. Na análise CV dos espectros das amostras foi determinado um número de 4 componentes principais representado por 4 "scores", por espectro. 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 8.

Tabela 8 - 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.

	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
PCR	0,107	0,113
PLS	0,107	0,113

A rede BP apresentou o menor erro RMS quando 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 obtidos pelos modelos PCR e PLS para os dois tipos de dados de entrada ($n = 44$, $F_{crit.} = 1,35$):

$$F_1 = \frac{(RMS_{PLS})^2}{(RMS_{BP})^2} = \frac{(0,107)^2}{(0,0741)^2} = 2,09 \quad F_2 = \frac{(RMS_{PLS})^2}{(RMS_{BP})^2} = \frac{(0,113)^2}{(0,0650)^2} = 3,02$$

As arquiteturas das redes que apresentaram os resultados da Tabela 8 são as seguintes: 4-27-4 para RBF com "scores"; 4-5-4 para BP e 4-3-4 para LM com "scores"; 4-3-4 para BP com integrais; 4-2-4 para LM com integrais e 4-25-4 para RBF com integrais.

No segundo tipo de aplicação, considerando o espectro completo, a análise CV das amostras determinou um número de 6 componentes principais representado por 6 "scores", por espectro. 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 a aplicação anterior, já mostra que realmente existe uma maior correlação entre os dados de entrada e saída. Foram utilizados os mesmos 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 9.

Tabela 9 - Erros RMS calculados para os modelos utilizados para amostras de solos, considerando-se o número de "scores" igual a 6.

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

A rede LM apresentou o menor erro RMS para esta segunda aplicação. As arquiteturas das redes que apresentaram os resultados da Tabela 9 são as seguintes: 6-21-4 para RBF, 6-8-4 para BP e 6-3-4 para LM. 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 pelos modelos PCR e PLS ($n = 44$, $F_{crit.} = 1,35$):

$$F_1 = \frac{(RMS_{PLS})^2}{(RMS_{BP})^2} = \frac{(0,0643)^2}{(0,0346)^2} = 3,45 \quad F_2 = \frac{(RMS_{PLS})^2}{(RMS_{LM})^2} = \frac{(0,0643)^2}{(0,0319)^2} = 4,06$$

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:

$$F_1 = \frac{(RMS_{BP})^2}{(RMS_{LM})^2} = \frac{(0,0650)^2}{(0,0319)^2} = 4,15$$

A Tabela 9 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 8). Pode-se concluir que existe, para este caso, uma contribuição das outras linhas espectrais no processo de estimação das concentrações do Rb, Sr, Y e Zr.

O erros relativos obtidos no cálculo das concentrações pela rede LM, que

foi o melhor modelo, estão mostrados na Tabela 10. Com exceção de seis valores de erros altos em 4 amostras (os quais apresentam o mesmo comportamento em todos os modelos, indicando um possível problema com a amostra) os erros relativos são menores que 20%.

Tabela 10 - Erros relativos obtidos para a rede LM (RMS: 0,0319).

	Rb			Sr			Y			Zr		
	real	calc.	err.	real	calc.	err.	real	calc.	err.	real	calc.	err.
1	30	17	43,3	170	131	22,9	42	33,	21,0	218	208	4,5
2	78	62	19,9	240	229	4,5	30	30	0	178	189	-6,7
3	85	59	29,6	220	207	5,7	44	40	8,2	254	272	-7,2
4	15,8	17,9	-13,3	26	93	-258	26,6	48	-82,7	318	326	-2,5
5	90	94	-5,2	35	82	-134	40,6	35	11,6	110	156	-42,1
6	69	74	-7,2	168	116	30,4	22	52	-138	270	330	-22,4
7	28	56	-101	110	65	40,4	23	32	-40,9	105	146	-39,2
8	257	251	2,2	30	32	-9,0	45,4	46	-2,90	101	118	-17,7
9	74	87	-17,7	336	337	-0,50	24	26	-10,8	185	187	-1,20
10	140	113	19,2	155	120	22,6	25	49	-98,8	245	279	-14,0
	RMS	0,0213			0,0481			0,0190			0,0394	

calc.: calculado; err.: erro relativo

Para estas amostras de solos, a rede neural BP foi aplicada também de uma terceira maneira distinta, em relação ao conjunto de dados de entrada, em que foram utilizadas diretamente pontos dos espectros, ou seja 131 pontos (131 neurônios na camada de entrada), correspondente a região entre 12,8 keV e 15,9 keV.

Os resultados obtidos para esta aplicação são mostrados na Figura 29. Eles não foram muito bons, mostrando um erro maior do que os que estão mostrados na Tabela 10. 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.

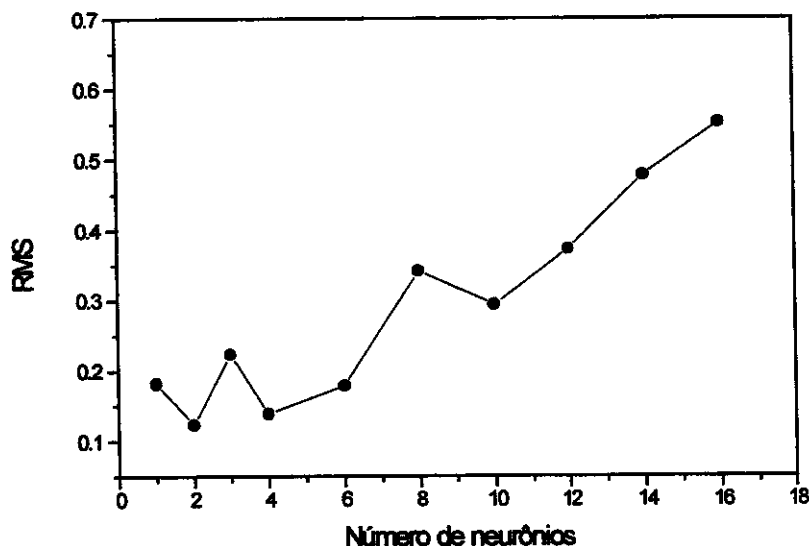


Figura 29 - Erros RMS calculados contra o número de neurônios na camada intermediária. A melhor configuração da rede BP é 131-2-4. O menor erro é de 0,123, bem superior ao encontrado utilizando-se "scores" ou integrais na entrada.

Os resultados da aplicação das redes neurais para "scores" e integrais como dados de entrada, para as amostras de solos, apresentam um comportamento muito semelhante ao observado na Figura 29 e eles estão mostrados na Figura 30.

Existe sempre um 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 "underfitting" e "overfitting", antes e depois de atingir a melhor configuração, respectivamente.

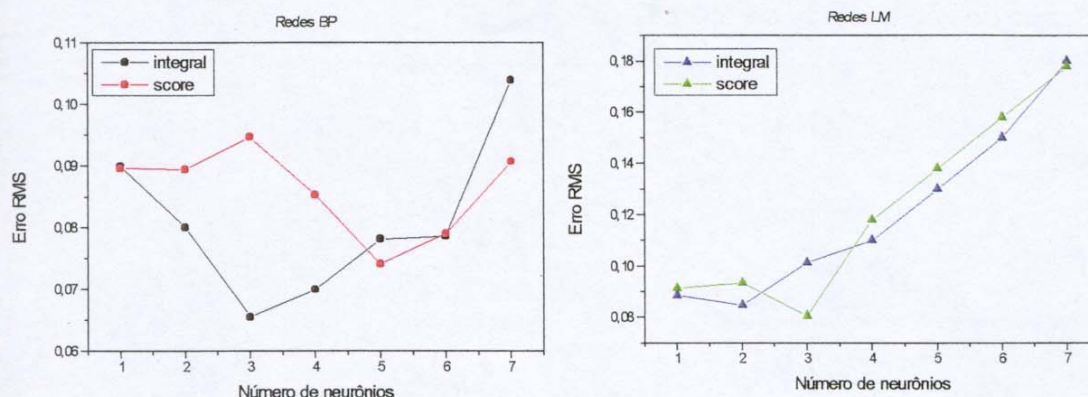


Figura 30 - Erros RMS calculados contra o número de neurônios na camada intermediária para as redes BP e LM, utilizando-se "scores" e integrais como conjunto de dados de entrada, considerando-se a região das linhas dos 4 metais.

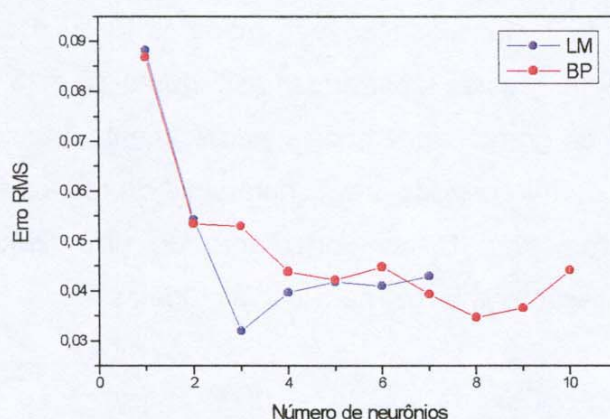


Figura 31 - Erros RMS calculados contra o número de neurônios na camada intermediária para as redes BP e LM, utilizando-se "scores" como conjunto de dados de entrada, considerando-se o espectro completo.

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 confirmado também, em função do tipo de amostra. Após este número, a possibilidade de ocorrer "overfitting" é muito grande. Como exemplo, foi feito um estudo da variação do erro RMS com o número de iterações para a rede neural BP apresentada na Figura 29 e de configuração 131-2-4, o resultado está mostrado na Figura 32. O comportamento da variação do erro apresentado nesta figura ocorre também de forma semelhante, quando se utiliza os "scores" ou as integrais como conjunto de dados de entrada, ou seja,

independe do tipo de dado de entrada.

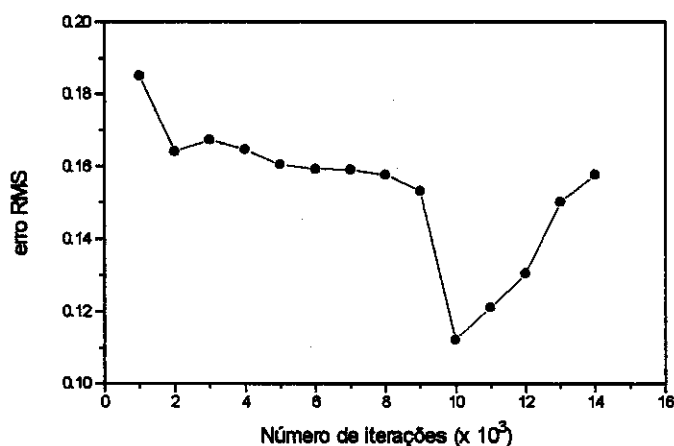


Figura 32 - Variação do erro RMS com o número de iterações, para uma rede neural 131-2-4.

O resultado gráfico da aplicação dos modelos fuzzy e da rede neural RBF, exibidos na Tabela 9, estão mostrados nas Figuras 32 e 33. Em ambos os casos algumas partes do gráfico onde ocorre o crescimento brusco do erro RMS foram omitidas, a fim de não prejudicar a constatação do menor erro em função da variação dos parâmetros. O comportamento do erro RMS não mostra “tendências” para a direção dos erros, para estes modelos.

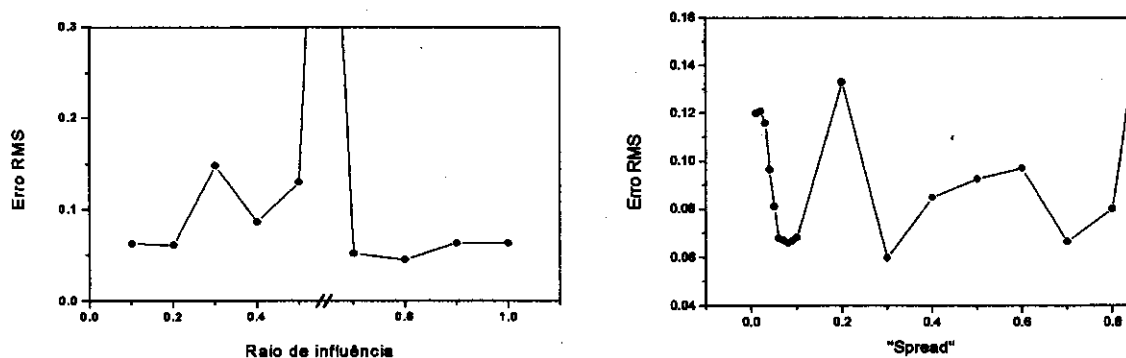


Figura 33 - Erros RMS calculados contra os parâmetros “raio de influência” para o modelo fuzzy e “spread” para a rede neural RBF. Somente “scores” foram como conjunto de entrada, utilizando todo o espectro.

Para as amostras de solos não foi feita a aplicação de extrapolação dos modelos, pois testes preliminares mostraram erros relativos muito altos no cálculo das concentrações.

4.3. Amostras de plantas (K e Ca)

Um outro estudo foi feito utilizando-se padrões de diversas 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 35. Assim como as amostras de solos, estas amostras possuem mais de 10 elementos químicos detectáveis (Figura 34) 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".

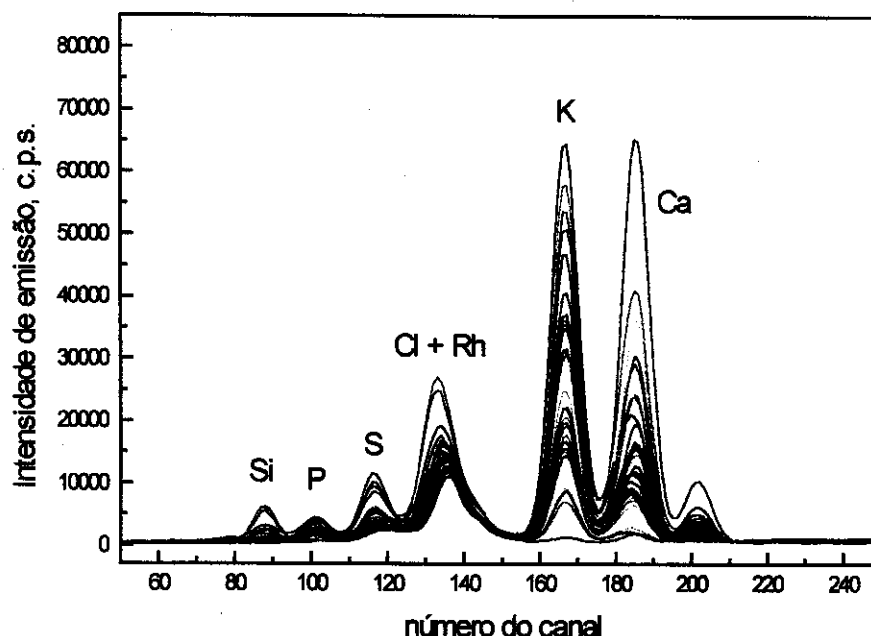


Figura 34 - Parte de espectros de EDXRF utilizados neste trabalho, mostrando as linhas espectrais dos diversos elementos presentes em amostras de plantas.

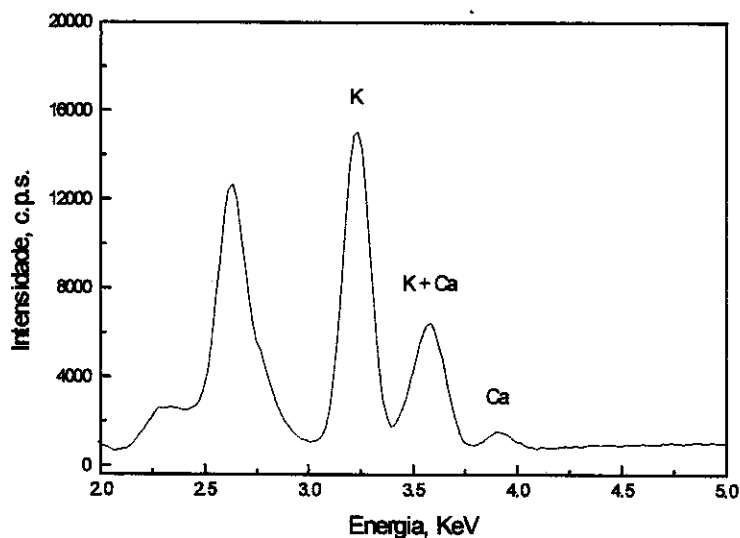


Figura 35 - Espectro de EDXRF típico para uma amostra de planta mostrando a sobreposição de linhas espectrais de Ca e K.

Na análise CV dos espectros das amostras foi determinado um número de 10 componentes principais representados por 10 "scores", por espectro. 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 11.

Tabela 11 - 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
PCR	0,153	0,0507
PLS	0,153	0,0507

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, embora a rede BP também tenha apresentado um bom resultado. Considerando-se as diferenças entre as concentrações reais e calculadas, elas fornecem erros relativos menores que 20%, como pode ser visto na Tabela 12. O teste F aplicado entre estes dois resultados mostra que não existe diferença entre eles ($F_{crit.} = 1,90$ para $n = 32$); apenas existe diferença em relação aos modelos PLS e PCR.

$$F = \frac{(RMS_{BP})^2}{(RMS_{RBF})^2} = \frac{(0,0247)^2}{(0,0222)^2} = 1,24 \quad F_2 = \frac{(RMS_{PLS})^2}{(RMS_{BP})^2} = \frac{(0,0507)^2}{(0,0222)^2} = 5,22$$

Tabela 12 - Erros relativos para a rede RBF com integrais (RMS: 0,0222).

	Ca			K		
	real	calc.	err.	real	calc.	err.
1	409	417	-2,0	576	525	8,7
2	484	387	19,9	1168	1190	-1,9
3	157	170	-8,3	240	263	-9,9
4	76,1	56,3	26,0	267	236	11,4
5	81,9	82,4	-0,60	635	622	2,0
6	135	124	7,5	354	404	-14,2
RMS	0,0195			0,0248		

calc.: calculado; err.: erro relativo

As arquiteturas das redes que apresentaram os resultados da Tabela 11 são as seguintes: 10-19-2 para RBF com "scores", 10-1-2 para LM com "scores" e 10-5-2 para BP com "scores", 3-10-2 para RBF com integrais, 3-4-2 para rede BP com integrais e 3-3-2 para LM com integrais.

As figuras apresentadas a seguir (Figuras 36 a 39) mostram, em uma forma gráfica, a origem dos resultados mostrados na Tabela 11. Novamente, as redes BP e LM possuem um comportamento em que o menor erro fica situado em um "vale" sendo que o erro tende a crescer nas vizinhanças do mesmo. Os modelos fuzzy e as redes RBF apresentam um tipo de comportamento

semelhante, mas a tendência do erro nas vizinhanças do “vale” não é muito bem definida.

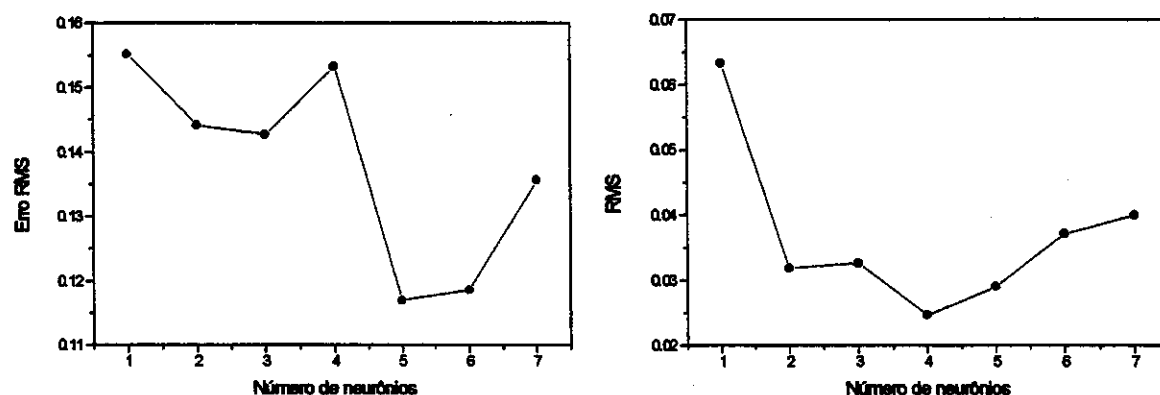


Figura 36 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP. O gráfico da esquerda representa o conjunto de entrada como integrais, e o da direita como “scores”.

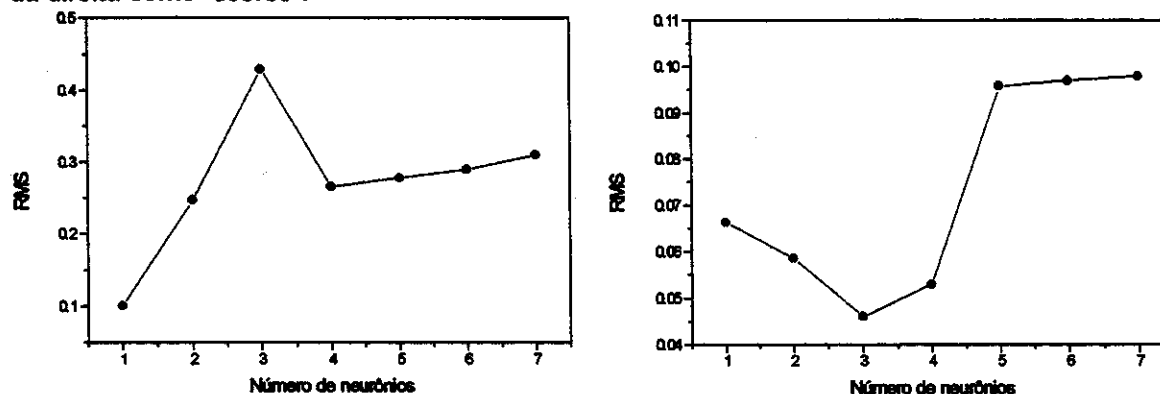


Figura 37 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural LM. O gráfico da esquerda representa o conjunto de entrada como “scores”, e o da direita como integrais.

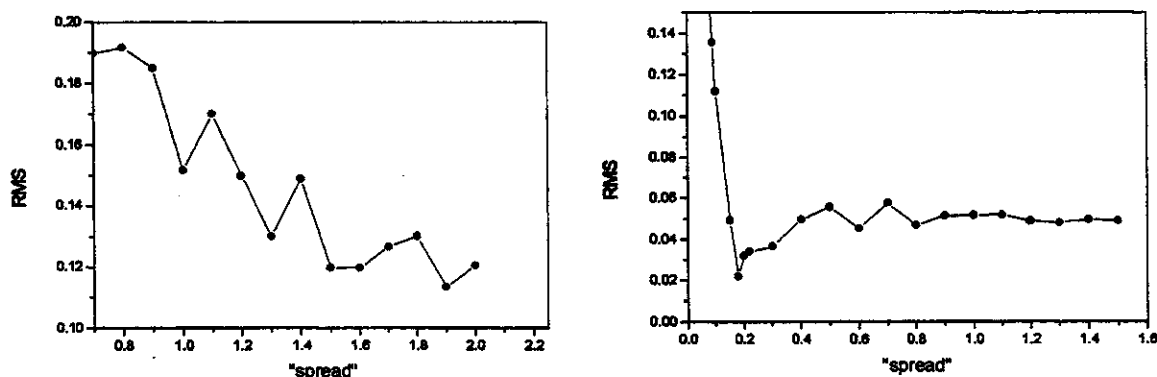


Figura 38 - Erros RMS calculados contra o parâmetro “spread”, para a rede neural RBF. O gráfico da esquerda representa o conjunto de entrada como “scores”, e o da direita como integrais. Em ambas as figuras algumas partes da escala dos parâmetros são omitidas, pois o erro RMS cresce além da escala do gráfico.

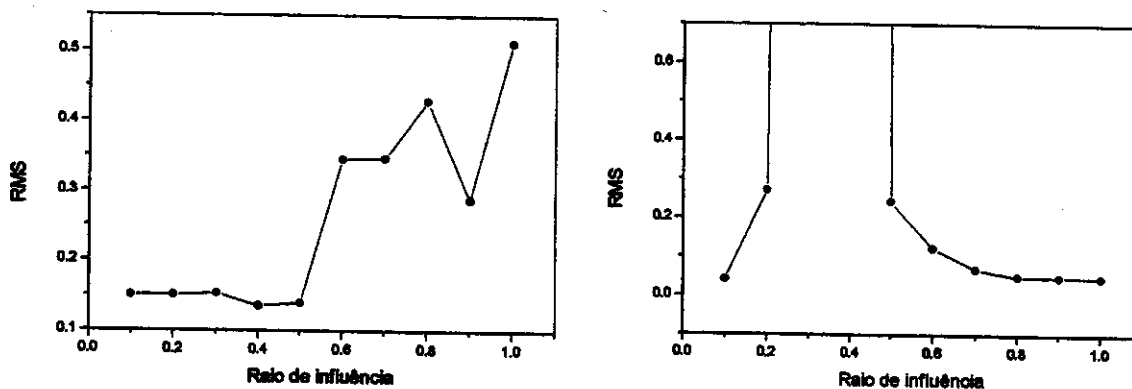


Figura 39 - Erros RMS calculados contra o parâmetro "raio de influência", para o modelo fuzzy. O gráfico da esquerda representa o conjunto de entrada como "scores", e o da direita como integrais. Na figura da direita não é mostrado a região entre 0.3 e 0.5, pois o erro RMS cresce além da escala do gráfico.

Os resultados de extrapolação dos modelos (erro RMS) para as amostras de plantas, considerando-se 4 espectros para validação e 28 para calibração, estão mostrado na Tabela 13. Apenas os "scores" foram utilizados como dados de entrada.

Tabela 13 - Erros RMS calculados para os modelos utilizados para amostras de plantas, para a extrapolação dos modelos .

rede neural BP	0,267
rede neural LM	0,336
rede neural RBF	0,367
modelo Fuzzy	0,336
PCR	0,619
PLS	0,619

Neste caso também as diferenças entre as concentrações produzem erros muito grandes, mesmo para a rede BP. Os modelos também não conseguiram fazer a extrapolação dos dados, como pode ser visto na ordem de grandeza dos erros da tabela acima.

As configurações das redes utilizadas foram os seguintes: rede RBF (10-18-2), rede LM (10-2-2) e rede BP (10-6-2). Este último modelo foi o que

apresentou um desempenho melhor em relação aos outros modelos, mas estatisticamente não há diferença entre a rede BP e a rede LM.

$$F = \frac{(\text{RMS}_{\text{LM}})^2}{(\text{RMS}_{\text{BP}})^2} = \frac{(0,336)^2}{(0,267)^2} = 1,58$$

4.4. Amostras de aços

Uma outra aplicação dos modelos quimiométricos feita neste trabalho foi a utilização de 16 amostras de aço a partir dos resultados obtidos da análise por ICP-OES (Espectrometria de emissão ótica com plasma indutivamente acoplado). Como o número de amostras neste caso foi pequeno, optou-se por utilizar dois conjuntos de validação distintos contendo 4 amostra cada, com as outras 12 amostras empregadas na calibração sendo utilizadas de maneira separada, sem a ocorrência de repetição das mesmas. Foram determinadas as concentrações dos elementos Cr, Mn e Fe, simultaneamente. Este tipo de amostra apresenta fluorescência secundária principalmente para o sistema Cr-Fe, bastante estudado pela literatura [12, 13], embora estas amostras apresentem outros 13 elementos detectáveis. Os 16 espectros utilizados estão mostrados sobrepostos na Figura 40:

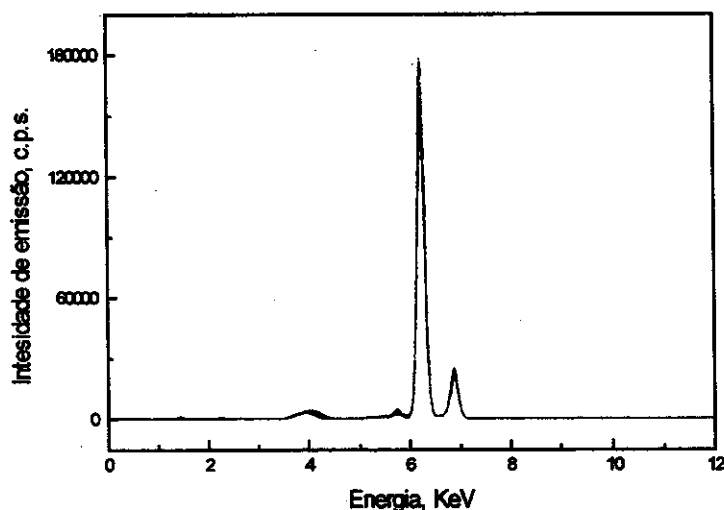


Figura 40 - Parte dos 16 espectros de EDXRF utilizados neste trabalho, mostrando as linhas espectrais do elemento Fe, sem reescalonamento.

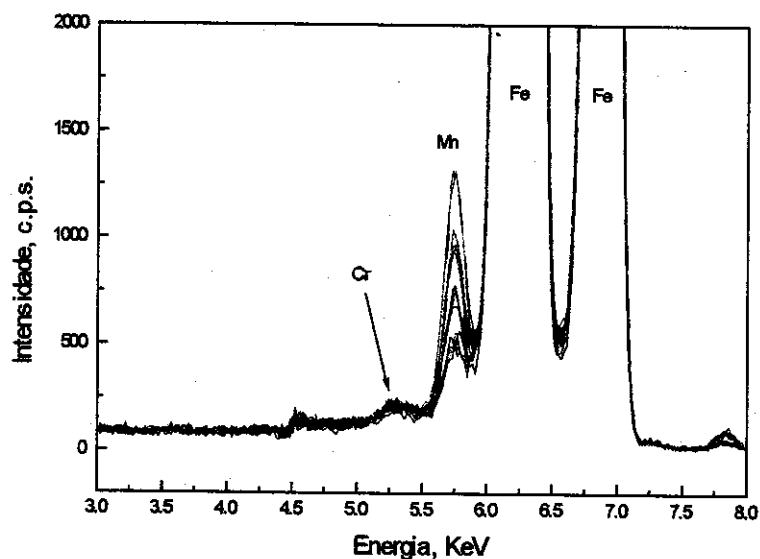


Figura 43 - Espectros de EDXRF para as amostras de aço mostrando as linhas dos elementos estudados, extrapolando a escala de intensidade de emissão, para se poder observar as outras linhas.

Na análise CV dos espectros das amostras foi determinado um número de 3 componentes principais representados por 3 "scores", por espectro. A análise mostrou uma variância capturada total de 100% para o bloco dos espectros e de 80,43% para o bloco das concentrações. Como nestas amostras, as linhas espectrais de Cr e Mn são muito menores que as de Fe, optou-se por utilizar apenas os "scores" como conjunto de dados de entrada. Os resultados das aplicações dos modelos para os dois grupos de validação distintos, estão mostrados na Tabela 14.

Tabela 14 - Erros RMS calculados para os modelos utilizados para amostras de aço, considerando-se apenas o número de "scores" igual a 3 para ambos os grupos.

	grupo 1	grupo 2
rede neural BP	0,166	0,323
rede neural LM	0,173	0,316
rede neural RBF	0,315	0,321
modelo Fuzzy	0,327	0,378
PCR	0,187	0,321
PLS	0,187	0,321

De maneira geral, as concentrações de Mn foram as que apresentaram os maiores erros, para ambos os grupos, de acordo com as diferenças entre as concentrações. Os erros relativos calculados sobre as concentrações, para os melhores modelos para ambos os grupos estão mostrados nas Tabelas 15 e 16.

Tabela 15 - Erros relativos para a rede BP (RMS: 0,166), para o GRUPO 1.

		amost.				RMS
		1	2	3	4	
Mn	real	0,570	0,608	0,255	0,232	0,200
	calc.	0,528	0,546	0,398	0,369	
	err.	7,37	10,2	-56,0	-59,0	
Cr	real	0,021	0,023	0,022	0,023	0,0201
	calc.	0,025	0,026	0,021	0,020	
	err.	-19,0	-13,0	4,55	13,0	
Fe	real	98,21	98,12	98,31	97,98	0,278
	calc.	97,92	97,90	98,05	98,08	
	err.	0,30	0,22	0,27	-0,10	

calc.: calculado; err.: erro relativo; amost.: amostra

Tabela 16 - Erros relativos para a rede LM (RMS: 0,316), para o GRUPO 2.

		amost.				RMS
		1	2	3	4	
Mn	real	0,531	0,451	0,277	0,230	0,251
	calc.	0,433	0,402	0,354	0,357	
	err.	18,4	10,8	-27,8	-55,2	
Cr	real	0,024	0,028	0,024	0,022	0,0807
	calc.	0,026	0,024	0,021	0,021	
	err.	-8,33	14,3	12,5	4,55	
Fe	real	97,62	97,69	97,63	97,74	0,615
	calc.	98,19	98,19	98,18	98,18	
	err.	-0,59	-0,51	-0,56	-0,45	

calc.: calculado; err.: erro relativo; amost.: amostra

As arquiteturas das redes que apresentaram os resultados da Tabela 14 para ambos os grupos são as seguintes: 3-1-3 para LM e BP, 3-11-3 para RBF e apenas 3-2-3 para a rede LM para o grupo 2.

De maneira geral (Tabela 14), o grupo 1 teve melhores resultados que o grupo 2 para todos os modelos. Para este tipo de amostra, as redes LM e BP foram as que apresentaram os melhores desempenhos, principalmente em relação ao grupo 1. Para o grupo 2 os resultados são quase idênticos. Aplicando-se o teste F, os resultados mostraram não serem estatisticamente diferentes, mas isto é devido ao conjunto de dados muito pequeno, com poucos graus de liberdade ($F_{crit} = 2,41$ para $n = 16$ com 95% de confiança).

$$F = \frac{(\text{RMS}_{\text{PLS}(G1)})^2}{(\text{RMS}_{\text{BP}(G1)})^2} = \frac{(0,187)^2}{(0,166)^2} = 1,27 \quad F = \frac{(\text{RMS}_{\text{PLS}(G2)})^2}{(\text{RMS}_{\text{LM}(G2)})^2} = \frac{(0,321)^2}{(0,316)^2} = 1,03$$

Os resultados gráficos que originaram a Tabela 14 estão mostrados nas Figuras 44 a 47 a seguir. As redes BP e LM possuem um comportamento em que o menor erro fica situado em um “vale” sendo que o erro tende a crescer nas vizinhanças do mesmo. Os modelos fuzzy e as redes RBF apresentam um tipo de comportamento semelhante, mas com variações dos erros mais suaves.

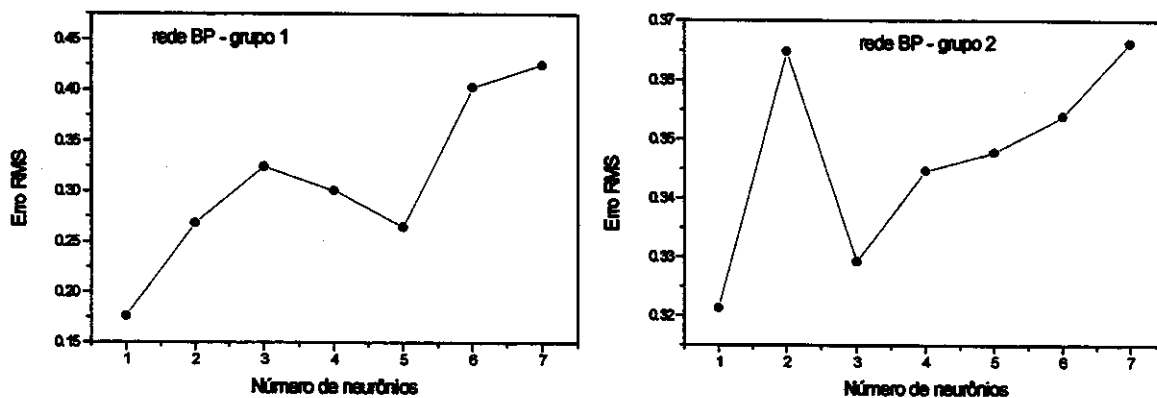


Figura 42 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP, para os dois grupos de validação.

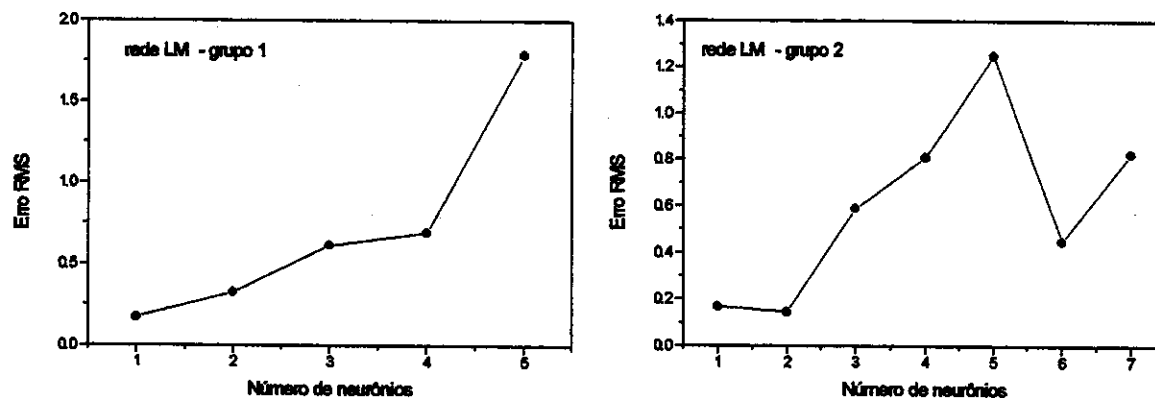


Figura 43 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural LM, para os dois grupos de validação.

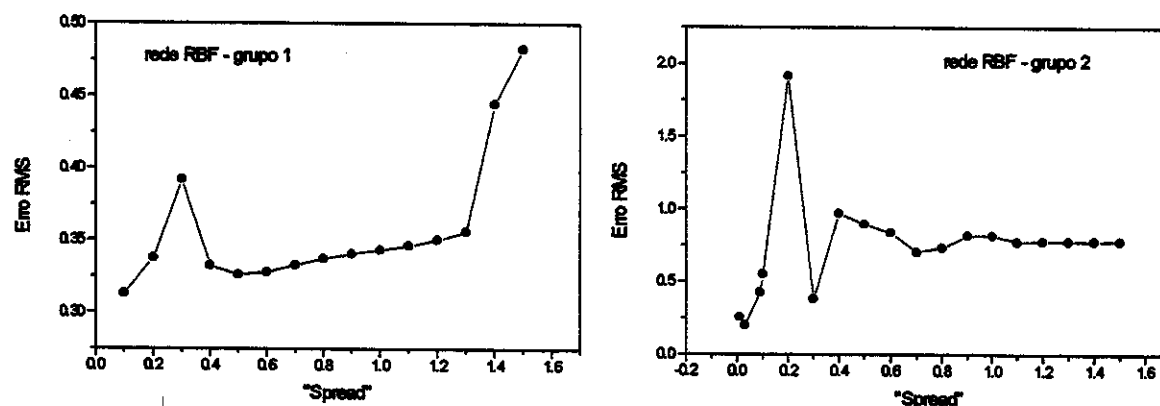


Figura 44 - Erros RMS calculados contra o parâmetro "spread", para a rede neural RBF, para os dois grupos de validação.

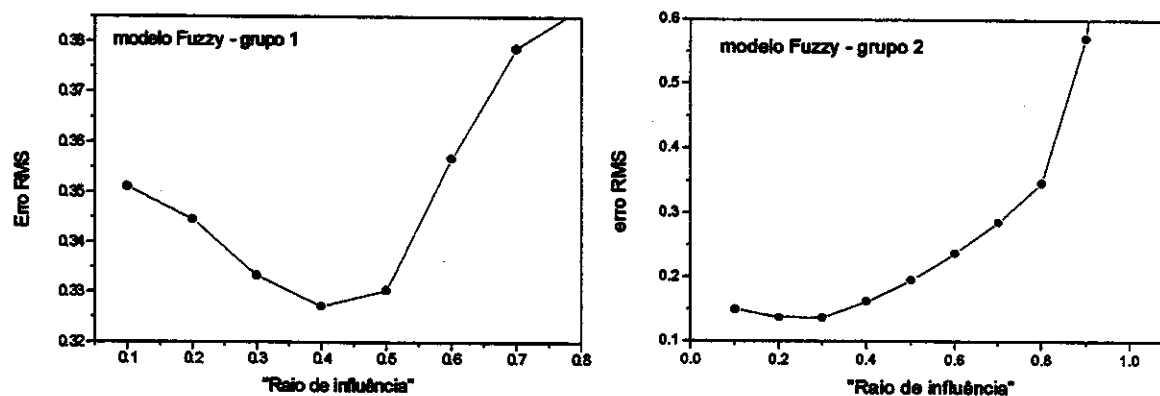


Figura 45 - Erros RMS calculados contra o parâmetro "raio de influência", para o modelo fuzzy, para os dois grupos de validação. Não é mostrado toda a variação do parâmetro "raio de influência", pois o erro RMS cresce além da escala do gráfico.

4.4.1. Reescalonamento

Um outro experimento realizado com as amostras de aço foi o de verificar a variação do erro RMS com a utilização de espectros reescalonados, ou seja, espectros onde foram realçadas as linhas espectrais dos elementos Cr e Mn em relação às linhas de Fe, mantendo as proporções destes dois elementos, mas sem provocar distorções. Este procedimento visa aumentar, de forma artificial, as linhas espectrais dos elementos menos intensos de modo a tentar facilitar a sua identificação matemática pelos modelos. O resultado espectral deste reescalonamento está mostrado na Figura 46.

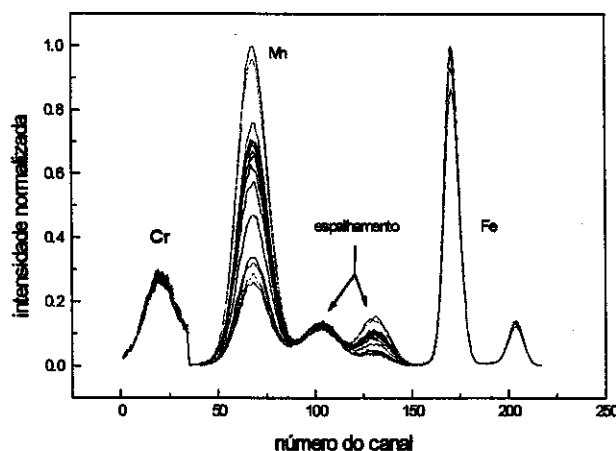


Figura 46- Espectros reescalonados das amostras de aço. Os picos de Fe e Cr foram obtidos por EDXRF sendo que as linhas de Cr foram aumentadas. As linhas de Mn foram obtidas através de fluorescência de raios-X por radiação síncrotron, a qual apresenta uma região de espalhamento entre os canais 100 e 150.

Estes novos 16 espectros foram obtidos utilizando-se partes de espectros de EDXRF aumentando-se a intensidade da linha de Cr (normalizada entre 0 e 0,3) e acrescentando partes de espectros obtidos por FRX de radiação síncrotron das mesmas amostras, com ênfase das linhas de Mn (Figura 47). A linha de Cr foi aumentada entre 0 e 0,3, pois ela é quase uma linha de ruído como pode-se observar na Figura 43. As linhas de Fe provêm de espectros de EDXRF

Para construção dos conjuntos de entrada, foram obtidos 3 componentes

principais representados por 3 “scores” , por espectro, através da aplicação de PCA a estes novos espectros reescalados. A variância total capturada foi de 100% para o bloco dos espectros e de 81,48% para o bloco das concentrações. Foram também utilizadas, como dados de entrada, as integrais das linhas de Cr (canais de 1 a 36), Mn (canais de 40 a 90) e Fe (canais de 156 a 186), em relação a Figura 46.

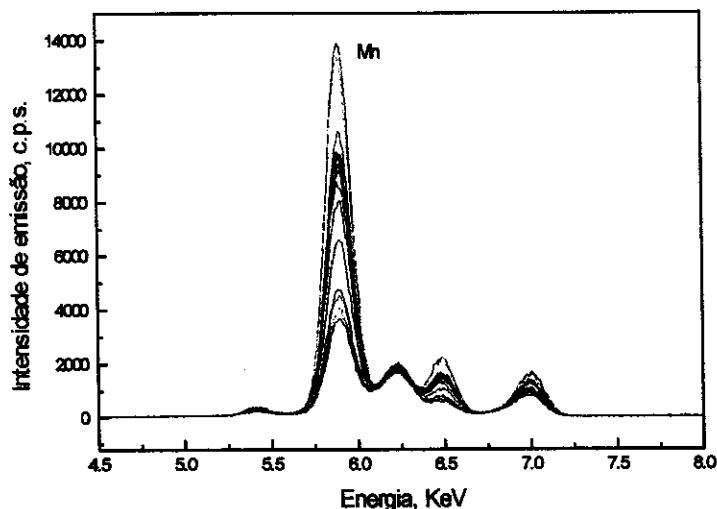


Figura 47 - Parte de espectros de fluorescência de raios-X por radiação síncrotron utilizados neste trabalho, mostrando as linhas espectrais do Mn.

Foram utilizados os mesmos dois grupos de validação de 4 espectros utilizados anteriormente. Os resultados das aplicações dos modelos estão mostrados na Tabela 17.

Tabela 17 - Erros RMS para os espectros reescalados, em relação ao conjunto de entrada (amostras de aço), considerando-se o número de “scores” igual a 3 e o número de integrais igual a 3.

	grupo 1		grupo 2	
	scores	integrais	scores	integrais
rede neural BP	0,213	0,185	0,275	0,223
rede neural LM	0,208	0,185	0,311	0,217
rede neural RBF	0,313	0,293	0,284	0,288
modelo Fuzzy	0,281	0,255	0,325	0,280
PCR	0,262	0,235	0,304	0,280
PLS	0,262	0,235	0,304	0,280

Os resultados da Tabela 17 não mostram para ambos os grupos, uma diferença significativa (teste F com 95%, para N = 16) entre os modelos PLS e BP, e também entre PLS e LM. Também não há uma diferença entre a rede BP e o modelo PLS utilizando integrais como dados de entrada.

$$F_1 = \frac{(\text{RMS}_{\text{PLS}(1)})^2}{(\text{RMS}_{\text{LM}(1)})^2} = \frac{(0,262)^2}{(0,208)^2} = 1,59 \quad F_2 = \frac{(\text{RMS}_{\text{PLS}(1)})^2}{(\text{RMS}_{\text{BP}(1)})^2} = \frac{(0,235)^2}{(0,185)^2} = 1,61$$

$$F_3 = \frac{(\text{RMS}_{\text{PLS}(2)})^2}{(\text{RMS}_{\text{BP}(2)})^2} = \frac{(0,304)^2}{(0,275)^2} = 1,22 \quad F_4 = \frac{(\text{RMS}_{\text{PLS}(2)})^2}{(\text{RMS}_{\text{LM}(2)})^2} = \frac{(0,280)^2}{(0,217)^2} = 1,66$$

Os resultados não mostraram de maneira geral uma melhora do erro RMS para os espectros reescalados, em relação ao que está apresentado na Tabela 14, sendo isto confirmado estatisticamente com o teste F. Isto ocorreu principalmente devido ao conjunto de amostras ser muito pequeno ($F_{\text{crit}} = 2,41$ para $n = 16$ com 95% de confiança). Abaixo mostramos a comparação entre os melhores resultados das Tabelas 14 e 17 para cada grupo:

$$F_1 = \frac{(\text{RMS}_{\text{BP}(1)})^2}{(\text{RMS}_{\text{LM}(1)})^2} = \frac{(0,187)^2}{(0,185)^2} = 1,02 \quad F_2 = \frac{(\text{RMS}_{\text{LM}(2)})^2}{(\text{RMS}_{\text{LM}(2)})^2} = \frac{(0,316)^2}{(0,217)^2} = 2,12$$

A principal informação que pode ser vista na Tabela 17 é que os erros obtidos com as integrais são menores do que os obtidos com os "scores" para quase todos os modelos.

Os erros relativos calculados sobre as concentrações, para os melhores modelos (rede BP com integrais) para ambos os grupos estão mostrados nas Tabelas 18 e 19. Neste caso também não parece haver alguma melhora em relação ao erros que estão mostrados nas Tabelas 15 e 16.

Tabela 18 - Erros relativos para a rede BP com integrais (RMS: 0,162), para o GRUPO 1.

		amost.				RMS
		1	2	3	4	
Mn	real	0,570	0,608	0,255	0,232	0,182
	calc.	0,511	0,529	0,373	0,365	
	err.	10,3	12,9	-46,2	-57,3	
Cr	real	0,021	0,023	0,022	0,023	0,0445
	calc.	0,025	0,025	0,021	0,021	
	err.	-19,0	-8,70	4,55	8,70	
Fe	real	98,21	98,12	98,31	97,98	0,260
	calc.	97,93	97,92	98,07	98,08	
	err.	0,29	0,21	0,25	-0,10	

calc.: calculado; err.: erro relativo; amost.: amostra

Tabela 19 - Erros relativos para a rede BP com integrais (RMS: 0,223), para o GRUPO 2.

		amost.				RMS
		1	2	3	4	
Mn	real	0,531	0,451	0,277	0,230	0,211
	calc.	0,727	0,590	0,570	0,324	
	err.	-36,9	-30,8	-105	-40,8	
Cr	real	0,024	0,028	0,024	0,022	0,0295
	calc.	0,025	0,024	0,021	0,020	
	err.	-4,17	14,3	12,5	9,09	
Fe	real	97,62	97,69	97,63	97,74	0,428
	calc.	97,80	98,03	97,99	98,23	
	err.	-0,19	-0,34	-0,36	-0,50	

calc.: calculado; err.: erro relativo; amost.: amostra

Para estas amostras de aço não foram feitos os cálculos de extrapolação dos modelos pois a variação de concentração dos 3 metais para as 16 amostras

é muito pequena.

Os resultados que originaram a Tabela 17 podem ser vistos nas Figuras 48 a 55, mostradas a seguir. As redes BP e LM possuem um comportamento em que o menor erro fica situado em um “vale” sendo que o erro tende a crescer bruscamente nas vizinhanças do mesmo. Os modelos fuzzy e as redes RBF apresentam um tipo de comportamento semelhante, mas a tendência do erro nas vizinhanças do “vale” é mais suave.

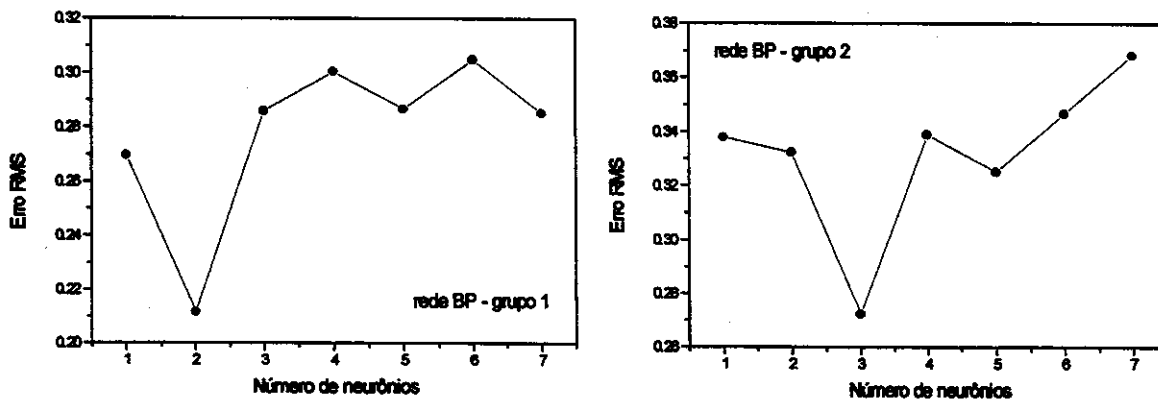


Figura 48 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP, para os dois grupos de validação usando “scores” como dados de entrada.

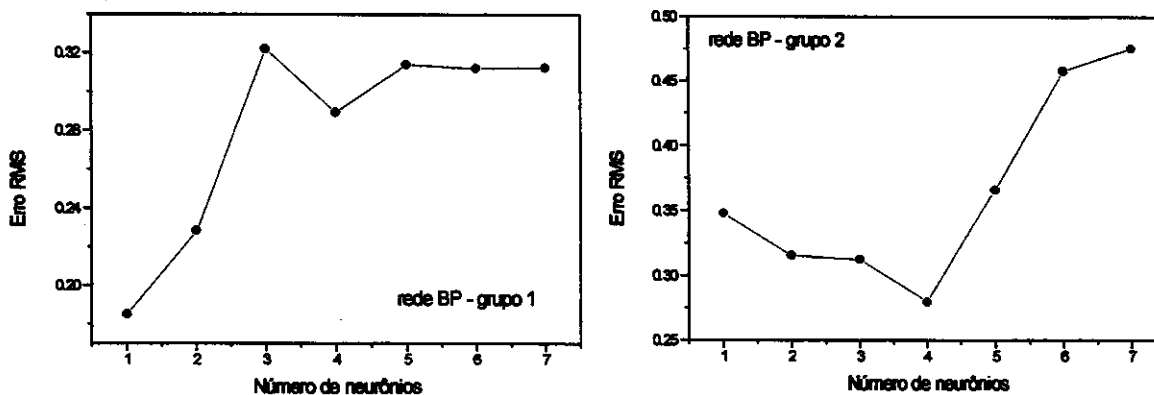


Figura 49 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP, para os dois grupos de validação usando integrais como dados de entrada.

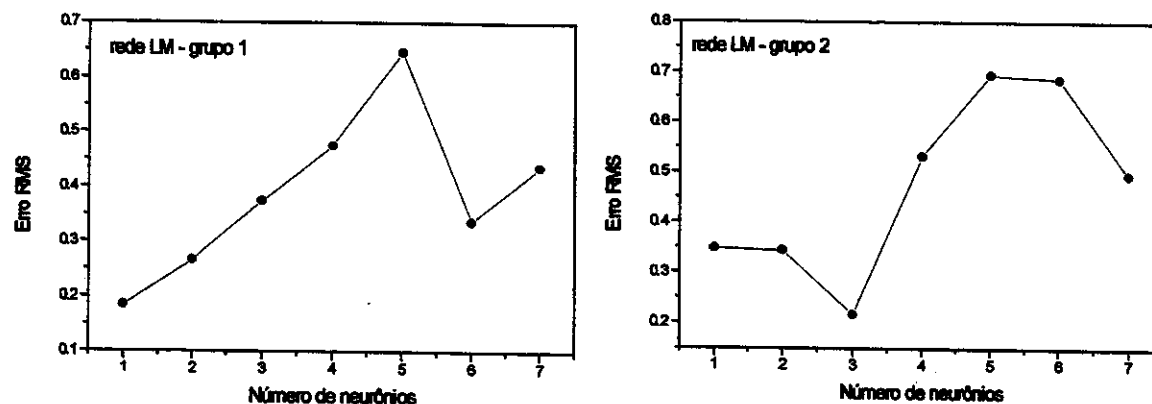


Figura 50 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural LM, para os dois grupos de validação usando integrais como dados de entrada.

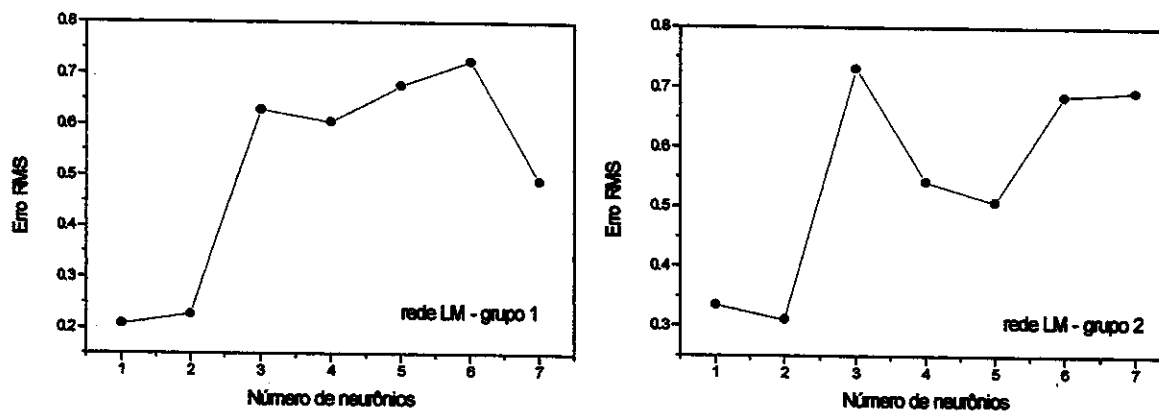


Figura 51 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural LM, para os dois grupos de validação usando "scores" como dados de entrada.

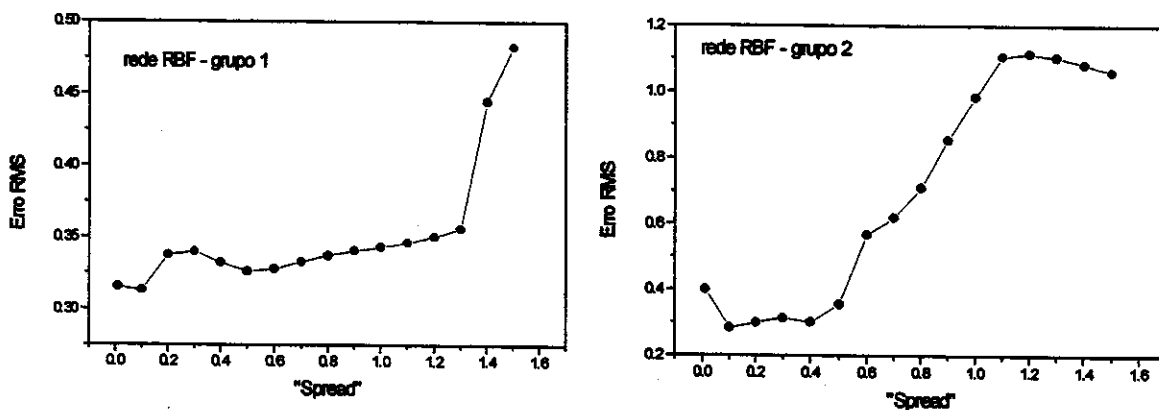


Figura 52 - Erros RMS calculados contra o parâmetro "spread", para a rede neural RBF, para os dois grupos de validação, utilizando os "scores" como dados de entrada.

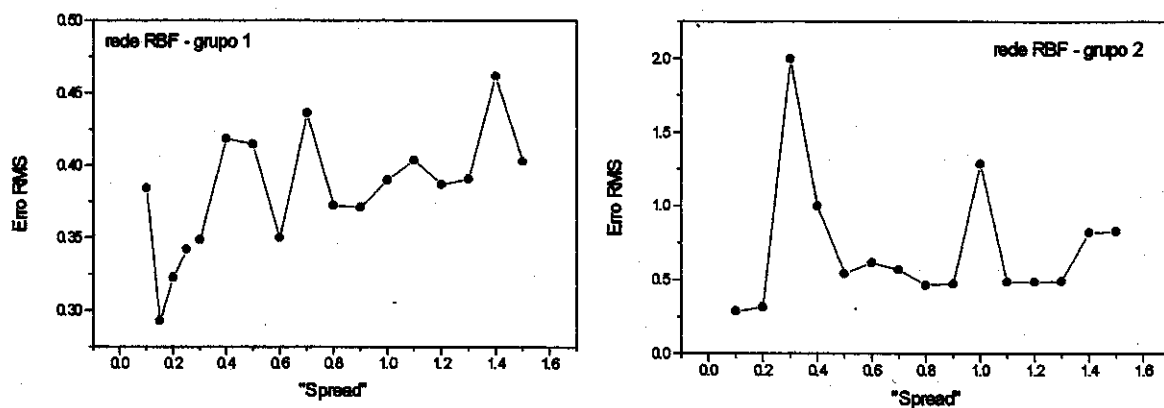


Figura 53 - Erros RMS calculados contra o parâmetro "spread", para a rede neural RBF, para os dois grupos de validação, utilizando integrais como dados de entrada.

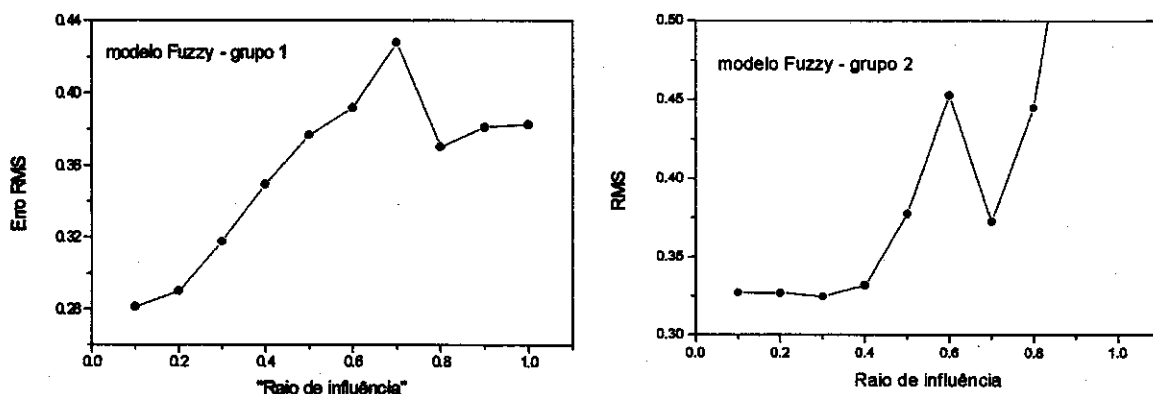


Figura 54 - Erros RMS calculados contra o parâmetro "raio de influência", para o modelo fuzzy, para os dois grupos de validação, utilizando os "scores" como dados de entrada.

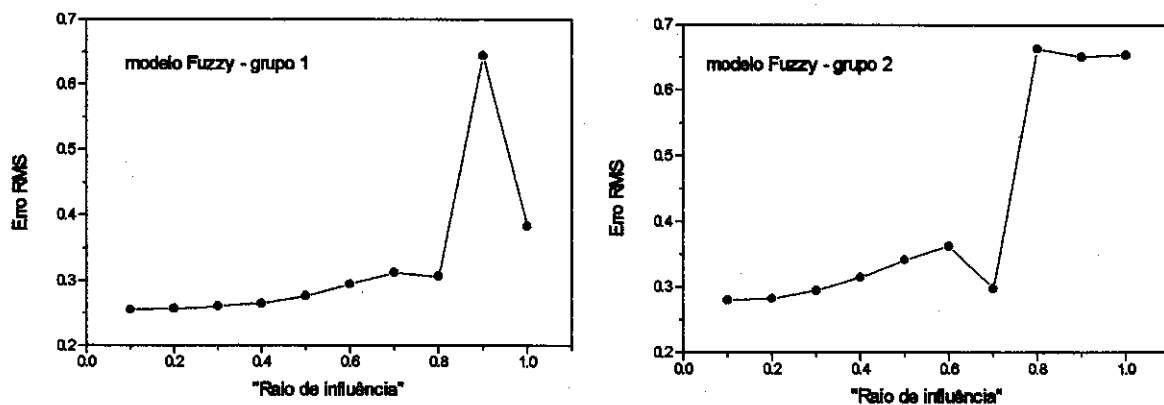


Figura 55 - Erros RMS calculados contra o parâmetro "raio de influência", para o modelo fuzzy, para os dois grupos de validação, utilizando integrais como dados de entrada.

4.5. Amostras de terras raras

Neste trabalho foi feita a determinação simultânea de Nd, Pr e Sm, através de misturas de óxidos destes metais em matriz de sílica. Os dados de entrada para os todos modelos utilizados foram apenas os “scores” e os dados de saída foram as concentrações de Nd, Pr e Sm. Dos 64 espectros das misturas originais, foram selecionados 45 espectros para calibração, 14 para previsão e foram retirados 4 “outliers”. Devido a grande sobreposição das linhas espectrais destes elementos, não foi possível calcular as integrais, por ser muito difícil a definição dos limites de integração. Isto fica claro quando se observa os espectros puros dos três elementos na Figura 56. A Figura 57 mostra alguns espectros sobrepostos com a região de sobreposição das linhas dos elementos de interesse.

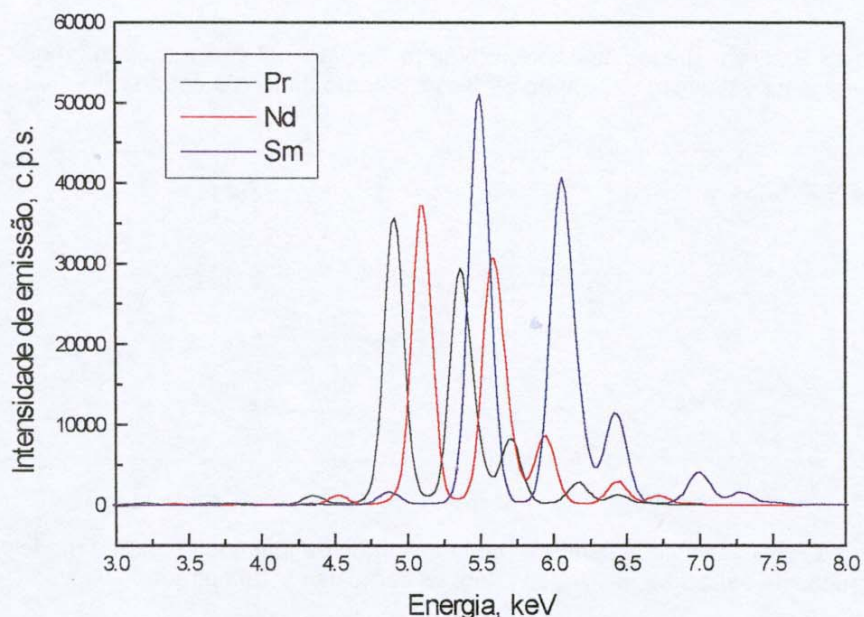


Figura 56 - Parte de espectros sobrepostos de EDXRF de amostras puras de Pr, Nd e Sm, mostrando as linhas espectrais características de cada elemento.

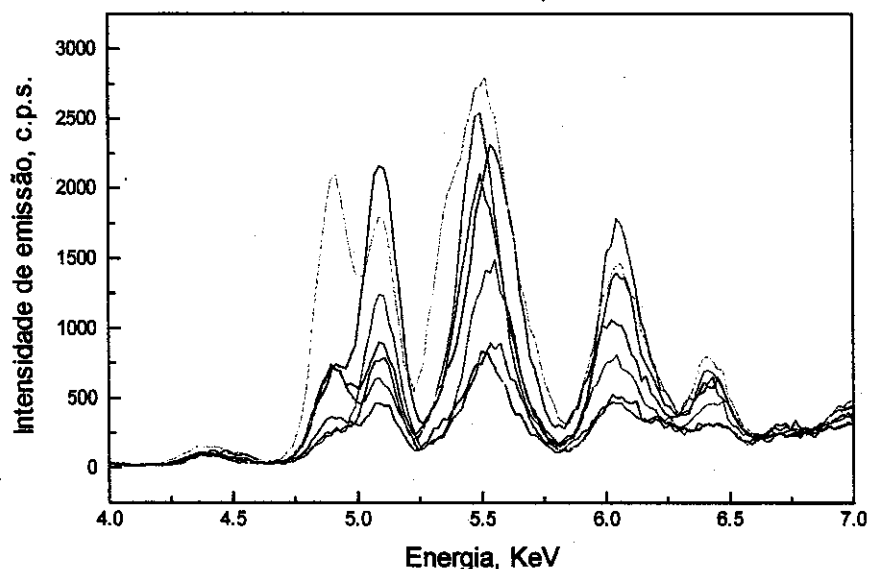


Figura 57 - Parte de espectros sobrepostos de EDXRF utilizados, mostrando a região onde aparecem as linhas espectrais do Pr, Nd e Sm, mostrando uma intensa sobreposição.

Na análise CV dos espectros das amostras foi determinado um número de 5 componentes principais, representados por 5 "scores" por espectro. A análise mostrou uma variância capturada total de 99,72% para o bloco dos espectros e de 95,68% para o bloco das concentrações. Os resultados da aplicação dos modelos estão mostrados na Tabela 20.

Tabela 20 - Erros RMS calculados para os modelos utilizados para amostras de terras raras, considerando-se apenas o número de "scores" igual a 3.

	scores
rede neural BP	0,147
rede neural LM	0,157
rede neural RBF	0,292
modelo Fuzzy	0,228
PCR	0,155
PLS	0,155

Para este tipo de amostra, os erros encontrados pelos modelos foram altos, em função da grande dificuldade de separar as linhas de cada elemento. Eles apresentam valores muito semelhantes, com uma vantagem para

a rede BP. Para este modelo, a diferença entre as concentrações reais e calculadas está em torno de 30% (Tabela 21).

Tabela 21 - Erros relativos para a rede neural BP (RMS: 0,147).

	Pr			Nd			Sm		
	real	calc.	err.	real	calc.	err.	real	calc.	err.
1	5,61	9,29	-65,5	10,18	14,39	-41,4	5,31	7,47	-40,6
2	5,53	10,44	-88,8	10,66	14,79	-38,8	15,26	19,04	-24,8
3	5,49	6,69	-21,8	16,17	21,12	-30,6	10,76	13,25	-23,1
4	5,56	5,99	-7,8	20,59	20,64	-0,30	16,03	17,73	-10,6
5	10,09	7,07	29,9	6,17	7,19	-16,5	20,43	18,06	11,6
6	10,34	11,64	-12,6	14,98	20,38	-36,1	5,11	5,82	-13,8
7	10,28	8,89	13,6	20,64	21,6	-4,7	15,11	14,72	2,6
8	15,36	13,93	9,3	5,06	4,6	9,0	15,2	15,8	-3,9
9	15,36	16,18	-5,3	10,38	12,74	-22,8	5,23	5,85	-11,9
10	15,39	14,35	6,7	10,63	11,41	-7,3	20,20	18,55	8,2
11	15,16	15,62	-3,0	20,33	21,79	-7,2	5,89	6,27	-6,5
12	20,40	19,73	3,3	4,94	2,59	47,6	10,03	8,89	11,4
13	20,20	15,45	23,5	10,29	6,93	32,7	15,01	13,07	12,9
14	20,28	17,9	11,7	20,26	17,21	15,1	20,14	18,25	9,4
RMS	0,149			0,182			0,111		

calc.: calculado; err.: erro relativo

Os erros relativos encontrados não foram muito bons mas um pouco menores que os usualmente encontrados para amostras de terras raras [31]. Os elementos terras raras possuem alguns problemas para a sua determinação quantitativa, devido a ausência de reagentes seletivos e por possuírem espectros de absorção e emissão muito complexos, além de interferências espectrais [31].

As arquiteturas das redes que apresentaram os resultados da Tabela 20 são as seguintes: 3-2-3 para a rede BP, 3-6-3 para a rede LM e 3-29-3 para a rede RBF.

Um teste F aplicado aos valores obtidos pela rede LM em relação a rede BP e aos modelos PLS e PCR, indica haver diferença significativa ao nível de 95% de confiança ($F_{crit} = 1,0$ para $n = 64$). A rede LM modelou melhor o conjunto de validação utilizado em relação aos outros modelos.

$$F_1 = \frac{(RMS_{BP})^2}{(RMS_{LM})^2} = \frac{(0,230)^2}{(0,200)^2} = 1,32$$

$$F_2 = \frac{(RMS_{PLS})^2}{(RMS_{LM})^2} = \frac{(0,223)^2}{(0,200)^2} = 1,24$$

As figuras apresentadas a seguir (Figuras 68 e 59) mostram, em forma gráfica, a origem dos resultados mostrados na Tabela 20.

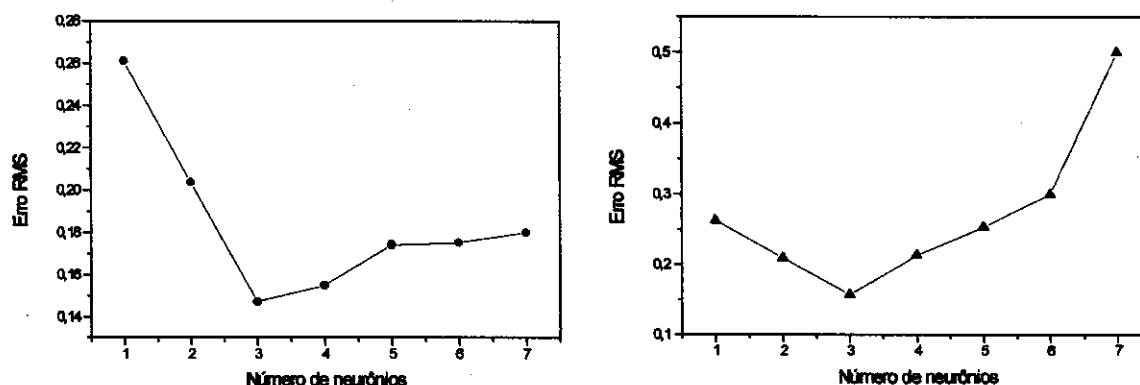


Figura 58 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP (a esquerda) e para a rede neural LM (a direita).

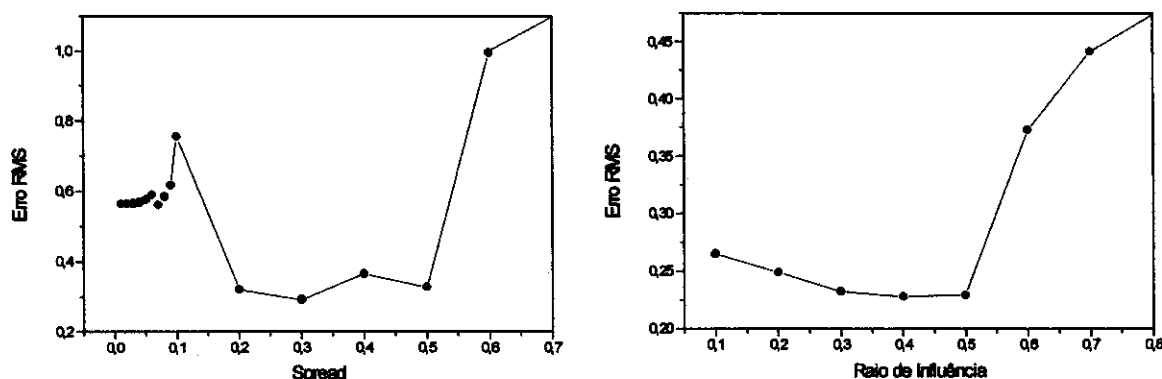


Figura 59 - Erros RMS calculados contra o parâmetro "spread", para a rede neural RBF (a esquerda). A direita, erro RMS calculado contra o parâmetro "raio de influência", para o modelo Fuzzy. Nesta figura, parte da escala entre 0,8 e 1,0 é omitida, pois o erro RMS cresce além da escala do gráfico.

Para estas amostras sintéticas de terras raras não foram feitos os cálculos de extrapolação dos modelos, pois a variação de concentração dos 3 elementos para as 64 amostras não é muito grande.

4.6. Amostras de plantas (Si, P, S, K e Ca)

4.6.1. Reescalamento

Retomamos um outro estudo utilizando-se os mesmos padrões de plantas do item 4.3 para verificar a validade da proposta de reescalamento dos espectros, e assim melhorar a determinação simultânea dos elementos. Neste novo trabalho foi feita a determinação das concentrações dos elementos Si, P, S, Ca e K.

As linhas espectrais dos elementos Si, P e S são bem menores do que as de Ca e K, como pode ser visto na Figura 60.

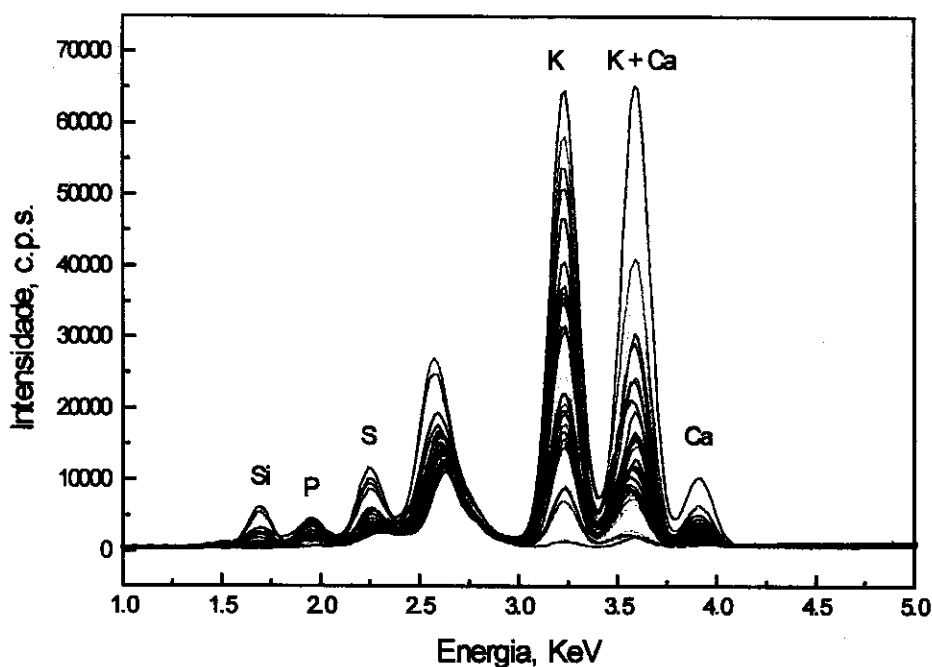


Figura 60 - Região de interesse dos 32 espectros de EDXRF para amostras de plantas utilizadas, mostrando as linhas espectrais dos cinco elementos estudados.

Os modelos foram aplicados para os espectros da Figura 60 utilizando-se apenas os "scores" como dados de entrada. Os resultados estão mostrados na Tabela 22. Na análise CV destes espectros foi determinado um número de 5

componentes principais representado por 5 "scores", por espectro, com uma variância capturada total de 99,90% para o bloco dos espectros e de 74,85% para o bloco das concentrações. Foram utilizados 20 espectros para calibração, 7 para validação e foram retirados 5 "outliers".

Tabela 22 - Erros RMS calculados para os modelos utilizados para amostras de plantas, considerando-se o número de "scores" igual a 5 (espectros comuns). O conjunto de validação utilizado aqui é diferente do outro usado apenas para o Ca e K.

	scores
rede neural BP	0,231
rede neural LM	0,257
rede neural RBF	0,290
modelo Fuzzy	0,327
PCR	0,275
PLS	0,275

Os erros encontrados são altos para todos os modelos, principalmente para o elemento Si. As configurações das redes neurais utilizadas na Tabela 22 são: rede BP (5-3-5), rede LM (5-3-5) e rede RBF (5-17-5). A rede BP apresentou o melhor resultado embora segundo um Teste F, não existe diferença em relação aos modelos PLS e PCR ($F_{crit.} = 1,90$ para $n = 32$).

$$F_1 = \frac{(RMS_{PLS})^2}{(RMS_{BP})^2} = \frac{(0,275)^2}{(0,231)^2} = 1,42$$

Os erros relativos calculados sobre as concentrações, para o melhor modelo (rede BP) estão mostrados na Tabela 23.

Tabela 23 - Erros relativos para os espectros comuns, para a rede BP (RMS: 0,231).

		amost.							RMS
		1	2	3	4	5	6	7	
Si	real	877	5233	1188	9475	3513	37509	1207	0,220
	calc.	1948	2689	5868	4071	1896	19129	9829	
	err.	-122	48,6	-393	57,0	46,0	49,0	-714	
Ca	real	163	484	567	160	81,9	2092	467	0,261
	calc.	249	424	315	24,1	149,1	689	475	
	err.	-53,1	12,2	44,3	84,9	-82,0	67,0	-1,71	
K	real	284	1168	672	290	635	1101	397	0,213
	calc.	532	990	653	232	722	1346	857	
	err.	-87,5	15,2	2,75	19,7	-13,7	-22,3	-116	
P	real	81,4	224	75,2	55,2	85,8	195	90,5	0,184
	calc.	97,8	141	103	23,6	89,2	156,8	126	
	err.	-20,1	36,8	-37,5	57,2	-3,96	19,5	-40,1	
S	real	49,9	130,7	149,1	76,6	54,9	451	125,8	0,278
	calc.	106,2	125	108,4	75,2	83	135	125,4	
	err.	-112	4,36	27,3	1,83	-51,1	69,8	0,32	

calc.: calculado; err.: erro relativo; amost.: amostra

Os resultados gráficos da Tabela 22 estão mostrados nas Figuras 61 e 62. As redes BP e LM possuem um comportamento em que o menor erro fica situado em um “vale”, sendo que o erro tende a crescer bruscamente nas vizinhanças do mesmo. Os modelos fuzzy e as redes RBF apresentam um tipo de comportamento semelhante, mas a tendência do erro nas vizinhanças do “vale” é mais suave.

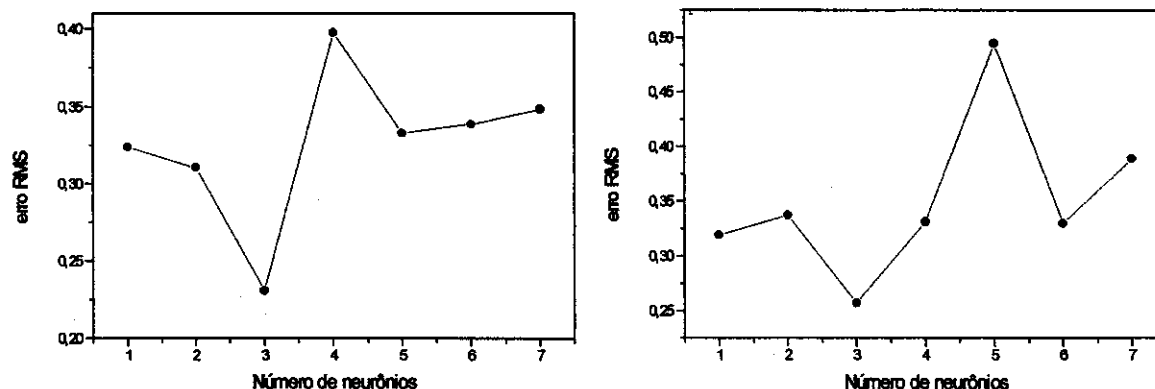


Figura 61 - Erros RMS calculados contra o número de neurônios na camada intermediária, para a rede neural BP (a esquerda) e para a rede neural LM (a direita).

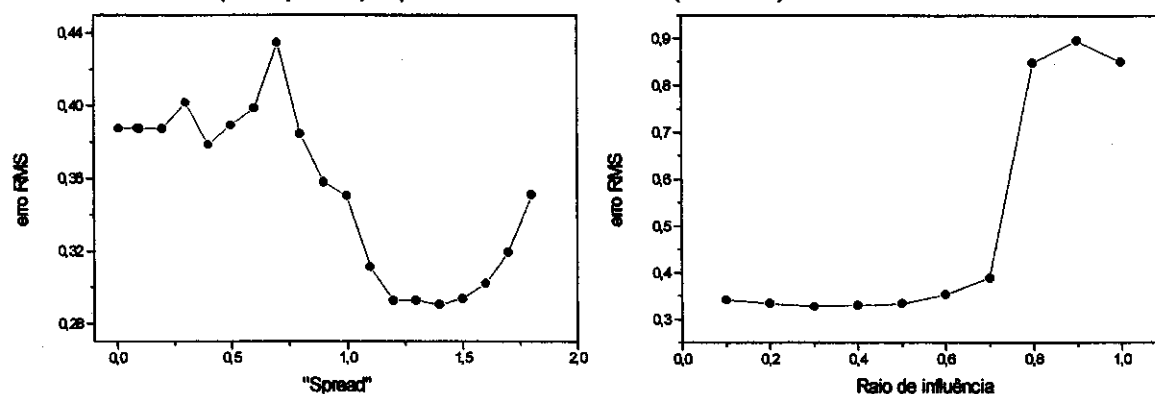


Figura 62 - Erros RMS calculados contra o parâmetro "spread", para a rede neural RBF (a esquerda). A direita, erro RMS calculado contra o parâmetro "raio de influência", para o modelo Fuzzy.

Para verificar se existe alguma melhora para estes resultados, foi feito o reescalonamento individual das linhas de Si, P e S entre os valores de 0 e 1. Então essas linhas foram adicionadas às linhas de Ca e K, sem alterar quaisquer outras características delas. Foi apenas retirada a linha correspondente ao Rh (entre 2,4 e 2,9 KeV). O resultado deste processo pode ser observado na Figura 63.

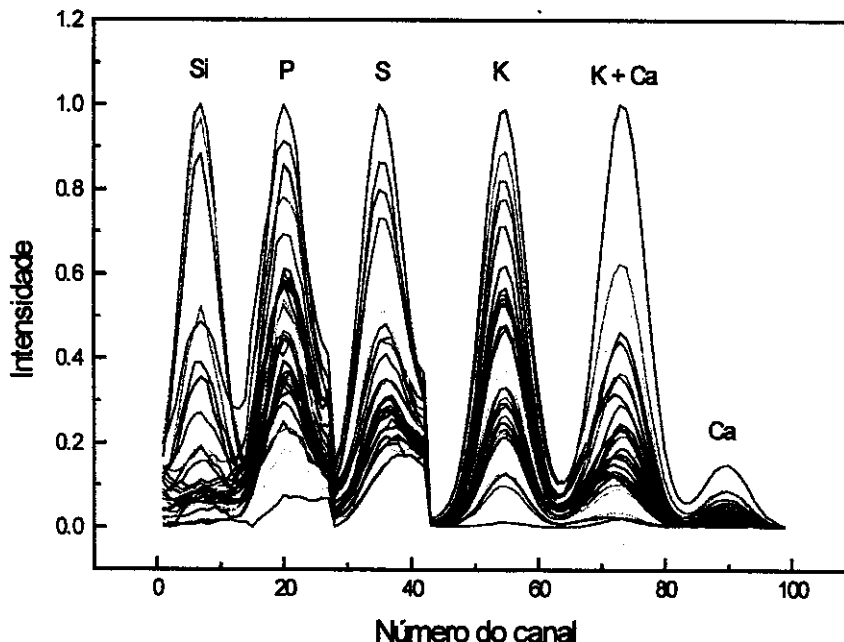


Figura 63- Espectros reescalados das amostras de plantas. As linhas de Si, P e S foram aumentadas e adicionadas às linhas de Ca e K.

Para esta nova situação, foram selecionados 20 espectros para calibração, 7 para previsão e foram retirados 5 "outliers".

Na análise CV destes espectros, foi determinado um número de 7 componentes principais representado por 7 "scores", com uma variância capturada total de 99,93% para o bloco dos espectros e de 89,78% para o bloco das concentrações. A determinação das concentrações foi feita também através das integrais calculadas sob as linhas dos elementos (6 linhas). Os resultados obtidos estão mostrados na Tabela 24.

Tabela 24 - Erros RMS calculados para os modelos utilizados para amostras de plantas (espectros reescalados), considerando-se o número de "scores" igual a 7 e o número de integrais igual a 6.

	scores	integrais
rede neural BP	0,195	0,184
rede neural LM	0,288	0,239
rede neural RBF	0,296	0,285
modelo Fuzzy	0,255	0,226
PCR	0,225	0,188
PLS	0,225	0,188

De acordo com a Tabela 24, a utilização de integrais proporcionou erros um pouco menores do que com o uso dos "scores", com pequena vantagem da rede BP sobre os outros modelos. As arquiteturas das redes neurais utilizadas que produziram os menores erros neste caso foram: rede BP com "scores" (7-5-5), rede LM com "scores" (7-5-5), rede RBF com "scores" (7-17-5), rede BP com integrais (6-4-5), rede LM com integrais (6-3-5) e rede RBF com integrais (6-15-5).

A aplicação do teste F ($F_{crit} = 1,90$, para $n = 32$) mostrou que não há diferença entre os melhores resultados para integrais (rede BP e PLS) e para "scores" (rede LM e PLS):

$$F_1 = \frac{(RMS_{PLS})^2}{(RMS_{LM})^2} = \frac{(0,225)^2}{(0,195)^2} = 1,33 \qquad F_2 = \frac{(RMS_{PLS})^2}{(RMS_{BP})^2} = \frac{(0,188)^2}{(0,184)^2} = 1,04$$

Os erros relativos calculados para as concentrações, para a rede BP com integrais, estão mostrados na Tabela 25.

Tabela 25 - Erros relativos para a rede BP, utilizando integrais, para o processo de reescalonamento (RMS: 0,184).

		amost.							RMS
		1	2	3	4	5	6	7	
Si	real	877	5233	1188	9475	3513	37509	1207	0,212
	calc.	422	4491	1052,2	6985	679,3	16665	662,5	
	err.	51,8	14,1	11,4	26,2	80,6	55,5	45,1	
Ca	real	163	484	567	160	81,9	2092	467	0,286
	calc.	120	378	464	148	9,5	521	369	
	err.	25,8	21,8	18,1	7,13	88,4	75,1	20,9	
K	real	284	1168	672	290	635	1101	397	0,0966
	calc.	340	1392	693	305	151	848	492	
	err.	-20,0	-19,2	-3,20	-5,21	76,2	22,9	-24,1	
P	real	81,4	224	75,2	55,2	85,8	195	90,5	0,145
	calc.	83,4	200	83,4	65,6	95	120	105	
	err.	-2,46	10,6	-10,9	-18,8	-10,7	38,0	-17,0	
S	real	49,9	130,7	149,1	76,6	54,9	451	125,8	0,178
	calc.	67,6	136,5	173,2	93,2	47,9	251,8	170,3	
	err.	-35,4	-4,40	-16,1	-21,6	12,7	44,1	-35,3	

calc.: calculado; err.: erro relativo; amost.: amostra

A Tabela 25 mostra uma diminuição dos erros relativos quando estes são comparados aos erros relativos da Tabela 23, embora a comparação dos erros RMS segundo os testes F não sejam estatisticamente diferentes.

Comparando-se os resultados dos erros RMS das Tabelas 22 (espectros comuns) e 24 (espectros reescalonados), verificamos que existe uma melhora geral nos mesmos, pois os erros diminuíram um pouco, mas considerando-se

apenas os melhores resultados de ambas e aplicando-se um Teste F, verifica-se que não existe uma diferença significativa entre eles ($F_{crit.} = 1,90$, para $n = 32$). Considerando-se os valores da rede BP da Tabela 24 contra a rede BP da Tabela 22:

$$F_1 = \frac{(RMS_{BP(comm)})^2}{(RMS_{BP(reescal)})^2} = \frac{(0,231)^2}{(0,184)^2} = 1,58$$

Existe apenas uma diferença estatisticamente significativa entre o desempenho dos modelos PLS e PCR quando aplicados aos espectros reescalados, em relação aos mesmos modelos para os espectros comuns ($F_{crit.} = 1,90$, para $n = 32$):

$$F_1 = \frac{(PLS_{comum})^2}{(PLS_{reescal})^2} = \frac{(0,275)^2}{(0,188)^2} = 2,14$$

Este último resultado mostra que a teoria do reescalamento tem alguma importância, embora estatisticamente as diferenças sejam muito pequenas. Acreditamos que esta teoria só poderia ser realmente confirmada com o uso de dados simulados, o que não foi objetivo deste trabalho, mas poderá ser concluído em investigações futuras.

Os resultados gráficos que originaram as Tabelas 22 e 24 estão mostrados nas Figuras 64 a 69. As redes BP e LM possuem um comportamento em que o menor erro fica situado em um "vale" sendo que o erro tende a crescer nas vizinhanças do mesmo. Os modelos fuzzy e as redes RBF apresentam um tipo de comportamento semelhante, mas a tendência do erro nas vizinhanças do "vale" não é muito bem definida.

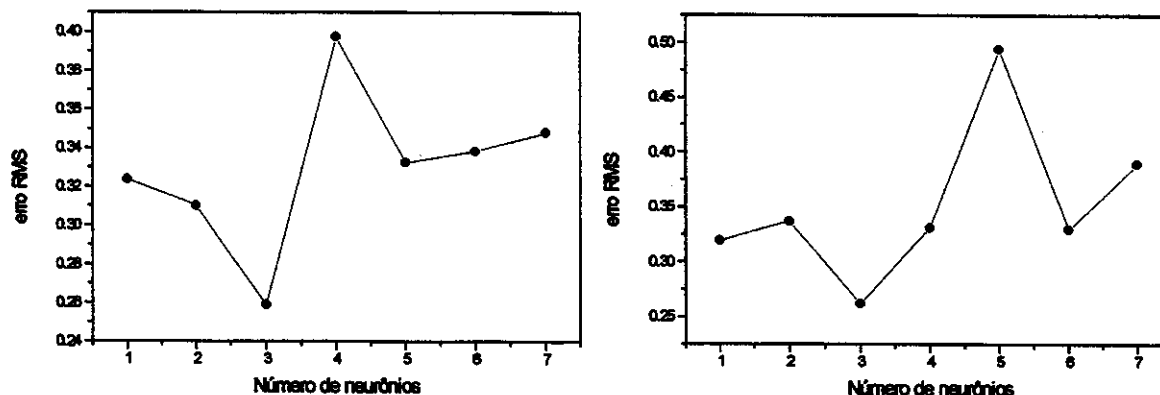


Figura 64- Erros RMS plotados contra o número de neurônios, para as redes BP (esquerda) e LM (direita), resultados mostrados na Tabela 22, utilizando apenas os "scores" (em número de 5) como dados de entrada.

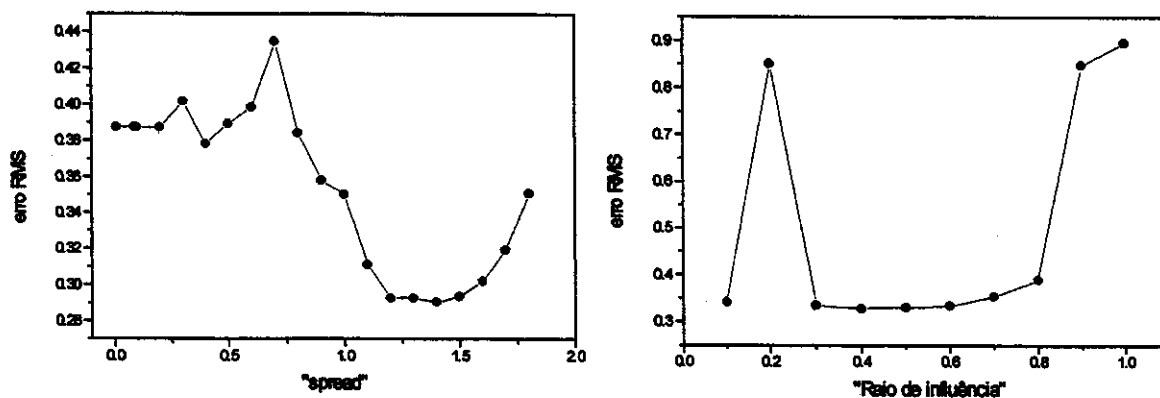


Figura 65 - Erros RMS para a redes RBF (esquerda) e modelo fuzzy (direita), resultados mostrados na Tabela 22, utilizando apenas os "scores" (em número de 5) como dados de entrada.

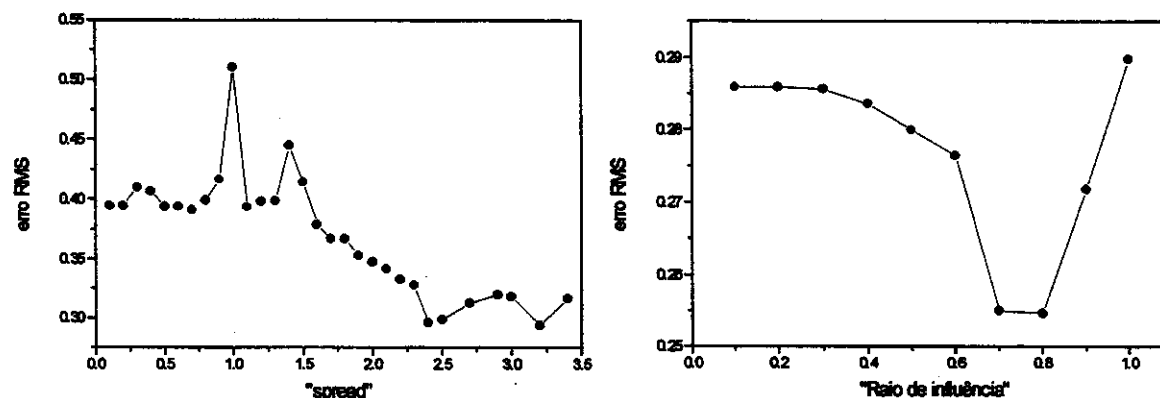


Figura 66 - Erros RMS para a redes RBF (esquerda) e modelo fuzzy (direita), resultados mostrados na Tabela 24, utilizando apenas os "scores" (em número de 7) como dados de entrada.

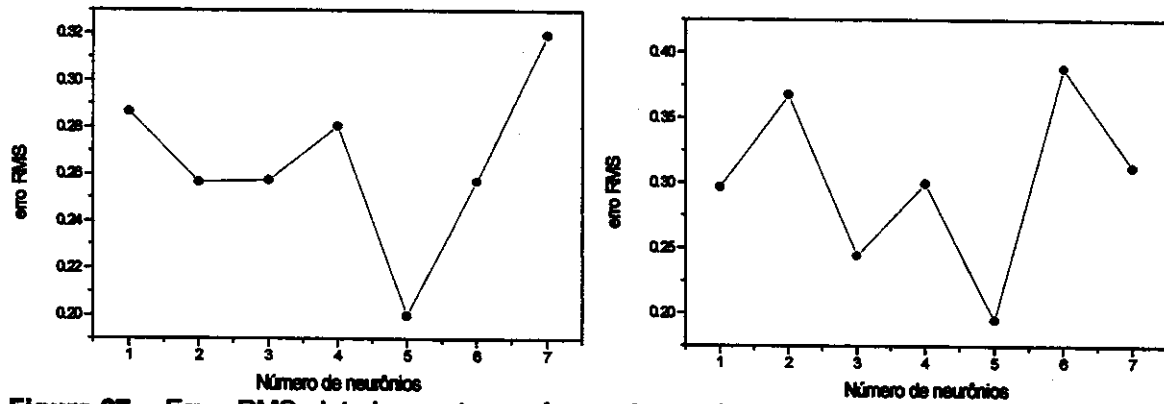


Figura 67 - Erros RMS plotados contra o número de neurônios, para as redes BP (esquerda) e LM (direita), resultados mostrados na Tabela 24, utilizando apenas os "scores" (em número de 7) como dados de entrada.

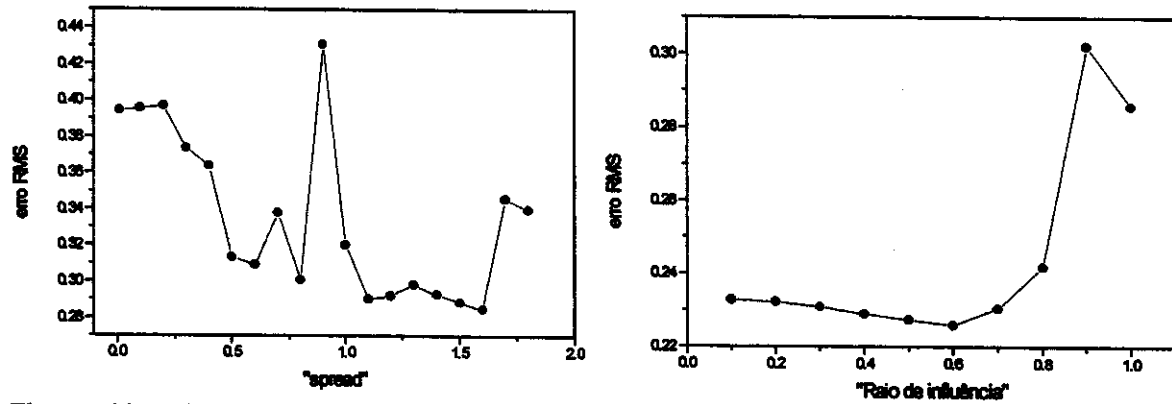


Figura 68 - Erros RMS para a redes RBF (esquerda) e modelo fuzzy (direita), resultados mostrados na Tabela 24, utilizando apenas integrais (em número de 6) como dados de entrada.

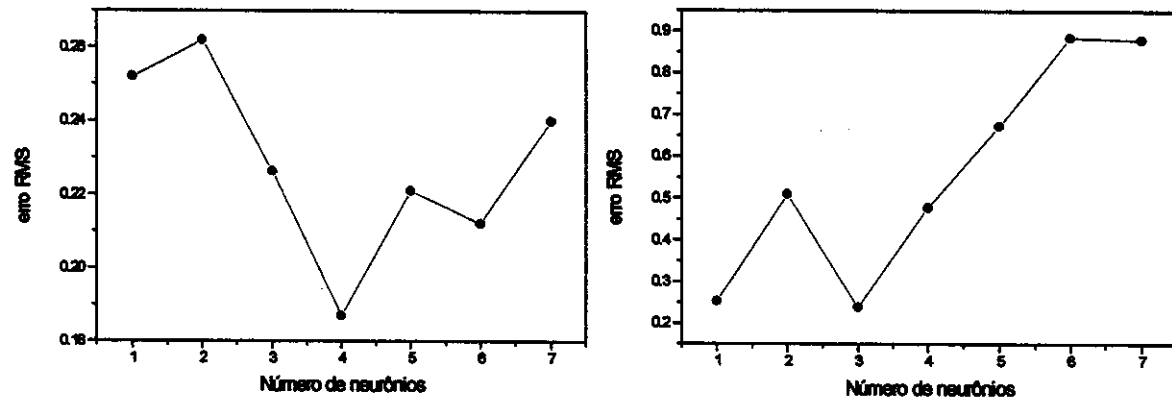


Figura 69 - Erros RMS plotados contra o número de neurônios, para as redes BP (esquerda) e LM (direita), resultados mostrados na Tabela 24, utilizando apenas integrais (em número de 6) como dados de entrada.

4.6.2. Determinação de elemento único

A idéia da determinação de um único elemento foi desenvolvida recentemente por Luo e colaboradores [15] e foi aplicada em calibração para redes neurais. Este método consiste em se determinar a concentração de apenas um elemento de cada vez, ao invés da determinação simultânea (ou multivariada). Segundo os autores, este processo produz uma simplificação nos cálculos para as redes, devido a uma melhor distribuição de pesos entre as camadas, além de um processamento mais rápido, pois há menos conexões entre os neurônios.

Com a finalidade de comprovar esta teoria e verificar sua aplicabilidade em um dos conjuntos de dados utilizados neste trabalho, resolvemos utilizar os espectros dos padrões de plantas e prever as concentrações dos elementos Si, P, S, Ca e K, em espectros comuns.

Para os 32 espectros de plantas e considerando-se os mesmo conjuntos de dados citados anteriormente (Ítem 4.6.1) sem modificações, é mostrado na Tabela 26 os resultados dos erros RMS para uma *rede neural BP*. Estes resultados são os mesmos que originaram a Figura 66. Os elementos da Tabela 26 que estão em *itálico* significam os melhores valores de RMS para cada linha da tabela (cada elemento) em triplicata, e os elementos da Tabela 26 em **negrito**, são os melhores resultados para cada triplicata.

Tabela 26 - Erros RMS calculados para a rede neural BP, mostrando o erro obtido para cada elemento em triplicada, segundo o número de neurônios da camada intermediária ("scores" como dados de entrada).

Neurônios na camada intermediária							
	1	2	3	4	5	6	7
Si	0,34566	0,33716	0,38999	0,38338	0,34022	0,45827	0,34305
Ca	0,35995	0,3704	0,32242	0,34217	0,3774	0,45995	0,31997
K	0,24963	0,21934	0,11878	0,21057	0,29114	0,70603	0,15987
P	0,30943	0,26077	0,16324	0,24825	0,24823	0,41041	0,24281
S	0,32102	0,32294	0,3040	0,30843	0,32528	0,45221	0,49936
Si	0,3655	0,34567	0,2231	0,31473	0,41152	0,21644	0,3647
Ca	0,38896	0,35994	0,39434	0,47081	0,32148	0,30064	0,4219
K	0,25319	0,24963	0,49113	0,93369	0,21591	0,33543	0,63763
P	0,35012	0,30943	0,18307	0,59855	0,2058	0,23327	0,27747
S	0,32836	0,32102	0,46507	0,60184	0,30047	0,15747	0,51491
Si	0,3455	0,34107	0,2454	0,33908	0,54353	0,37633	0,41461
Ca	0,3599	0,37488	0,32095	0,37497	0,19638	0,31176	0,27292
K	0,24946	0,26625	0,22909	0,23657	0,41103	0,12826	0,14617
P	0,3092	0,24997	0,23264	0,27313	0,45289	0,16108	0,18271
S	0,32082	0,32446	0,33729	0,32493	0,34738	0,36741	0,42313

Resumindo, os valores em negrito são os seguintes: 0,11878 para o elemento K; 0,21644 para o elemento Si; 0,15747 para o elemento S; 0,19638 para o elemento Ca e 0,16108 para o elemento P. Estes valores fornecem uma média de 0,170 e este número é significativamente melhor que os resultados mostrados na Tabela 22. Aplicando-se um teste F para o resultado da rede BP (0,231) para os espectros comuns, o valor encontrado é bem próximo do valor crítico ($F_{crit} = 1,90$, para $n = 32$):

$$F_1 = \frac{(RMS_{BP(comm)})^2}{(RMS_{BP(unico)})^2} = \frac{(0,231)^2}{(0,170)^2} = 1,85$$

Um inconveniente deste processo é que torna-se necessário despende um grande tempo de processamento para as redes, e também gera-se vários modelos de rede neural, um para cada elemento químico a ser determinado, necessitando ter um certo cuidado na escolha dos melhores resultados (menores erros) na etapa de treinamento da rede.

Este processo de escolha do menor ocorre apenas para as redes BP e LM, pois está associado a variação do erro em função da adição de neurônios a camada intermediária. No caso dos modelos rede RBF e modelo fuzzy, esta variação do erro em função dos parâmetros escolhidos para cada elemento é muito pequena.

Este processo de determinação de elemento único mostrou alguns resultados interessantes para a rede BP utilizada na proposta de reescalonamento dos espectros de plantas. Nas Tabelas 27 e 28, que serão mostradas na sequência, pode-se observar estes resultados utilizando-se as mesmas redes neurais BP da Tabela 24 com "scores" e integrais como dados de entrada, mas mostrando os erros RMS em triplicata para cada neurônio da camada intermediária.

Os melhores resultados das triplicatas para a Tabela 27 são os seguintes: 0,06236 para o elemento K; 0,04559 para o elemento Si; 0,11857 para o elemento S; 0,11859 para o elemento Ca e 0,06363 para o elemento P. Estes valores fornecem uma média de 0,0817. Aplicando-se um teste F (95%) comparando este resultado com o da rede neural BP para os espectros comuns (Tabela 22) obtém-se um resultado que é significativamente diferente:

$$F_1 = \frac{(\text{RMS}_{\text{BP}(\text{comum})})^2}{(\text{RMS}_{\text{BP}(\text{unico})})^2} = \frac{(0,231)^2}{(0,0817)^2} = 7,99$$

Tabela 27 - Erros RMS calculados para a rede neural BP, mostrando o erro obtido para cada elemento, em triplicada, segundo o número de neurônios da camada intermediária (integrais como dados de entrada) para os espectros reescalados.

Neurônios na camada intermediária							
	1	2	3	4	5	6	7
Si	0,35971	0,07791	0,06403	0,42552	0,08129	0,2953	0,34203
Ca	0,30081	0,3638	0,3612	0,25525	0,24953	0,25103	0,11859
K	0,13345	0,24672	0,20338	0,10253	0,29689	0,22283	0,47113
P	0,19115	0,28344	0,26919	0,11841	0,26546	0,1165	0,15617
S	0,27408	0,24567	0,24518	0,22262	0,3023	0,32016	0,16347
Si	0,35976	0,13095	0,09293	0,15244	0,40334	0,24533	0,31038
Ca	0,30044	0,36851	0,3553	0,32413	0,17639	0,16293	0,23575
K	0,13463	0,27475	0,16295	0,06236	0,17823	0,28501	0,31308
P	0,1911	0,30081	0,31467	0,15641	0,08722	0,13222	0,10067
S	0,27376	0,23114	0,25927	0,23779	0,14872	0,19181	0,2337
Si	0,35976	0,17838	0,04559	0,05576	0,43981	0,41734	0,26492
Ca	0,30044	0,37552	0,35173	0,33922	0,25935	0,18102	0,20921
K	0,13463	0,30824	0,17129	0,16655	0,13665	0,15946	0,27339
P	0,1911	0,31662	0,26046	0,20203	0,17182	0,06363	0,21794
S	0,27376	0,22733	0,23705	0,23802	0,11857	0,13841	0,1917

Tabela 28 - Erros RMS calculados para a rede neural BP, mostrando o erro obtido para cada elemento, em triplicada, segundo o número de neurônios da camada intermediária ("scores" como dados de entrada) para os espectros reescalados.

Neurônios na camada intermediária							
	1	2	3	4	5	6	7
Si	0,34064	0,27864	0,3417	0,35461	0,28653	0,34783	0,34766
Ca	0,36138	0,34115	0,3069	0,35827	0,22786	0,19166	0,35834
K	0,27423	0,14494	0,11855	0,27562	0,13295	0,12028	0,27452
P	0,30329	0,26533	0,23347	0,2857	0,17033	0,22034	0,28922
S	0,32535	0,22518	0,26731	0,32254	0,18213	0,2605	0,32405
Si	0,36064	0,29285	0,36785	0,35378	0,22293	0,34488	0,41793
Ca	0,31776	0,31645	0,31497	0,35842	0,28993	0,27152	0,33677
K	0,11278	0,14639	0,1205	0,27551	0,09254	0,13354	0,16974
P	0,24483	0,25351	0,24628	0,28651	0,21817	0,1249	0,23478
S	0,28246	0,26007	0,27874	0,32268	0,21778	0,25114	0,23022
Si	0,37055	0,36784	0,3426	0,27977	0,2769	0,57278	0,72521
Ca	0,31634	0,31417	0,31344	0,28336	0,29889	0,22605	0,25411
K	0,15788	0,12019	0,13857	0,12118	0,17762	0,47939	0,26966
P	0,25239	0,24463	0,22701	0,16999	0,16003	0,19647	0,37976
S	0,27787	0,27803	0,24428	0,16227	0,2057	0,1141	0,18214

Os melhores resultados das triplicatas para a Tabela 28 são os seguintes: 0,09254 para o elemento K; 0,22293 para o elemento Si; 0,1141 para o elemento S; 0,19166 para o elemento Ca e 0,1249 para o elemento P. Estes valores fornecem uma média de 0,149. Mais uma vez, aplicando-se um teste F (95%) comparando este resultado com o da rede neural BP para os espectros comuns (Tabela 22) obtém-se um resultado que é significativamente diferente:

$$F_1 = \frac{(\text{RMS}_{\text{BP}(\text{comun})})^2}{(\text{RMS}_{\text{BP}(\text{unico})})^2} = \frac{(0,231)^2}{(0,149)^2} = 2,40$$

Estes dois últimos valores médios dos erros RMS, 0,0817 (Tabela 27) e 0,149 (Tabela 28), são significativamente melhores que os resultados mostrados nas Tabelas 22, 24 e 26, o que pode ser comprovado pela aplicação dos testes F. O erros relativos para as concentrações correspondentes aos valores em negrito da Tabela 27 (Tabela 30 do apêndice) são um pouco menores que os da Tabela 25.

Em função disso é possível concluir que o processo de reescalonamento de espectros associado a esta determinação de elemento único pode gerar resultados melhores que os outros modelos aplicados de maneira multivariada a espectros normais, mas são necessários mais estudos para uma conclusão mais significativa.

5. Conclusão

Neste trabalho podemos concluir que:

1) 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 as análises tradicionais, embora as vezes com erros um pouco maiores.

2) 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. Verificamos também que, para uma boa eficiência dos modelos, a maior relação (R) que deve existir entre a maior e a menor concentração deve ser de 260 vezes:

$$R = \frac{\text{maior_conc.}}{\text{menor_conc.}} = 260$$

Por exemplo, este valor de 260 vezes atribuído a uma matriz de concentração normalizada entre 0 e 1, permite detectar variações a partir de 0,0038 unidades de concentração e representa um limite para cálculos convencionais no Matlab™. Se esta relação (R) possuir um valor maior, o limite é ultrapassado e compromete a eficiência da previsibilidade dos modelos quimiométricos. Em função disso, os conjuntos de calibração e validação

utilizados nos modelos devem respeitar esta relação (R) antes da sua construção.

3) 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.

4) Para as redes neurais, a observação do erro RMS em função da adição de neurônios as camadas intermediárias permite que se encontre o menor erro de uma maneira segura e eficiente. Nos trabalhos encontrados na literatura sobre redes neurais, sempre existem algumas dúvidas sobre qual o número ideal de neurônios nesta camada. A maioria dos autores não explicam como chegaram a este número.

Este trabalho não visou estudar os detalhes estruturais das redes neurais e dos modelos fuzzy, pois isso geraria uma quantidade muito grande de parâmetros a serem controlados. Então optou-se por considerar apenas os valores padrão sugeridos pelos manuais específicos, visando apenas aplicações em calibração multivariada.

5) Em relação a utilização do melhor tipo de conjunto de dados de entrada para dados de EDXRF, baseado nos resultados deste trabalho, podemos até o momento afirmar que os "scores" são melhores do que a utilização de partes do espectro, e de modo geral, a utilização das integrais das linhas espectrais são mais eficientes que o emprego dos "scores" (quando determinados por CV), embora acreditemos que seja necessária a utilização de mais amostras reais com muitos elementos para confirmação deste fato. A utilização das linhas espectrais produziu erros menores para os espectros de EDXRF, pois elas geralmente são fixas em torno de um valor. A sua aplicação em outros tipos de espectros ficaria restrita a este fator.

6) As redes neurais RBF e os modelos fuzzy, cujas aplicações em fluorescência de raios-X são inéditas em quimiometria, 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.

7) O reescalonamento de espectros por si só não apresentou grandes vantagens, mas associado a determinação de elemento único parece ser mais promissor. Ele necessita ser estudado com um conjunto grande de espectros e com dados simulados, para verificar a sua viabilidade.

8) A determinação de elemento único aplicada a redes neurais parece ter um campo mais promissor que a determinação multivariada. Como afirmou Luo e colaboradores [13], este processo produz uma simplificação nos cálculos para as redes, devido a uma melhor distribuição de pesos entre as camadas, além de um processamento mais rápido, pois há menos conexões entre os neurônios.

9) Finalmente, este trabalho permite que seja feita uma extensão da aplicação destes modelos a outros tipos de espectros ou conjuntos de dados, oriundos de outras técnicas analíticas.

6. Referências bibliográficas

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. Wang, Z., Hwang, J.N. and Kovalski, B.R., *Anal. Chem.*, **1995**, (67)9, 1497.
4. Arnold, T., Otto, M. and Wegscheider, W., *Talanta*, **1994**, (41)7, 1169-1184.
5. Shimosaka, T., Kitamori, T., Harata, A. and Sawada, T., *Anal. Sci.*, **1996**, 12, 385-391.
6. Ferreira, M.C., Antunes, A.M., Melgo, M.S. and Volpe, P.L.O., *Química Nova*, **1999**, (22)5, 724-731.
7. Otto, M., "Chemometrics", Wiley-VCH Verlag, Weinheim, 1999.
8. Zuppan, J. and Gasteiger, J., "Neural Networks for Chemists", VCH, 1st ed., 1993, New York.
9. Zuppan, J. and Gasteiger, J., *Anal. Chim. Acta*, **1991**, 248, 1.
10. Cabral Jr, E.F (editor). "Redes Neurais Artificiais", Grupo Comunicação Homen - Máquina, Escola Politécnica da USP, São Paulo, 1999.
11. Demuth, H. and Beale, M., "Neural Network Toolbox, for use with MATLABTM - User's Guide", versão 2.0, Mathworks Inc., 1994.
12. Bos, A., Bos, M. and van der Linden, W.E., *Anal. Chim. Acta*, **1993**, 277, 289-295.
13. Bos, M. and Weber, H.T., *Anal. Chim. Acta*, **1991**, 247, 97-105.
14. Luo, L., Guo, C., Ma, G. and Ji, A., *X-Ray Spectrom.*, **1997**, 26, 15-22.
15. Luo, L., Ji, A., Ma, G. and Guo, C., *X-Ray Spectrom.*, **1998**, 27, 17-22.
16. Boger, Z. and Karpas, Z., *Anal. Chim. Acta*, **1994**, 292, 243-251.

17. Facchin, I., Mello, C., Bueno, M.I.M.S., and Poppi, R.J., *X-Ray Spectrom.*, **1999**, 28, 173-177.
18. Pedrycz, W. and Gomide, F., "An Introduction to Fuzzy Sets: Analysis and Design", MIT Press, Cambridge, 1998.
19. "Fuzzy Logic Toolbox User's Guide", versão 2.0, The Mathworks Inc., 1998.
20. Kantrowitz, M. et al, "Answers to Frequently Asked Fuzzy Questions", Arquivo texto versão 1.27, 14/03/97, Organization: School of Computer Science, Carnegie Mellon University, URL: <http://www.cs.cmu.edu/Web/Groups/AI/html/faqs/ai/fuzzy/part1/faq.html>
21. Chiu S., *Journal of Intelligent & Fuzzy Systems*, **1994**, 2, 267-278.
22. Tagaki, T and Sugeno, M., *IEEE Transactions on Systems, Man and Cybernetics*, **1985**, (15)1, 116-132.
23. Otto, M., *Anal. Chim. Acta*, **1990**, 235, 169-175.
24. Otto, M., *Trends in Anal. Chem.*, **1990**, (9)2, 69-72.
25. Otto, M., *Chem. Intel. Lab. Sys.*, **1988**, 4, 101-120.
26. Otto, M. and Bandemer, H., *Anal. Chim. Acta*, **1986**, 191, 193-204.
27. Wise, B.M. and Gallagher, N.B., "PLS Toolbox, for use with MATLABTM", versão 1.5, Eigenvector Technologies , 1995.
28. Wang, Z., Hwang, J. and Kowalski, B.R., *Anal. Chem.*, **1995**, 67, 1497-1504.
29. Skoog, D.A., West, D.M. and Holler, F.J., "Fundamentals of Analytical Chemistry", Saunders College Publishing, 6th ed., Florida, 1992.
30. Thomas, E.V., *Anal. Chem.*, **1994**, 66, 795A-804A.
31. Cornejo-Ponce, L., Peralta-Zamora, P. and Bueno, M.I.M.S., *Talanta*, **1998**, 46, 1371-1378.

7. Apêndice

7.1. Determinação de elemento único

Tabela 29 - Erros relativos para a rede BP com "scores" (espectros comuns), correspondente aos melhores resultados da Tabela 26.

RMS			amost.						
			1	2	3	4	5	6	7
0,2164	Si	real	877	5233	1188	9475	3513	37509	1207
		calc.	2781	4503	3880	3924	3075	22628	3561
		err.	-217	13,9	-226	58,5	12,4	39,6	-195
0,1963	Ca	real	163	484	567	160	81,9	2092	467
		calc.	123,9	332,1	257,8	108,4	131	1651,9	832,7
		err.	23,9	31,3	54,5	32,2	-59,9	21,0	-78,3
0,1187	K	real	284	1168	672	290	635	1101	397
		calc.	427,8	1281,7	667,6	135,1	547,7	1267,5	628,4
		err.	-50,6	-9,73	0,65	53,4	13,7	-15,1	-58,2
0,1610	P	real	81,4	224	75,2	55,2	85,8	195	90,5
		calc.	85	163,8	104	59,8	78,1	143	131,1
		err.	-4,42	26,8	-38,3	-8,33	8,97	26,6	-44,8
0,1574	S	real	49,9	130,7	149,1	76,6	54,9	451	125,8
		calc.	80,8	128,7	94	135,7	81,3	308,1	230,1
		err.	-61,9	1,53	36,9	-77,1	-48,0	31,6	-82,9

média: 0,170

Tabela 30 - Erros relativos para a rede BP com integrais (espectros reescalados), correspondente aos melhores resultados da Tabela 27.

RMS			amost.						
			1	2	3	4	5	6	7
0,0455	Si	real	877	5233	1188	9475	3513	37509	1207
		calc.	545	5885	939	6259	749	38743	1910
		err.	37,8	-12,4	20,9	33,9	78,6	-3,29	-58,3
0,1185	Ca	real	163	484	567	160	81,9	2092	467
		calc.	173	598	553	165	20,9	900	495
		err.	-6,2	-23,7	2,43	-3,63	74,4	56,9	-6,04
0,0623	K	real	284	1168	672	290	635	1101	397
		calc.	311,9	1270,3	716,6	264,8	152,9	1188	454,8
		err.	-9,82	-8,76	-6,64	8,69	75,9	-7,9	-14,5
0,0636	P	real	81,4	224	75,2	55,2	85,8	195	90,5
		calc.	85,1	219,4	76,2	73,3	96,5	182,4	105,1
		err.	-4,55	2,05	-1,33	-32,7	-12,4	6,46	-16,1
0,1185	S	real	49,9	130,7	149,1	76,6	54,9	451	125,8
		calc.	63,5	152,9	193,1	74	45,6	417,8	231,6
		err.	-27,2	-16,9	-29,5	3,39	16,9	7,36	-84,1

média: 0,0817

Tabela 31 - Erros relativos para a rede BP com "scores" (espectros reescalados), correspondente aos melhores resultados da Tabela 28.

RMS			amost.						
			1	2	3	4	5	6	7
0,2229	Si	real	877	5233	1188	9475	3513	37509	1207
		calc.	182,7	4599	2976	9431	5455	2170	7890
		err.	79,1	12,1	-150	0,46	-55,2	94,2	-553
0,1916	Ca	real	163	484	567	160	81,9	2092	467
		calc.	164,6	375,2	952,8	60,4	33	1295,1	616,8
		err.	-0,98	22,4	-68,0	62,2	59,7	38,1	-32,1
0,0925	K	real	284	1168	672	290	635	1101	397
		calc.	435,3	689,8	833	539,6	162,5	445,4	35,8
		err.	-53,2	40,9	-23,9	-86,0	74,4	59,5	90,9
0,1249	P	real	81,4	224	75,2	55,2	85,8	195	90,5
		calc.	105	106,4	114	85,1	107,2	165,8	85,9
		err.	-28,9	52,5	-51,6	-54,1	-24,9	14,9	5,08
0,1141	S	real	49,9	130,7	149,1	76,6	54,9	451	125,8
		calc.	83,7	113,5	161,5	106,8	54,6	364,9	155
		err.	-67,7	13,1	-8,32	-39,4	0,55	19,1	-23,2

média: 0,149