



UNIVERSIDADE
ESTADUAL DE LONDRINA

EVANDRO BONA

**INTEGRAÇÃO DE REDES NEURAS ARTIFICIAIS AO NARIZ
ELETRÔNICO: AVALIAÇÃO AROMÁTICA DE CAFÉ
SOLÚVEL**

Londrina
2008

EVANDRO BONA

**INTEGRAÇÃO DE REDES NEURAIIS ARTIFICIAIS AO NARIZ
ELETRÔNICO: AVALIAÇÃO AROMÁTICA DE CAFÉ
SOLÚVEL**

Tese apresentada ao Programa de Mestrado e Doutorado em Ciência de Alimentos da Universidade Estadual de Londrina, como requisito parcial à obtenção do título de Doutor em Ciência de Alimentos.

Orientador: Prof. Dr. Rui Sérgio dos Santos Ferreira da Silva

Londrina
2008

EVANDRO BONA

**INTEGRAÇÃO DE REDES NEURAIS ARTIFICIAIS AO NARIZ
ELETRÔNICO: AVALIAÇÃO AROMÁTICA DE CAFÉ SOLÚVEL**

Tese apresentada ao Programa de Mestrado e Doutorado em Ciência de Alimentos da Universidade Estadual de Londrina, como requisito parcial à obtenção do título de Doutor em Ciência de Alimentos.

BANCA EXAMINADORA

Prof. Dr. Rui Sérgio dos Santos Ferreira da
Silva

Prof. Dr. Dionísio Borsato

Prof. Dr. Luiz Henry Monken e Silva

. Dr. Ernesto Fernando Ferreyra Ramírez

Dr. Denisley Gentil Bassoli

Londrina, 25 de agosto de 2008.

DEDICATÓRIA

A Deus, a minha esposa Adriana, aos meus pais e irmãos pelo constante apoio e carinho incondicional durante todos os momentos de minha caminhada acadêmica.

AGRADECIMENTOS

Ao professor **Dr. Rui Sérgio dos Santos Ferreira da Silva**, pela orientação, conselhos e pela inestimável amizade construída ao longo destes 11 anos.

Ao professor **Dr. Dionísio Borsato**, pela valiosa e indispensável colaboração durante a elaboração desta Tese.

Ao professor **Dr. Ernesto Fernando Ferreyra Ramírez**, pelas aulas e dúvidas esclarecidas sobre redes neurais artificiais.

Aos **professores do programa de Doutorado em Ciência de Alimentos** e para secretária **Sandra Rezende** pelo apoio.

À **Companhia Iguaçu de Café Solúvel** que, representada pelo **Dr. Denisley Gentil Bassoli**, forneceu todo o suporte experimental e analítico necessário para este trabalho.

À **Fundação Araucária**, pelo suporte financeiro ao projeto.

À **Coordenação de Aperfeiçoamento de Pessoal em Nível Superior (CAPES)**, pela concessão da bolsa de doutorado.

BONA, Evandro. **Integração de redes neurais artificiais ao nariz eletrônico: avaliação aromática de café solúvel.** 2008. 159 f. Tese (Doutorado em Ciência de Alimentos) – Universidade Estadual de Londrina, Londrina – PR.

RESUMO

O aroma é uma das características mais importantes dos alimentos, principalmente para bebidas como o café. Sua avaliação é complexa e primordial no controle de qualidade da indústria de café solúvel. O nariz eletrônico é um instrumento, constituído por um arranjo de sensores parcialmente seletivos, que tem sido muito utilizado na análise de voláteis alimentares. É necessário, também, que ao equipamento seja acoplada uma interface multivariada capaz de reconhecer padrões aromáticos. As redes neurais artificiais, que são um conjunto de técnicas capazes de realizar mapeamentos entrada-saída complexos e não-lineares, estão entre as técnicas mais empregadas para a análise dos perfis aromáticos obtidos através do nariz eletrônico. Neste trabalho foram utilizados dois tipos de rede neural artificial, o perceptron de múltiplas camadas (aprendizagem supervisionada) e o mapa auto-organizável (aprendizagem não-supervisionada), para o reconhecimento e classificação de padrões aromáticos de café solúvel em dois bancos de dados. No desenvolvimento das redes, foram aplicadas outras metodologias para: geração de conjuntos de treinamento (reamostragem *bootstrap*), escolha dos parâmetros de rede (otimização simplex seqüencial) e para verificar a confiabilidade e variabilidade dos resultados obtidos (estatística descritiva e testes de hipóteses). As redes perceptron de múltiplas camadas reconheceram todos os sete cafés do primeiro banco de dados. Já para o segundo banco com quarenta produtos industriais, foi obtido um desempenho de classificação correta de 90% nas amostras de validação e 75% para as de teste. Esta performance só foi possível em virtude da utilização das metodologias de reamostragem, otimização e a correta representação das classes. Os mapas auto-organizáveis possibilitaram a separação das amostras, de ambos os bancos de dados, em grupos de acordo com similaridade aromática. Para o primeiro banco de dados os resultados obtidos pelo mapa convergiram com os apresentados pela análise de componentes principais. O mesmo não foi possível para o banco industrial, pois, não foi possível visualizar uma separação dos produtos usando a análise de componentes principais. Além disso, o mapa permitiu testar estatisticamente a proximidade entre produtos e associar a distribuição das amostras com a dispersão dos valores observados para cada sensor do nariz eletrônico. A combinação do nariz eletrônico com as redes neurais artificiais mostrou ser uma alternativa promissora para o desenvolvimento de novos produtos, comparação de similaridades entre produtos concorrentes e para o controle da qualidade aromática de café solúvel.

Palavras-chave: Perceptron de múltiplas camadas. Mapas auto-organizáveis. Reamostragem *bootstrap*. Média de ensemble.

BONA, Evandro. **Artificial Neural networks integrated to electronic nose system**: aromatic evaluation of soluble coffee. 2008. 159 f. Thesis (Doctorate in Food Science) – Universidade Estadual de Londrina, Londrina – PR.

ABSTRACT

Flavor is one of the most important features for beverages like coffee. The volatile compounds evaluation is complex and important for quality control in soluble coffee industry. The electronic nose is an instrument, which comprises an array of chemical sensors with partial specificity, very used for food flavor analysis. It is also necessary to couple the equipment with a suitable multivariate pattern recognition system. The artificial neural networks are a group of methods that can represent complex and nonlinear input-output relationships. The neural network integrated to electronic nose system is greatly used for food aroma analysis. In this work were used two types of neural networks for flavor recognition and classification of two soluble coffee databases, the multilayer perceptron (supervised learning) and the self-organizing maps (unsupervised learning). Other techniques were also applied during neural network development: bootstrap resample to generate training sets; sequential simplex optimization to select parameters; and descriptive statistics and hypothesis tests to assess the variability and reliability of obtained results. The multilayer perceptrons recognized all coffees in the first database with seven products. For the industrial database, with forty products, 90% of samples in validation set and 75% in test set were correctly classified. This performance was only achieved due to correct class representation, bootstrap resample and optimization methodology. The self-organizing maps made possible the separation of coffee samples according to aromatic similarity in both databases. In the first database the results obtained by the self-organizing maps are the same ones obtained by principal component analysis. But, in the industrial database was not possible to visualize the products separation using the principal component analysis. Furthermore, in the self-organizing map it was possible to statistically test the neighborhood relation of products and associate the samples distribution with electronic nose sensor values dispersion. The neural network integrated to electronic nose system showed as a promising choice for new products development, similarity check among concurrent products and quality control of soluble coffee.

Keywords: Multilayer perceptron. Self-organizing maps. Bootstrapped samples. Ensemble average.

LISTA DE FIGURAS

CAPÍTULO 1

- Figura 1** – Esquema ilustrativo de um nariz eletrônico22
- Figura 2** – Representação de um neurônio biológico25
- Figura 3** – Representação gráfica da rede perceptron de múltiplas camadas com duas camadas ocultas34
- Figura 4** – Mapa auto-organizável segundo o modelo de Kohonen39
- Figura 5** – Representação de vizinhança topológica: (a) gráfico unidimensional da função gaussiana; (b) representação de uma vizinhança quadrada40

CAPÍTULO 2

- Figura 1** – Mapa auto-organizável segundo o modelo de Kohonen56
- Figura 2** – Representação de vizinhança topológica: (a) gráfico unidimensional da função gaussiana; (b) representação de uma vizinhança quadrada57
- Figura 3** – Gráfico de autovalor para o critério do teste scree65
- Figura 4** – Projeção dos escores (a) e das cargas (b) fatoriais nas duas componentes principais selecionadas66
- Figura 5** – Comportamento do coeficiente de variação do EMQ e coeficiente de correlação do gráfico $dy-dx$ em função do número de neurônios em cada dimensão do mapa auto-organizável69
- Figura 6** – Mapa auto-organizável 4 x 4: variação ao longo do treinamento da taxa de aprendizagem (a), da largura da função de vizinhança (b) e do erro médio de quantização (c); gráfico $dy-dx$ (d).....70
- Figura 7** – Gráficos de probabilidade acumulada para um raio de vizinhança igual a 1(a) e 0 (b)71
- Figura 8** – Gráficos com os valores de $EVIZ_{i,j}(r)$ para as amostras 1 (a), 19 (b) e 16 (c)73
- Figura 9** – Gráficos com os valores de $EVIZ_{i,j}(r)$ para as amostras 4 (a), 6 (b) e 7 (c).....74

Figura 10 – Gráficos com os valores de $EVZ_{i,j}(r)$ para as amostras 10 (a), 13 (b) e 15 (c).....	75
Figura 11 – Mapas topológicos 4 x 4 de duas repetições de treinamento	76
Figura 12 – Mapas de peso de alguns sensores selecionados (20ª repetição de treinamento)	77

CAPÍTULO 3

Figura 1 – Representação esquemática do olfato humano e nariz eletrônico, adaptado de Craven et al. (1996).....	85
Figura 2 – Representação gráfica da rede perceptron de múltiplas camadas com duas camadas ocultas	88
Figura 3 – Representação de uma RNA com duas entradas, duas saídas e uma camada oculta de dois neurônios	93
Figura 4 – Erro quadrático médio e porcentagem de classificação das redes otimizadas	95
Figura 5 – Porcentagem de confiança média na classificação realizada pelas redes 2 e 4	97
Figura 6 – Gráficos de dispersão das áreas médias de alguns sensores	99
Figura 7 – Gráficos de dispersão das áreas médias de alguns sensores	100
Figura 8 – Dendrograma das amostras de café	101

CAPÍTULO 4

Figura 1 – Representação gráfica da rede perceptron de múltiplas camadas com duas camadas ocultas	113
Figura 2 – Mapa auto-organizável segundo o modelo de Kohonen	115
Figura 3 – Projeção dos escores e das cargas fatoriais nas três componentes principais selecionadas	121
Figura 4 – Mapa topológico de uma das repetições	126
Figura 5 – Mapas de pesos de um dos sensores de cada grupo identificado ..	127

APÊNDICE

Figura 1 – Interpretação geométrica do simplex (a) para duas variáveis (b) três variáveis	136
Figura 2 – Movimento de reflexão em um simplex para duas variáveis	138

Figura 3 – Intervalos proibidos para os valores de β (BARROS NETO et al., 1995)	139
Figura 4 – Gráfico das respostas transformadas para diversos valores de s ...	140

LISTA DE TABELAS

CAPÍTULO 1

Tabela 1 – Grupos de voláteis de impacto no aroma do café torrado (MELLO; TRUGO, 2003)	20
Tabela 2 – Descrição simplificada das amostras de cafés solúveis analisadas	30
Tabela 3 – Identificação dos produtos e quantidade de lotes avaliados	31
Tabela 4 – Especificidade dos sensores utilizados no nariz eletrônico PEN2.....	32

CAPÍTULO 2

Tabela 1 – Descrição resumida das amostras de café solúvel analisadas	53
Tabela 2 – Especificidade dos sensores utilizados no nariz eletrônico PEN2	54
Tabela 3 – Comparação entre as áreas médias obtidas pelo nariz eletrônico para cada sensor e café	64
Tabela 4 – Correlação entre as áreas de cada sensor	67
Tabela 5 – Resumo dos mapas auto-organizáveis treinados	67
Tabela 6 – Propriedades dos raios de vizinhança quadrada testados no mapa 4 x 4	70

CAPÍTULO 3

Tabela 1 – Descrição simplificada das amostras de cafés solúveis analisadas	86
Tabela 2 – Especificidade dos sensores utilizados no nariz eletrônico PEN2	87
Tabela 3 – Planejamento fatorial 2 x 3 com 5 repetições para cada experimento	91
Tabela 4 – Parâmetros ótimos obtidos para as diferentes redes propostas	94
Tabela 5 – Índice de casualidade médio para a rede 4	98
Tabela 6 – Comparação entre as áreas médias de cada sensor para cada amostra	100

CAPÍTULO 4

Tabela 1 – Identificação dos produtos e quantidade de lotes avaliados	110
Tabela 2 – Especificidade dos sensores utilizados no nariz eletrônico PEN2.....	111
Tabela 3 – Estatísticas descritivas dos produtos analisados em função do sensor do nariz eletrônico	119
Tabela 4 – Correlação entre as áreas de cada sensor	120
Tabela 5 – Comparação de desempenho das redes MLP utilizadas	123
Tabela 6 – Influência da quantidade de repetições na resposta do ensemble	124
Tabela 7 – Matriz de confundimento das redes MLP utilizadas	125

SUMÁRIO

INTRODUÇÃO	15
CAPÍTULO I	18
1.1 REVISÃO BIBLIOGRÁFICA.....	19
1.1.1 Café	19
1.1.2 Nariz Eletrônico	21
1.1.3 Sistemas Especialistas	23
1.1.4 Redes Neurais Artificiais.....	24
1.2 OBJETIVOS	29
1.2.1 Objetivo Geral.....	29
1.2.2 Objetivos Específicos	29
1.3 MATERIAIS E MÉTODOS.....	30
1.3.1 Café Solúvel	30
1.3.2 Nariz Eletrônico	31
1.3.3 Integração e Normalização	32
1.3.4 Rede Perceptron de Múltiplas Camadas	33
1.3.5 Mapa Auto-Organizável	38
1.3.6 Bootstrap	42
1.3.7 Média de Ensemble	42
1.3.8 Otimização dos Parâmetros da Rede (MLP)	43
1.3.9 Implementação Computacional	43
1.4 REFERÊNCIAS	44
CAPÍTULO II – EMPREGO DE MAPAS AUTO-ORGANIZÁVEIS NO RECONHECIMENTO DE PADRÕES AROMÁTICOS DE CAFÉ SOLÚVEL	49
2.1 INTRODUÇÃO	50
2.2 MATERIAIS E MÉTODOS	53
2.2.1 Café Solúvel	53
2.2.2 Nariz Eletrônico	54
2.2.3 Integração e Autoescalonamento	54
2.2.4 Mapa Auto-Organizável (SOM)	55
2.2.5 Preservação da Topologia	59

2.2.6	Confiabilidade do Mapa Auto-Organizável	59
2.2.7	Análise de Componentes Principais (ACP)	62
2.2.8	Implementação Computacional	63
2.3	RESULTADOS E DISCUSSÃO	63
2.3.1	Pré-Processamento	63
2.3.2	Análise de Componentes Principais	64
2.3.3	Mapa Auto-Organizável	67
2.4	CONCLUSÕES	78
2.5	REFERÊNCIAS	79
2.6	AGRADECIMENTOS	81

CAPÍTULO III – REDE NEURAL OTIMIZADA PARA A CLASSIFICAÇÃO DE CAFÉ SOLÚVEL ATRAVÉS DE UM NARIZ ELETRÔNICO:

I) DESENVOLVIMENTO PARA ALGUMAS AMOSTRAS DE CAFÉ

SOLÚVEL	82
3.1 INTRODUÇÃO	83
3.2 MATERIAIS E MÉTODOS	86
3.2.1 Café Solúvel	86
3.2.2 Nariz Eletrônico	86
3.2.3 Normalização	87
3.2.4 Rede Neural Artificial	87
3.2.5 Bootstrap	90
3.2.6 Média de Ensemble	90
3.2.7 Otimização dos Parâmetros de Rede	91
3.2.8 Índice de Casualidade	92
3.2.9 Implementação Computacional	93
3.3 RESULTADOS E DISCUSSÃO	93
3.4 CONCLUSÕES	102
3.5 REFERÊNCIAS	103
3.6 AGRADECIMENTOS	106

CAPÍTULO IV – REDE NEURAL OTIMIZADA PARA A CLASSIFICAÇÃO DE CAFÉ SOLÚVEL ATRAVÉS DE UM NARIZ ELETRÔNICO: II) GENERALIZAÇÃO PARA UM BANCO DE DADOS INDUSTRIAL

107

4.1	INTRODUÇÃO	108
4.2	MATERIAIS E MÉTODOS	110
4.2.1	Café Solúvel	110
4.2.2	Nariz Eletrônico	111
4.2.3	Variáveis de Entrada: Integração e Autoescalamento	111
4.2.4	Variáveis de Saída: Representação dos Produtos	112
4.2.5	Redes Neurais Artificiais	112
4.2.6	Amostragem e Bootstrap	116
4.2.7	Média de Ensemble para a Rede MLP	117
4.2.8	Otimização dos Parâmetros da Rede MLP	117
4.2.9	Implementação Computacional	118
4.3	RESULTADOS E DISCUSSÃO	118
4.3.1	Pré-Processamento	118
4.3.2	Redes Neurais	121
4.4	CONCLUSÕES	128
4.5	REFERÊNCIAS	128
4.6	AGRADECIMENTOS	132
5	CONCLUSÃO GERAL	133
5.1	CONCLUSÕES GERAIS	133
5.2	SUGESTÕES PARA TRABALHOS FUTUROS	134
	APÊNDICES	135
	APÊNDICE A.....	136
	APÊNDICE B.....	140
	APÊNDICE C	142
	ANEXOS	143
	ANEXO A	144
	ANEXO B	148
	ANEXO C	155

INTRODUÇÃO

O aroma é um dos atributos mais importantes em alimentos, assim, sua avaliação objetiva é uma das tarefas mais importantes na indústria de alimentos. Para o café, o aroma é o critério mais importante na avaliação de qualidade e um dos parâmetros determinantes na escolha pelo consumidor (FARAH et al., 2006). O aroma do café é formado por inúmeros componentes com variados grupos funcionais e essa composição depende de fatores como espécie e variedade, condições de crescimento e colheita, armazenagem, intensidade de torra e tipo de torrador, além de outras condições de processo (MELLO; TRUGO, 2003). A partir do café torrado, são utilizadas colunas extratoras na industrialização do café solúvel para a obtenção do extrato hidrossolúvel de café (CLARKE, 2001). A busca pela aproximação do aroma do café solúvel ao do café de coador tem sido constante, e em geral, meta importante dos processos de aumento da intensidade dos aromas presentes no café solúvel ou mesmo da sua aromatização (BASSOLI, 2006).

A percepção dos compostos voláteis do café solúvel pelo nariz de provadores especializados é uma técnica usada habitualmente na indústria de alimentos e bebidas (DUTCOSKY, 2007). Entretanto, nas últimas décadas, o nariz eletrônico tem sido usado com frequência na análise de voláteis em alimentos (SCHALLER et al., 1998; DEISINGH et al., 2004) e até mesmo em café torrado e moído (PARDO; SBERVEGLIERI, 2002). O nariz eletrônico é um instrumento contendo um grupo de sensores gasosos, com especificidade parcial. Cabe ao equipamento emular o funcionamento de um nariz humano, porém, é necessário também um sistema multivariado e “inteligente” capaz de desempenhar um papel semelhante ao córtex cerebral no reconhecimento de padrões aromáticos (CRAVEN et al., 1996). Existe uma relação entre os narizes eletrônico e humano. Pois, cada um desses é composto de três regiões principais. No humano existe um grupo de células que atuam como receptores dos odores, o bulbo olfativo e o cérebro. No nariz eletrônico o equivalente a estas regiões seria o grupo de sensores de odor, um pré-processador de sinal e o sistema de reconhecimento de padrões (DEISINGH et al., 2004).

O reconhecimento e a classificação de padrões é uma tarefa naturalmente realizada pelos humanos graças à evolução e adaptação, durante

milhares de anos, de nosso sistema nervoso central, mais especificamente o córtex cerebral. Entretanto, a solução de problemas de classificação através de sistemas automatizados, na maioria dos casos, é de extrema complexidade. Principalmente se os padrões forem descritos por um grande número de variáveis independentes (BISHOP, 1995; HAYKIN, 2001). Existem várias metodologias estatísticas tradicionais e consagradas, tanto paramétricas quanto não paramétricas, usadas para a classificação e reconhecimento de padrões. Um dos métodos multivariados mais tradicionais é a análise de componentes principais (ACP). Nesta técnica, dados de alta dimensionalidade são reduzidos a um sistema de coordenadas lineares e de baixa dimensão com o objetivo de facilitar a visualização de tendências ou agrupamentos (HAIR et al., 2005). Entretanto, algumas desvantagens do método são bem conhecidas (MELSSSEN et al., 2006). Primeiramente, considera-se que os dados podem ser descritos por combinações lineares. Conseqüentemente, sistemas não-lineares não serão bem representados. Um segundo ponto crítico na ACP é a qualidade do resultado que pode ser muito influenciada por amostras discrepantes. Somados a estes dois fatos, existe a possibilidade de após as transformações, o número de componentes significativas ainda ser elevado dificultando a extração de conhecimento útil dos dados.

As redes neurais artificiais (RNA) são um conjunto de técnicas baseadas em princípios estatísticos não tão convencionais, que vem crescentemente ganhando espaço no reconhecimento e classificação de padrões (BISHOP, 1995; WARNER; MISRA, 1996). As RNA são extremamente versáteis para realizar o mapeamento de relações complexas e não-lineares entre múltiplas variáveis de entrada e saída. A construção de regras de classificação baseadas apenas nos dados disponíveis, sem a imposição de um modelo a priori, é outro atrativo da técnica (SABLANI, 2008). Entre as desvantagens da metodologia citam-se: a necessidade de uma grande quantidade de dados de treinamento (BISHOP, 1995); dificuldade na escolha dos parâmetros de treinamento e do tipo de rede mais adequada ao problema (CURRY; MORGAN, 2006); variabilidade dos resultados devido à inicialização e amostragem (BODT et al., 2002). Porém, vários trabalhos vêm apresentando propostas promissoras para solucionar ou amenizar as desvantagens associadas às RNA (LOONEY, 1996; KOLMAN; MARGALLOT, 2005; LUDERMIR et al., 2006; WINDEATT, 2006). Como pode ser observado nas revisões realizadas por Liao (2004) e Huang et al. (2007), as RNA tem sido empregadas com

sucesso dentro da área de ciência, tecnologia e engenharia de alimentos. Em vários trabalhos (DÍAZ et al., 2003; LINDER; PÖPPL, 2003; PANIGRAHI et al., 2006; CHANDRARATNE et al., 2007; MARINI et al., 2007; TIGRINE-KORDJANI, 2007) foi obtido um desempenho superior ou similar dos classificadores neurais quando comparados com técnicas estatísticas tradicionais. Além disso, os modelos neurais estão sendo apontados como ferramentas promissoras para o controle de qualidade (NI; GUNASEKARAN, 1998) e análise sensorial (BOMIO, 1998; DELLA LUCIA; MINIM, 2006).

Considerando a importância comercial e a necessidade do controle de qualidade aromática do café solúvel, propôs-se a integração das redes neurais artificiais ao nariz eletrônico. Foram testados tanto os perceptrons de múltiplas camadas e como os mapas auto-organizáveis. Além disso, para o desenvolvimento das redes, foram aplicadas outras metodologias para: a geração de conjuntos de treinamento (reamostragem bootstrap), escolha dos parâmetros de rede (otimização simplex seqüencial) e para verificar a confiabilidade e variabilidade dos resultados obtidos (estatística descritiva e testes de hipótese).

No capítulo 1 foi apresentada uma visão geral da tese sob a forma de revisão bibliográfica, materiais e métodos, objetivos e referências consultadas.

Os demais capítulos foram elaborados na forma de artigos para publicação, e serão posteriormente colocados nas normas das revistas para onde serão encaminhados.

No capítulo 2 foi realizada uma comparação entre o mapa auto-organizável e a análise de componentes principais em um banco de dados contendo sete cafés solúveis.

No capítulo 3 foi desenvolvido um classificador neural otimizado, para o mesmo banco de dados usado no capítulo 2, usando um perceptron de múltiplas camadas.

No capítulo 4 foi utilizado tanto o perceptron como o mapa auto-organizável para o reconhecimento e classificação de quarenta tipos de café solúvel de um banco de dados industrial.

CAPÍTULO I

1.1 REVISÃO BIBLIOGRÁFICA

1.1.1 Café

A partir do café torrado e moído é produzida uma bebida que apresenta aroma e sabor bastante apreciados, à ponto de ser considerada uma das bebidas mais populares do mundo (MARIA et al., 1999). Uma característica interessante da bebida do café é o fato de não possuir valor nutricional relevante, sendo consumida basicamente devido aos efeitos fisiológicos e psicológicos relacionados à presença da cafeína e, principalmente, pelo prazer e satisfação que seu aroma e sabor são capazes de proporcionar (GROSCH, 2001). Desses dois atributos fundamentais, o aroma é sem dúvida o mais complexo e interessante e por isso tem sido alvo de estudos sistemáticos. O aroma do café é formado por uma mistura extremamente complexa de inúmeros compostos voláteis que apresentam qualidades de aroma, intensidades e concentrações diferentes. Dessa maneira, a contribuição de cada um desses compostos voláteis para o aroma final do café é bem variada, podendo ainda ocorrer interações sinérgicas e antagônicas entre esses diferentes compostos (MOREIRA et al., 2000). A composição química da semente do café verde é bastante complexa sendo essa complexidade aumentada durante o processo de torrefação em elevadas temperaturas. Uma série de compostos originalmente presentes na semente serão precursores para a geração de novos compostos extremamente importantes para a formação do aroma típico da bebida obtida do grão de café torrado e moído (MELLO; TRUGO, 2003). Alguns dos principais grupos voláteis de impacto no aroma do café torrado e seus respectivos precursores são apresentados na Tabela 1. A combinação de odores entre os diferentes compostos presentes será responsável pela característica final do produto. A complexidade dos estudos sobre o papel dos compostos voláteis do café ainda é aumentada, em relação à concentração dos mesmos, tendo em vista que os teores de determinados compostos podem exercer também forte influência na característica sensorial. A determinação da concentração e dos limiares de percepção dos diferentes compostos voláteis do café tem mostrado ser uma árdua tarefa, devido à enorme complexidade da matriz. Nem sempre os resultados

encontrados na literatura são concordantes em relação aos compostos considerados de impacto aromático, como em relação às suas concentrações nas amostras (MELLO; TRUGO, 2003).

Tabela 1 – Grupos de voláteis de impacto no aroma do café torrado (MELLO; TRUGO, 2003).

Grupos de voláteis	Precursores	Exemplos
Pirrol e alquil pirróis	- Aminoácidos + carboidratos - Hidróxi-aminoácidos	N-metil pirrol
Furanonas	- Sacarose	2-metil-3-tetrahidrofuranona
Furaldeídos	- Sacarose - Arabinogalactana	2-furfural
Alquilfuranos	- Arabinogalactana - Aminoácidos + carboidratos	5-metil-2-vinilfurano
Pirazinas e Alquilpirazinas	- Hidróxi-aminoácidos	Metil-pirazina
Acilpirróis	- Aminoácidos + carboidratos	1-metil-2-acetilpirrol
Compostos carbocíclicos	- Sacarose	Cicloteno
Piridina	- Trigonelina - Hidróxi-aminoácidos	Piridina
Fenóis	- Ácido clorogênico	Fenol

A partir do café torrado, são utilizadas colunas extratoras na industrialização do café solúvel para a obtenção do extrato hidrossolúvel de café. O teor de sólidos solúveis do extrato de café é relativamente baixo, e por razões econômicas é concentrado em processo de evaporação, pela remoção da água e posterior secagem por atomização ou por liofilização. A quantidade de voláteis do café solúvel é inferior quando comparada à do café torrado, pois, além do café torrado apresentar a matriz não volátil juntamente com os compostos voláteis, o teor de material não volátil solubilizado é cerca de duas e meia vezes superior no caso do café solúvel (CLARKE, 2001). Em Viegas e Bassoli (2007) foram identificados 88 compostos importantes para a qualidade do café solúvel utilizando-se a técnica de GC-MS. Entre os componentes reconhecidos destaca-se o metional, furaneol, 2-furfuriltiol e ácido isovalérico. Destacam-se também, outras moléculas altamente voláteis e que possuem descritivos sensoriais avaliados como amanteigados, tais como 2,3-butanodiona e 2,3-pentanodiona. A degradação de Strecker leva à

formação de vários dos compostos relatados como importantes na formação do aroma, tais como 2-metil-butanal e 3-metil-butanal, que são perceptíveis sensorialmente mesmo em altas diluições. Muitas pirazinas resultantes de reações de Maillard foram identificadas neste mesmo trabalho, dentre elas pode-se citar 2,3-dimetilpirazina, 2,6-dimetil-pirazina.

No Brasil, o consumo interno evoluiu 24,8% desde 2003, chegando as 17,1 milhões de sacas de 60kg em 2007. O mercado brasileiro também representa 14% da demanda mundial, e 32% do consumo de toda a Europa. O consumo brasileiro de café verde no período de novembro de 2006 a outubro de 2007 foi de 5,53 kg/hab/ano. Esta quantidade é comparável à dos maiores consumidores do mundo: Alemanha (5,86 kg/hab/ano), Itália (5,63 kg/hab/ano) e França (5,07 kg/hab/ano). Do consumo total de 2007, cerca de 6% foram comercializados na forma de café solúvel. Entretanto, nos últimos anos a evolução de consumo do café solúvel tem sido maior do que a de café torrado e moído (ABIC, 2008).

1.1.2 Nariz Eletrônico

O olfato humano pode reconhecer e discriminar compostos voláteis com alta sensibilidade e precisão, sendo de grande importância na avaliação da qualidade de alimentos no cotidiano (SCHALLER et al., 1998). Alguns odores são detectados em concentrações de partes por trilhão e até mesmo estereoisômeros podem ser diferenciados. Uma área de cerca de 5 cm² do epitélio nasal pode conter entre 10 e 100 milhões de receptores olfativos. E, diferentemente da gustação, os odores não podem ser separados facilmente em grupos diferentes. As moléculas responsáveis pelos odores, geralmente, são leves (massa molecular menor que 300 Da), pequenas, polares e frequentemente hidrofóbicas (CRAVEN et al., 1996). Devido a tais fatos, não é surpreendente que vários esforços sejam direcionados na tentativa de mimetizar o olfato através de um instrumento (DEISINGH et al., 2004). Tais sensores poderiam ser usados para alertar os consumidores do potencial risco

de alimentos contaminados (PANIGRAHI et al., 2006) e no próprio controle de qualidade das indústrias (TRIHAAS; NIELSEN, 2005).

Nariz eletrônico é um instrumento contendo um grupo de sensores gasosos, com especificidade parcial, acoplado a um sistema apropriado de reconhecimento de padrões capaz de identificar odores simples e complexos. Um sensor químico é um dispositivo, utilizado tanto em análises quantitativas como qualitativas, que responde a um composto específico através de uma reação reversível. Os sensores são compostos por duas regiões principais: a primeira é onde ocorre a reação química que produz um sinal como mudança de cor, fluorescência, calor ou mudança da frequência de oscilação de um cristal; na segunda, chamada de transdutor, ocorre a conversão de uma forma de energia em outra. Outras partes do sensor incluem um processador eletrônico de sinal e um mostrador para a saída do resultado. Diversos tipos de transdutores podem ser utilizados tais como: eletroquímicos, piezelétricos, ópticos, sistemas térmicos e biossensores, os quais combinam um elemento biológico específico para um determinado composto de interesse. Existe uma relação entre os narizes eletrônico e humano. Pois, cada um desses é composto de três regiões principais. No humano existe um grupo de células que atuam como receptores dos odores, o bulbo olfativo e o cérebro. No nariz eletrônico o equivalente a estas regiões seria o grupo de sensores de odor, um pré-processador de sinal e o sistema de reconhecimento de padrões (DEISINGH et al., 2004). Uma estrutura geral de um nariz eletrônico pode ser observada na Figura 1.

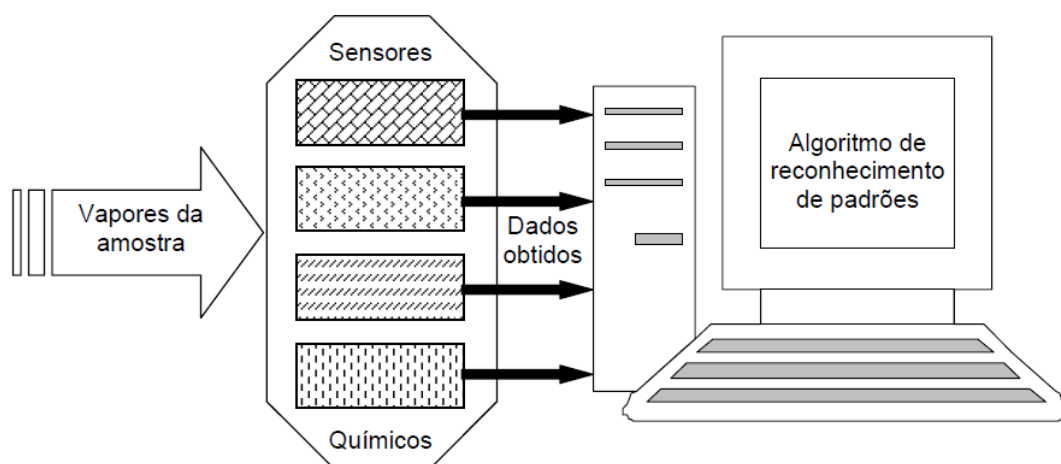


Figura 1 – Esquema ilustrativo de um nariz eletrônico.

Com o objetivo de avaliar os resultados obtidos, os dados gerados por cada sensor são processados por um algoritmo de reconhecimento de padrões. O algoritmo de reconhecimento de padrões utiliza a matriz formada pelos padrões associados a amostras previamente conhecidas para classificar um padrão desconhecido. O sucesso da classificação não depende apenas da metodologia de reconhecimento, mas também do pré-processamento dos dados, que deve selecionar e extrair as características mais importantes de cada padrão. Os métodos de reconhecimento de padrões podem ser divididos em não supervisionados e supervisionados, embora a combinação de ambos também possa ser utilizada. O método de análise de componentes principais é o método não supervisionado mais utilizado, já as redes neurais artificiais são amplamente utilizadas como metodologia supervisionada (DEISINGH et al., 2004).

Existem trabalhos recentes de revisão (SCHALLER et al., 1998; DEISINGH et al., 2004) que relatam a aplicação dos narizes eletrônicos nas mais variadas áreas da indústria de alimentos. Além disso, trabalhos atuais vêm confirmando a versatilidade do instrumento na avaliação de aromas de alimentos: monitoramento da maturação de queijo (TRIHAAS; NIELSEN, 2005); identificação de carne bovina contaminada por microrganismos (PANIGRAHI et al., 2006); reconhecimento de tipos diferentes de óleos comestíveis e méis (LINDER; PÖPPL, 2003); estudo da liberação de aromas em géis de pectina (MONGE et al., 2004).

1.1.3 Sistemas Especialistas

A idéia básica dos sistemas especialistas (SE) é transferir a experiência e perícia humana em um determinado assunto para o computador e, a partir de técnicas matemáticas extrair conclusões que auxiliam na tomada de decisão (LIAO, 2005). Apesar da existência de grandes bancos de dados com muitas informações, ainda são encontradas dificuldades na descoberta de conhecimento baseada nessas informações. Essas dificuldades podem estar relacionadas a um dos seguintes fatores: falta de conhecimento da existência de técnicas de mineração de dados; complexidade na implementação e aplicação

dessas técnicas; falta de ferramentas adequadas; alto custo das ferramentas de mineração de dados (DIAS, 2002).

As técnicas de mineração de dados são aplicadas a sistemas de descoberta de conhecimento em bancos de dados com o objetivo de extrair padrões ou regras com o objetivo de auxiliar na tomada de decisões. Essas ferramentas podem ser aplicadas a tarefas como classificação, estimativa, associação, segmentação e sumarização. Não há uma técnica que resolva todos os problemas de mineração de dados. Diferentes métodos servem para diferentes propósitos; cada método oferece suas vantagens e suas desvantagens. Sendo que, as redes neurais artificiais são indicadas para os problemas multivariados como a classificação e segmentação (DIAS, 2002). Devido a todas essas capacidades associadas aos sistemas especialistas, atualmente mais atenção e empenho, tanto por parte da academia como nas indústrias, vindo sendo dado a essas técnicas de auxílio à tomada de decisão (LIAO, 2005).

1.1.4 Redes Neurais Artificiais

A origem da teoria de redes neurais artificiais remonta aos modelos matemáticos e aos modelos da engenharia de neurônios biológicos (KOVÁCS, 1996). O neurônio é a unidade celular fundamental do cérebro que é o centro do sistema nervoso humano (MUKESH, 1996; HAYKIN, 2001). Cada neurônio (Figura 2) é formado por um corpo celular (centro dos processos metabólicos da célula nervosa) de onde se projetam extensões filamentosas, os dendritos (zonas receptivas) e o axônio (linhas de transmissão). O neurônio possui geralmente um único axônio, embora este possa apresentar algumas ramificações e estender-se por vários metros (KOVÁCS, 1996).

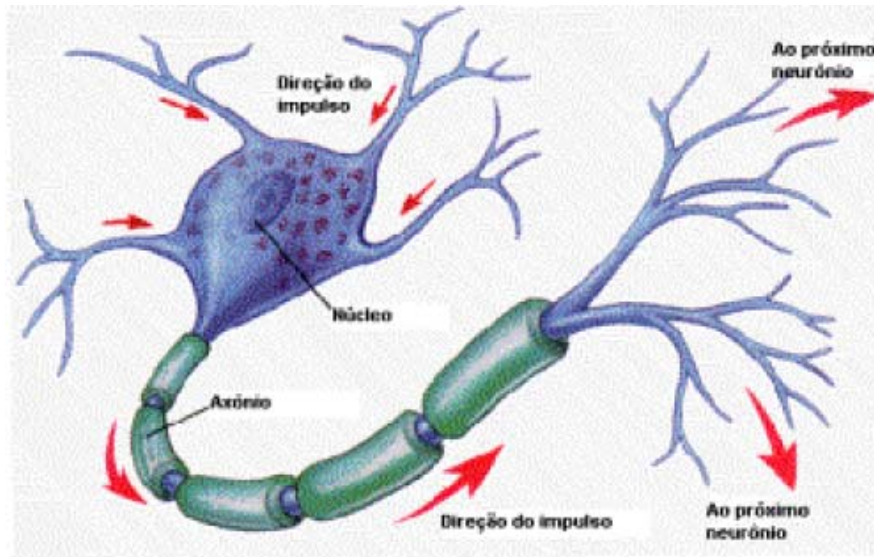


Figura 2 – Representação de um neurônio biológico.

A conexão entre o axônio de um neurônio e os dendritos de outros neurônios (cerca de 60 trilhões no córtex humano) é chamada de sinapse. E pode ser resumida como uma conexão simples que impõe ao neurônio receptivo excitação ou inibição, mas não ambas (HAYKIN, 2001). Os sinais que chegam por estes axônios são pulsos elétricos conhecidos como impulsos nervosos (ou potenciais de ação), e constituem a informação que o neurônio processará de alguma forma para produzir como saída um impulso nervoso no seu axônio (KOVÁCS, 1996). Os eventos em um circuito de silício acontecem na ordem de nanossegundos (10^{-9} s), enquanto que os eventos neurais acontecem na ordem de milissegundos (10^{-3} s). Entretanto, o cérebro compensa a taxa de operação relativamente baixa de um neurônio pelo número realmente espantoso de neurônios, com conexões maciças entre si. O resultado é que o cérebro é uma estrutura extremamente eficiente (HAYKIN, 2001).

[...] No momento, os computadores têm a vantagem da velocidade (um processador Pentium IV pode fazer 42 milhões de cálculos por segundo [expressão acrescida]), mas não mostram sinal de inteligência. Isso não surpreende, já que os nossos computadores atuais são menos complexos do que o 'cérebro' de uma minhoca – uma espécie que não se destaca por seus dotes intelectuais (HAWKING, 2002, p.165).

A capacidade de adaptação e armazenamento de conhecimento em um cérebro adulto pode ser atribuída a dois mecanismos: a criação de novas conexões sinápticas entre os neurônios e a modificação das sinapses existentes. Assim, defini-se o conhecimento: informação armazenada ou modelos utilizados por uma pessoa ou máquina para interpretar; prever e responder apropriadamente ao mundo exterior (HAYKIN, 2001).

Nos algoritmos computacionais tradicionais o conhecimento é construído de forma explícita sob a forma de regras definidas a priori. Já as redes neurais são sistemas capazes de fazer generalizações, ou seja, geram suas próprias regras, para associar as variáveis de entrada às de saída, a partir do aprendizado com dados de treinamento (LINDER; PÖPPL, 2003). Isto é possível através de uma regra (algoritmo de aprendizagem), que muda os pesos das conexões entre os neurônios em função das respostas obtidas a partir das variáveis de entrada (MUKESH, 1996; HAYKIN, 2001).

As redes neurais artificiais procuram modelar, mesmo que primitivamente, as operações lógicas pelas quais o cérebro executa as mais variadas tarefas (DEISINGH et al., 2004, LIAO, 2005). Para tanto, são criadas unidades de processamento com funções similares às dos neurônios biológicos. Cada neurônio artificial pode receber vários estímulos de entrada e gerar vários estímulos de saída. Sendo que, este fluxo de estímulos é propagado através de uma rede de conexões, tal qual no modelo biológico. A intensidade de propagação dos estímulos entre uma determinada conexão é feita através de pesos de ponderação (MUKESH, 1996; HAYKIN, 2001).

Podemos dizer que uma boa solução depende de uma boa representação do conhecimento. Assim, para as redes neurais que possuem diversas formas de representação, desde as entradas até os parâmetros internos, o desenvolvimento de uma solução satisfatória é um desafio real de projeto. A convergência do método está muito relacionada com a topologia da rede e os parâmetros adotados no algoritmo de treinamento. Dependendo da escolha, a aprendizagem pode ficar presa em um mínimo local ou convergir lentamente (LINDER; PÖPPL, 2003). Além disso, em sistemas que trabalham em tempo real é importante reduzir o tamanho da amostra de treinamento, através da retirada de dados redundantes, para aumentar a velocidade de processamento (HAMMOND et al., 2004)

O tema representação do conhecimento no interior de uma rede artificial é muito complicado. Apesar disso, existem quatro regras para a representação do conhecimento que são de senso comum (HAYKIN,2001):

Regra 1 – Entradas similares de classes similares normalmente devem produzir representações similares no interior da rede e, portanto, devem ser classificadas como pertencentes à mesma categoria.

Regra 2 – Devem ser atribuídas representações bem diferentes na rede a itens que devem ser categorizados em classes separadas.

Regra 3 – Se uma característica particular é importante, então deve haver um grande número de neurônios envolvidos na representação daquele item na rede.

Regra 4 – Informação prévia e invariâncias devem ser incorporadas no projeto de uma rede neural, simplificando com isso o projeto da rede por não ter que aprendê-las.

O conhecimento, armazenado através dos pesos sinápticos, é adquirido pela rede através de um processo de aprendizagem. Aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre. Esta definição do processo de aprendizagem implica a seguinte seqüência de eventos na rede neural: (i) estímulo por um ambiente; (ii) modificações dos parâmetros livres como resultado do estímulo e (iii) nova resposta ao ambiente devida às modificações realizadas. Um conjunto preestabelecido de regras bem-definidas para a solução de um problema de aprendizagem é denominado um algoritmo de aprendizagem. Como se pode esperar, não há um algoritmo de aprendizagem único para o projeto de redes neurais. Em vez disso, temos um conjunto de ferramentas representado por uma variedade de algoritmos de aprendizagem, cada qual oferecendo vantagens específicas. Basicamente, os algoritmos de aprendizagem diferem entre si pela forma como é formulado o ajuste de um peso sináptico de um neurônio (HAYKIN, 2001), sendo que o algoritmo de retropropagação é o mais conhecido e utilizado (DEISINGH et al., 2004).

O uso de redes neurais oferece as seguintes propriedades úteis e capacidades (HAYKIN, 2001; DEISINGH et al., 2004; PÉREZ-MAGARIÑO et al., 2004):

1) Sua natureza estruturada amplia as possibilidades de aplicação eficiente tanto em problemas não-lineares, que envolvam grande quantidade de variáveis.

2) Existe uma analogia próxima entre o mapeamento de entrada-saída realizado por uma rede neural e a inferência estatística não-paramétrica, ou seja, não são feitas suposições prévias sobre o modelo estatístico dos dados de entrada.

3) Uma rede neural treinada para operar em um problema específico pode ser facilmente retreinada para lidar com pequenas modificações nas condições operacionais do problema.

4) No contexto de classificação de padrões, uma rede neural pode ser projetada para fornecer informação não somente sobre qual padrão particular selecionar, mas também sobre a confiança na decisão tomada.

5) Cada neurônio da rede é potencialmente afetado pela atividade de todos os neurônios na rede. Conseqüentemente, as informações contextuais são tratadas naturalmente pela rede neural.

6) A natureza maciçamente paralela de uma rede neural a torna potencialmente rápida na computação de certas tarefas.

7) As redes neurais desfrutam de universalidade como processadores de informação. Esta uniformidade torna possível compartilhar teorias e algoritmos de aprendizagem em aplicações diferentes.

8) O projeto de uma rede neural é motivado pela analogia com o cérebro, que é uma prova viva de que o processamento paralelo tolerante a falhas é não somente possível fisicamente, mas também rápido e poderoso.

9) É possível trabalhar com dados incompletos e com ruídos.

Pelo fato das redes neurais artificiais serem aptas a resolver problemas de cunho geral, tais como aproximação, classificação, categorização e predição (BRAGA; CARVALHO; LUDERMIR, 2000), fez com que este conjunto de técnicas fosse aplicado numa vasta gama de áreas, com destaque para controle de processos, navegação por satélites, previsão do tempo, processamento de sinais, reconhecimento de fala, tratamento de resíduos, engenharia de cerâmicas, detecção de incêndios, mercado financeiro, reconhecimento de padrões, auxílio ao diagnóstico de doenças e previsão dos parâmetros de entrada a partir de um dado conjunto de respostas (MUKESH, 1996; AZEVEDO; BRASIL; OLIVEIRA, 2000;

HUNTER et al., 2000; HAMMOND et al., 2004; LIAO, 2005). A extrapolação de alguns destes exemplos para outras áreas do conhecimento pode ser obtida, já que alguns deles foram tratados de forma mais genérica (BRAGA; CARVALHO; LUDERMIR, 2000). Na área de alimentos as redes já foram utilizadas com sucesso na interpretação e classificação de sinais obtidos por narizes eletrônicos em óleos comestíveis e mel (LINDER; PÖPPL, 2003), carnes contaminadas por microrganismos (PANIGRAHI et al., 2006), liberação de aromas em géis de pectina (MONGE et al., 2004), monitoramento da maturação de queijo (TRIHASS; NIELSEN, 2005). E na classificação de vinhos (PÉREZ-MAGARIÑO et al., 2004), previsão do volume de pães em testes de panificação (HORIMOTO et al., 1995).

Apesar da extensa utilização, as redes neurais são muitas vezes referidas como ferramentas heurísticas do tipo “caixa preta”. Entretanto alguns trabalhos recentes (BOGER, 2003; KOLMAN; MARGALOT, 2005) têm apresentado maneiras para extrair conhecimento a partir do modelo do mapeamento entrada-saída construído por uma rede neural.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Avaliar o desempenho uma rede neural artificial para a identificação e classificação de amostras de café solúvel a partir de suas características aromáticas obtidas através de um nariz eletrônico.

1.2.2 Objetivos Específicos

1. Avaliação de métodos de normalização dos dados obtidos pelo nariz eletrônico.

2. Otimização dos parâmetros da arquitetura de rede escolhida e aplicação de técnicas de poda de rede.
3. Avaliar o procedimento de reamostragem *bootstrap* para a geração de conjuntos de treinamento.
4. Aplicação de métodos heurísticos na extração de conhecimento à rede neural artificial treinada.
5. Comparar, diferenciar e validar o *software* que será desenvolvido com o módulo de redes neurais artificiais que acompanha o pacote *Statistica 7.1* (STATSOFT, 2005).

1.3 MATERIAIS E MÉTODOS

1.3.1 Café Solúvel

Foram disponibilizados dois bancos de dados pela Companhia Iguaçu de Café Solúvel (Cornélio Procópio – Paraná – Brasil).

No primeiro banco foram usados sete tipos de café solúvel (Tabela 2), sendo três amostras experimentais produzidas na planta piloto do laboratório de pesquisa e desenvolvimento da Companhia Iguaçu de Café Solúvel. As demais amostras são cafés solúveis comerciais adquiridos no varejo do Brasil e Inglaterra (BASSOLI, 2006).

Tabela 2 – Descrição simplificada das amostras de cafés solúveis analisadas

Amostra	Origem	Classificação	Composição	Intensidade de torração
1 ^a	Brasil	Experimental	Arábica	Média
2 ^b	Brasil	Experimental	Arábica	Média
3 ^c	Brasil	Experimental	Conilon	Média
4	Brasil	Comercial	Conilon/arábica	Escura
5	Brasil	Comercial	Desconhecida	Escura ^d
6	Inglaterra	Comercial	Desconhecida	Média ^d
7	Inglaterra	Comercial	Desconhecida	Média/clara ^d

^a Café, classificado como bebida dura, proveniente da região sul de Minas Gerais.

^b Café, classificado como bebida dura, proveniente da região norte do Paraná.

^c Café proveniente do estado de Rondônia.

^d Intensidade de torração estimada por avaliação sensorial.

No segundo banco de dados foram analisados 40 cafés solúveis produzidos pela Companhia Iguaçu de Café Solúvel. Cada produto foi selecionado e identificado pela própria indústria usando uma letra maiúscula e dois algarismos arábicos. No total foram avaliados 71 lotes conforme apresentado na Tabela 3.

Tabela 3 – Identificação dos produtos e quantidade de lotes avaliados.

Produto	Lotes	Produto	Lotes	Produto	Lotes	Produto	Lotes
A01	1	A02	2	A03	2	A04	3
B01	1	B02	4	C02	1	C04	1
C05	1	C09	1	C10	1	C11	1
C13	1	D01	1	E01	1	E02	3
E03	1	E04	1	F01	1	H01	1
K01	1	K02	4	N01	4	N02	3
N03	3	N05	1	N06	1	N07	1
N08	3	N09	1	R01	3	S01	2
S02	1	S03	1	X02	1	X03	3
X04	1	X05	2	X07	3	X08	3

1.3.2 Nariz Eletrônico

Todas as análises foram realizadas nos laboratórios de pesquisa e desenvolvimento da Companhia Iguaçu de Café Solúvel. Foi utilizado o equipamento Portable Electronic Nose Type PEN2 (Airsense Analytics, Alemanha), que monitora a variação da condutância em cada um de seus sensores (Tabela 4) conforme a passagem do fluxo gasoso com a amostra (HAI; WANG, 2006; ZHANG et al., 2007). Essa variação é, então, comparada com àquela obtida pelo branco (ar ambiente) gerando, assim, um sinal adimensional.

Tabela 4 – Especificidade dos sensores utilizados no nariz eletrônico PEN2.

Sensor	Especificidade
S1	Compostos aromáticos
S2	Amplo espectro, muito sensível, reage com óxidos de nitrogênio
S3	Amônia, usado para compostos aromáticos
S4	Hidrogênio
S5	Alcanos e compostos aromáticos menos polares
S6	Metano, amplo espectro
S7	Compostos sulfurados, terpenos, limonenos e pirazinas
S8	Álcoois, amplo espectro
S9	Compostos orgânicos sulfurados
S10	Metanos alifáticos em altas concentrações (> 100 ppm)

No preparo das amostras 2,8g de café solúvel foram colocados em frascos específicos para análise de headspace. Cada amostra, de ambos os bancos de dados, foi avaliada em três repetições genuínas. Os parâmetros operacionais do equipamento foram (BASSOLI, 2006): 1,0 s de intervalo entre as amostras; 60 s de tempo de medição; 120s para limpeza ou purga dos sensores; 400 mL/min de fluxo de injeção.

1.3.3 Integração e Normalização

Após a aquisição dos dados, a área abaixo da curva gerada por cada sensor foi numericamente integrada, formando um vetor de entrada com 10 valores para cada amostra. Antes de ser alimentado na rede neural os vetores de entrada foram pré-processados para que as funções de ativação dos neurônios artificiais não sejam facilmente saturadas ou ocorra um erro de overflow (HAYKIN, 2001). Os métodos de pré-processamento usados foram: máximo e mínimo (minimax), transformação para uma escala entre 0 e 1; autoescalamento, vetor de entrada com média zero e variância unitária; vetor unitário, norma euclidiana igual a 1 (BASHEER; HAJMEER, 2000; PÉREZ-MAGARIÑO et al., 2004). Outro benefício associado à normalização é a garantia de que todas as variáveis de entrada apresentaram a mesma ordem de grandeza. Evitando assim que a rede neural

artificial atribua um importância maior a uma determinada variável devido apenas a sua magnitude (BISHOP, 1995).

1.3.4 Rede Perceptron de Múltiplas Camadas

Foi empregada uma rede neural artificial do tipo perceptron de múltiplas camadas (MLP – multi-layer perceptron) que é amplamente utilizada para a classificação de padrões em conjunto com o nariz eletrônico (DEISINGH et al., 2004). A arquitetura deste tipo de rede (Figura 3) é composta por uma camada de entrada com um neurônio para cada variável de entrada. Uma ou duas camadas intermediárias, responsáveis pela separação dos padrões através da formação de fronteiras de decisão, contendo uma quantidade de neurônios a ser definida. E uma camada de saída, que constrói combinações lineares das fronteiras de decisão formadas pelos neurônios ocultos, cuja quantidade de neurônios depende de quantos produtos serão classificados e de como serão representados (HAYKIN, 2001).

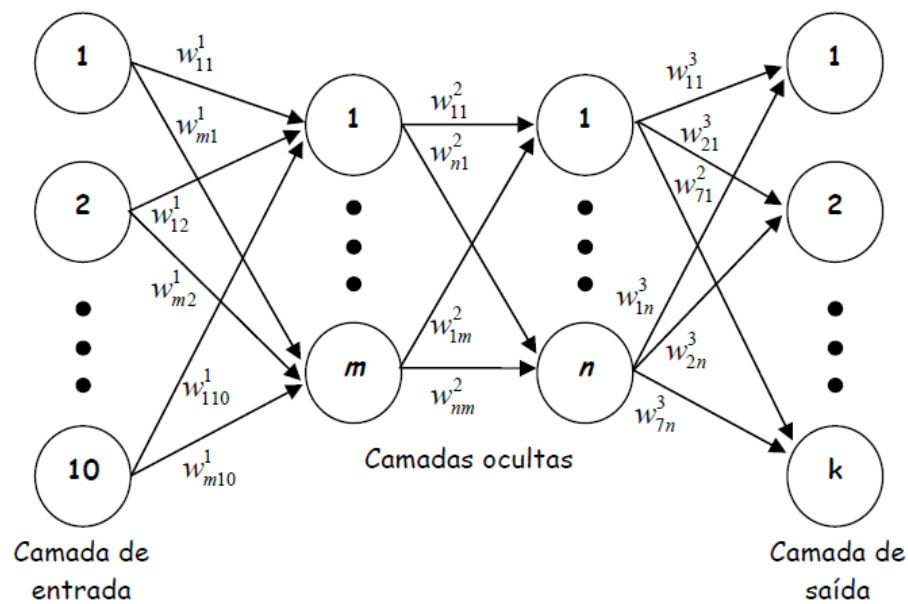


Figura 3 – Representação gráfica da rede perceptron de múltiplas camadas com duas camadas ocultas. Para cada conexão entre os neurônios existe um peso w_{ij}^l associado, sendo i o índice do neurônio de entrada do sinal, j o neurônio de saída do sinal e l camada onde esta localizado o neurônio de entrada.

Em cada um dos N^l neurônios das L camadas da rede MLP é realizado um somatório ponderado pelos pesos sinápticos dos sinais provenientes dos neurônios da camada anterior. Esta soma, chamada de campo local induzido (1), é aplicada a uma função de ativação não linear (2) que irá produzir a saída do neurônio (HAYKIN, 2001).

$$v_i^l = \sum_{j=0}^{m_i} w_{ij}^l y_j^{l-1} \quad i = 1, 2, \dots, N^l; \quad l = 1, 2, \dots, L. \quad (1)$$

$$y_i^l = \varphi(v_i^l) \quad i = 1, 2, \dots, N^l; \quad l = 1, 2, \dots, L. \quad (2)$$

Nas equações acima v_i^l é o campo local induzido do neurônio i na camada l , w_{ij}^l é peso sináptico associado ao sinal de saída y_j^{l-1} – do neurônio j da camada anterior ($l-1$), m_i é o número de entradas associadas ao neurônio i . Para $j = 0$, temos $y_j^{l-1} = 1$ e w_{i0}^l é chamado de termo de polarização, ou *bias*, que tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo. Nas redes MLP as funções de ativação mais utilizadas

(BASHEER; HAJMEER, 2000) são as sigmóides do tipo logística e tangente hiperbólica, definidas respectivamente por,

$$\varphi(v_i^l) = \frac{1}{1 + e^{-av_i^l}} \quad i = 1, 2, \dots, N^l; \quad l = 1, 2, \dots, L, \quad (3)$$

$$\varphi(v_i^l) = a \tanh(bv_i^l) \quad i = 1, 2, \dots, N^l; \quad l = 1, 2, \dots, L. \quad (4)$$

Para a função logística (3) o símbolo a representa a inclinação e pode ser considerado como um parâmetro a ser otimizado pelo projetista da rede, embora em muitos casos é usado um valor unitário para este parâmetro (HUANG et al., 2007). Na função tangente hiperbólica (4) as constantes $a = 1,7159$ e $b = 0,6667$ já foram otimizadas para melhorar as propriedades desejadas desta função de ativação (HAYKIN, 2001).

Nos algoritmos computacionais tradicionais, o conhecimento é construído de forma explícita sob a forma de regras definidas a priori. Já as redes neurais são sistemas capazes de fazer generalizações, ou seja, geram suas próprias regras, para associar as variáveis de entrada às de saída, a partir do aprendizado com dados de treinamento (LINDER; PÖPPL, 2003). Isto é possível através de uma regra, algoritmo de aprendizagem, que muda os pesos das conexões entre os neurônios em função das respostas obtidas a partir das variáveis de entrada. Para a rede construída foi utilizado o algoritmo de retropropagação, que consiste de dois passos através das diferentes camadas da rede: um passo a frente, a propagação, e um passo atrás, a retropropagação. No passo a frente, um padrão de atividade (vetor de entrada) é aplicado aos neurônios de entrada da rede e seu efeito se propaga através da rede, camada por camada. Finalmente, um conjunto de saídas é produzido como resposta real da rede. Durante o passo de propagação, os pesos sinápticos da rede são todos fixos. No passo atrás, por outro lado, os pesos são ajustados de acordo com uma regra de correção de erro (5). Especificamente, a resposta real da rede é subtraída de uma resposta desejada (alvo) para produzir um sinal de erro (7). Este é então propagado para trás através da rede, contra a direção das conexões sinápticas (HAYKIN, 2001).

$$w_{ij}^l(n+1) = w_{ij}^l(n) + \alpha \Delta w_{ij}^l(n-1) + \eta \delta_i^l y_j^{l-1}$$

$$i = 1, 2, \dots, N^l; l = 1, 2, \dots, L; j = 1, 2, \dots, m_i; n = 0, 1, 2, \dots \quad (5)$$

$$\text{Sendo, } \delta_i^l = \begin{cases} e_i \phi_i'(v_i) & \text{para o neurônio } i \text{ da camada de saída} \\ \phi_i'(v_i) \sum_j \delta_j^{l+1} w_{ij}^{l+1}(n) & \text{para o neurônio } i \text{ da camada oculta } l \end{cases}$$

$$i = 1, 2, \dots, N^l; l = 1, 2, \dots, L; j = 1, 2, \dots, m_i. \quad (6)$$

$$e_i = d_i - o_i \quad i = 1, 2, \dots, N^L \quad (7)$$

Nas equações acima n refere-se a uma iteração do algoritmo de retropropagação. Assim, $w_{ij}^l(n+1)$ é o novo peso sináptico (5) que será utilizado na próxima iteração e $\Delta w_{ij}^l(n-1)$ é a correção de pesos realizada na iteração passada. Nesta mesma equação α é chamada de constante de momento e η é a taxa de aprendizagem. Os valores destes parâmetros devem ser otimizados para cada problema e em geral a escolha é feita através de tentativa e erro (BRAGA *et al.*, 2000). O gradiente local (δ_i^{l-1}) de cada neurônio é calculado diferentemente para a camada de saída e ocultas. Para os neurônios da última camada (6) o gradiente local é igual ao produto da derivada da função de ativação em relação ao campo local induzido, ϕ_i' em v_i , pelo sinal de erro (e_i), que é igual a resposta desejada (d_i) menos àquela produzida pela rede (o_i). Quando o neurônio está localizado em uma camada oculta da rede, não existe uma resposta desejada. Correspondentemente, o sinal de erro para um neurônio oculto deve ser determinado, recursivamente, em termos dos sinais de erro de todos os neurônios aos quais está diretamente conectado. Ou seja, δ_i^{l-1} é igual ao produto da derivada da função de ativação pela soma ponderada dos gradientes calculados para os neurônios da próxima camada oculta ou camada de saída que estão diretamente conectados.

Uma apresentação completa do conjunto de treinamento é denominada época e o processo de aprendizagem é mantido até os pesos sinápticos e os níveis de *bias* se estabilizarem e o erro quadrático médio convergir para um valor mínimo (AZEVEDO *et al.*, 2000).

$$EQM = \frac{1}{2N} \sum_{n=1}^N \sum_{i=1}^C e_i^2(n) \quad (8)$$

Na equação acima, N é igual ao tamanho do conjunto de treinamento e C igual ao número de neurônios na camada de saída. Para a rede empregada, foi utilizado o modo seqüencial de treinamento, uma vez que, os pesos são atualizados a cada novo exemplo apresentado a rede. No modo em lote, o ajuste é realizado após cada época. O modo seqüencial é preferível em relação ao modo lote, porque os exemplos de treinamento são apresentados à rede na forma aleatória tornando estocástica a busca no espaço de pesos. Por sua vez, isto torna menos provável que o algoritmo de retropropagação fique preso em um mínimo local. Além disso, o modo seqüencial é de simples implementação e fornece soluções efetivas a problemas complexos e de grande porte (HAYKIN, 2001).

Tendo como objetivo uma boa generalização, foi utilizado o método de treinamento com parada antecipada para evitar que a rede acabe sendo excessivamente ajustada aos dados de treinamento. É possível identificar o início do excesso de treinamento através do uso de amostras de validação, onde além dos dados de treinamento é usado um conjunto de teste que após cada época é submetido à rede para verificar a capacidade de generalização desta com dados não utilizados para o ajuste dos pesos. Tipicamente, o modelo não funciona tão bem sobre o conjunto de teste quanto sobre o de treinamento, em relação ao qual foi baseado. A curva do erro quadrático médio, também chamada de curva de aprendizagem, para o treinamento decresce monotonamente para um número crescente de épocas. Já para o conjunto de teste, esta curva decresce para um mínimo e então começa a crescer conforme o treinamento continua, ou seja, o que a rede aprende após este ponto é essencialmente o ruído contido nos dados de treinamento. Nesta heurística, o ponto mínimo na curva de aprendizagem de teste é usado como o critério de parada ou armazenamento do conjunto de pesos (HAYKIN, 2001).

1.3.5 Mapa Auto-Organizável

Neste trabalho foi implementado o algoritmo SOM (em inglês self-organizing map) bidimensional proposto e discutido em maiores detalhes por Haykin (2001), que serviu de suporte metodológico. O principal objetivo do mapa auto-organizável é transformar um padrão de sinal incidente de dimensão arbitrária em um mapa discreto bidimensional e realizar esta transformação adaptativamente de uma maneira topologicamente ordenada. No modelo de Kohonen (Figura 4) cada neurônio da grade está totalmente conectado com todos os nós de fonte da camada de entrada. Esta grade representa uma estrutura alimentada adiante com uma única camada computacional consistindo de neurônios arranjados em linhas e colunas. Os neurônios apresentam a mesma quantidade de pesos, que por sua vez é igual ao número de variáveis analisadas. Cada variável pode ser representada por um plano de pesos (Figura 4), que após a fase de treinamento pode ser utilizado na tentativa de extrair regras sobre os grupos formados. O algoritmo responsável pela formação do SOM começa, primeiramente, inicializando os pesos sinápticos da grade. Isto é feito atribuindo-lhes valores pequenos (entre 0 e 1) tomados de um gerador de números aleatórios. Fazendo dessa forma, nenhuma organização prévia é imposta ao mapa de características. Após a inicialização, há três processos essenciais: competição, cooperação e adaptação sináptica.

No processo competitivo, para cada padrão de entrada, os neurônios da grade calculam seus respectivos valores de uma função discriminante (distância euclidiana). O neurônio que apresentar a menor distância euclidiana é chamado de neurônio vencedor, ou centróide, para o padrão de entrada considerado. O neurônio vencedor localiza o centro de uma função que representa a vizinhança topológica de neurônios cooperativos. A razão do processo cooperativo provém da evidência neurobiológica de interação lateral entre os neurônios.

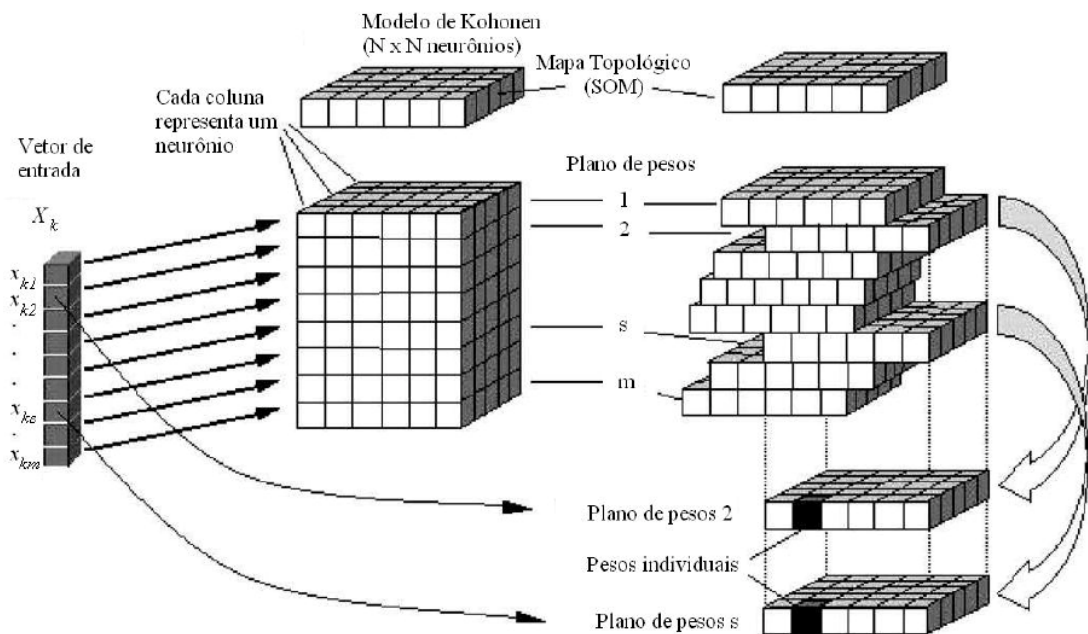


Figura 4 – Mapa auto-organizável segundo o modelo de Kohonen. Na figura k representa o número de padrões de entrada, m é a quantidade variáveis de entrada e N é o número de neurônios em cada dimensão (MARINI; ZUPAN; MAGRÌ, 2005).

A função escolhida para representar a vizinhança topológica deve satisfazer as seguintes propriedades: (i) ser simétrica em relação ao ponto de máximo definido pela posição do neurônio vencedor; (ii) sua amplitude deve ser monotonamente decrescente com o aumento da distância em relação ao neurônio vencedor. Uma escolha típica que satisfaz estas exigências é a função gaussiana apresentada na equação (9).

$$h_{j,i} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right) \quad (9)$$

Na equação acima o parâmetro σ é o raio efetivo da vizinhança topológica (Figura 5a). Ele mede o grau com o qual os neurônios vizinhos ao vencedor participam do processo de aprendizagem. Em um sentido qualitativo, a vizinhança topológica gaussiana é mais biologicamente apropriada que uma vizinhança quadrada (Figura 5b). Seu uso também faz com que o algoritmo SOM convirja mais rapidamente que com uma vizinhança topológica retangular. Ainda na equação (9), $d_{j,i}$ representa a distância lateral entre o neurônio vencedor i e o neurônio excitado j . No caso de uma grade bidimensional esta distância é definida

através da distância euclidiana entre as posições de i e j , ambas determinadas no espaço de saída discreto, definidas pelos vetores r_i e r_j respectivamente.

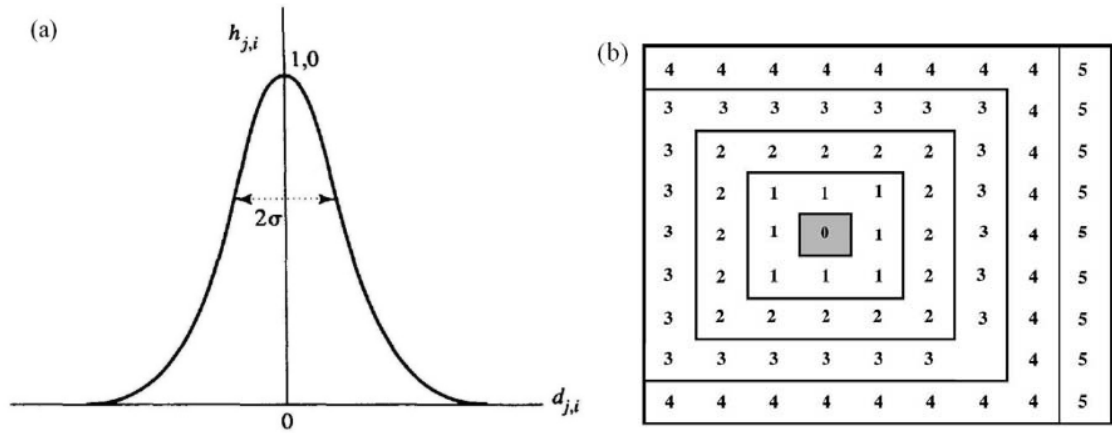


Figura 5 – Representação de vizinhança topológica: (a) gráfico unidimensional da função gaussiana; (b) representação de uma vizinhança quadrada ilustrando a distância de cada neurônio (raio) para aquele marcado em cinza.

Um ponto importante do algoritmo SOM é a redução do tamanho da vizinhança ao longo do treinamento. No início a função deve ser suficientemente larga para cobrir todo o mapa, porém a medida que o treinamento prossegue a largura deve ser diminuída para garantir a formação de regiões topológicas especializadas. Uma maneira de implementar a redução é através de um decaimento exponencial do tipo representado pela equação (10).

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_1}\right), \quad n = 0, 1, 2, \dots \quad (10)$$

Onde σ_0 é o raio efetivo de inicialização do algoritmo e τ_1 é uma constante de tempo, sendo recomendado o uso de $\tau_1 = 1000 / \log \sigma_0$. O tempo discreto n representa a época de treinamento, ou seja, uma apresentação completa de todos os padrões de entrada.

Durante o processo adaptativo é necessário que o vetor de peso sináptico w_j do neurônio j da grade se modifique em relação ao vetor de entrada x . O processo de correção ocorre através de uma modificação no postulado de

aprendizagem de Hebb. Matematicamente e usando o formalismo de tempo discreto, o processo adaptativo é realizado por:

$$\mathbf{w}_j(n+1) = \mathbf{w}_j + \eta(n)h_{j,i}(\mathbf{x})(\mathbf{x} - \mathbf{w}_j(n)) \quad n = 0, 1, 2, \dots \quad (11)$$

onde $\eta(n)$ é a taxa de aprendizagem que, como representado, deve ser variável e decrescente ao longo das n épocas de treinamento. O processo de adaptação pode ser dividido em duas etapas: ordenação e convergência. Na etapa de ordenação a taxa de aprendizagem deve ser maior para garantir que o mapa topológico, inicialmente desordenado, sofra maiores ajustes em seus pesos sinápticos. Na fase de convergência é realizada uma sintonia fina do mapa usando uma taxa de aprendizagem menor. O decréscimo da taxa de aprendizagem também pode ser realizado através de um decaimento exponencial, como apresentado na equação (12). Nesta equação, η_0 é a taxa de aprendizagem inicial e τ_2 é uma outra constante de tempo, sendo os valores recomendados, respectivamente, 0,1 e 1000.

$$\eta(n) = \eta_0 \exp\left(-\frac{n}{\tau_2}\right), \quad n = 0, 1, 2, \dots \quad (12)$$

Após o treinamento, o resultado pode ser visualizado através de um mapa topológico e mapas de peso. No primeiro, cada padrão de entrada é posicionado sobre o mapa de características de acordo com a posição do neurônio vencedor da amostra considerada (Figura 4). A função do mapa topológico é facilitar a visualização de agrupamentos e também a relação de vizinhança entre os grupos formados. Grupos próximos partilham de alguma similaridade, assim como, quanto maior a distância maior será a diferença de comportamento. O mapa de pesos (ZUPAN; NOVIČ; RUISÁNCHEZ, 1997) é um gráfico de contorno para cada plano de pesos (Figura 4). Em conjunto com o mapa topológico é possível extrair regras de comportamento para cada grupo formado e inferir a influência de cada variável sobre o resultado obtido.

1.3.6 Bootstrap e Amostragem

O bootstrap (reamostragem) consiste na obtenção de medidas de dispersão, (como a média, variância ou desvio-padrão) através da aplicação de métodos tradicionais, de inferência e estimação, a diversos conjuntos de dados obtidos a partir de um único conjunto original mediante a geração de dados aleatórios com a mesma média e desvio padrão dos originais (FERREYRA RAMÍREZ, 2005).

O primeiro banco de dados era composto por sete tipos de café analisados em triplicatas genuínas, totalizando 21 amostras, que é uma quantidade pequena para ser subdividida em um conjunto de treinamento e outro de teste. Usando a metodologia de bootstrap foram criadas mais 63 amostras a partir do conjunto original, sendo estas utilizadas para o treinamento e as 21 originais para validação/teste das redes perceptron de múltiplas camadas.

No banco de dados industrial 71 lotes de café solúvel em triplicatas genuínas totalizando 213 amostras. Estas foram subdivididas em dois conjuntos: um de validação composto por 166 amostras e outro de teste com 47. O conjunto de teste foi formado através do sorteio de uma das repetições de cada lote para os produtos com mais de um lote. Pois, caso fossem retiradas amostras dos cafés com apenas um lote restariam apenas duas para empregar a metodologia bootstrap. Usando, então, o bootstrap foram criadas mais 664 amostras a partir do conjunto de validação. Em resumo, para cada rede foram usadas 75,7% das amostras para treinamento (geradas aleatoriamente), 18,9% para validação e 5,4% para teste (amostras originais).

1.3.7 Média de Ensemble

Uma maneira de melhorar a capacidade de generalização das redes neurais é a combinação das respostas de várias RNAs, também chamada de máquinas de comitê (HAYKIN 2001), as quais foram treinadas de maneira semelhante de modo que, o resultado final seja melhor do que aquele obtido ao ser

utilizado apenas uma única RNA. Uma maneira fácil de implementar o comitê de máquinas é a média de ensemble, que consiste em combinar as respostas das redes treinadas através de uma média. O uso de um ensemble diminui a probabilidade de erro final, pois será coberto um espaço maior de hipóteses dentro das possíveis soluções para o problema. A desvantagem deste método reside no aumento do tempo computacional (BISHOP, 1995).

1.3.8 Otimização dos Parâmetros da Rede MLP

Nas redes MLP alguns parâmetros de arquitetura e do algoritmo de aprendizagem foram otimizados através do método simplex seqüencial (BONA et al., 2000). Os parâmetros escolhidos para otimização foram: a quantidade de neurônios na primeira e segunda camada ocultas, taxa de aprendizagem, constante de momento e a inclinação da função de ativação logística. A otimização foi realizada visando maximizar a porcentagem de classificação correta e minimizar o erro quadrático médio para o conjunto de validação. Além disso, estas metas deveriam ser alcançadas com a menor rede possível, entende-se por isso aquela que apresentar a menor dimensão para o vetor de pesos (HAYKIN, 2001). Para combinar estas três exigências em uma única resposta global, que será maximizada pelo método simplex seqüencial, foram utilizadas as funções de desejabilidade (DERRINGER; SUICH, 1980). Maiores informações sobre a metodologia de otimização estão disponíveis no apêndice.

1.3.9 Implementação Computacional

A integração numérica dos dados obtidos pelo nariz eletrônico, o pré-processamento, a reamostragem bootstrap, as redes MLP os SOM e o procedimento de otimização usando o simplex seqüencial acoplado às funções de desejabilidade foram desenvolvidos em linguagem FORTRAN 90 para Windows. As estatísticas descritivas, ANOVA, comparação de médias (teste Tukey) e análise de

componentes principais foram realizadas no software Statistica 7.1. Por uma questão de qualidade de apresentação, todos os gráficos dos resultados obtidos pelo programa desenvolvido também foram gerados no software Statistica 7.1.

1.4 REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE CAFÉ. Disponível em: <www.abic.com.br>. Acesso em: 10 fev. 2008.

AZEVEDO, F. M.; BRASIL, L. M.; OLIVEIRA, R. C. L. **Redes Neurais com Aplicações em Controle e em Sistemas Especialistas**. Florianópolis: Bookstore, 2000, 401p.

BASSOLI, D. G. **Impacto aromático dos componentes voláteis do café solúvel: uma abordagem analítica e sensorial**. 2006. 214 f. Tese (Doutorado em Ciência de Alimentos), Universidade Estadual de Londrina, Londrina, 2006.

BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of Microbiological Methods**, v. 43, p. 3-31, 2000.

BISHOP, C. M. **Neural Networks for Pattern Recognition**. Oxford: Oxford University, 1995. 482p.

BODT, E.; COTTRELL, M.; VERLEYSEN, M. Statistical tools to assess the reliability of self-organizing maps. **Neural Networks**, v. 15, p. 967-978, 2002.

BOGER, Z. Selection of quasi-optimal inputs in chemometrics modeling by artificial neural network analysis. **Analytica Chimica Acta**, v. 490, p. 31-40, 2003.

BOMIO, M. Neural networks and the future of sensory evaluation. **Food Technology**, v. 52, n. 8, p. 62-63, 1998.

BONA, E. et al. Aplicativo para otimização empregando o método simplex seqüencial. **Acta Scientiarum**, v. 22, p. 1201-1206, 2000.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC, 2000. 262p.

CHANDRARATNE, M. R.; KULASIRI, D.; SAMARASINGHE, S. Classification of lamb carcass using machine vision: comparison of statistical and neural network analyses. **Journal of Food Engineering**, v. 82, p. 26-34, 2007.

CLARKE, R. J. Technology III: Instant coffee. In: CLARKE, R. J.; VITZTHUM, O. G. (ed.) **Coffe: Recent Developments**. London: Blackwell Science, 2001. p.125-139.

CRAVEN, M. A.; GARDNER, J. W.; BARTLETT, P. N. Eletronic noses: development and future prospects. **Trends in Analytical Chemistry**, v. 15, n. 9, 1996.

CURRY, B.; MORGAN, P. H. Model selection in Neural Networks: some difficulties. **European Journal of Operational Research**, v. 170, p. 567-577, 2006.

DEISINGH, A. K.; STONE, D. C.; THOMPSON, M. Applications of eletronic noses and tongues in food analysis. **International Journal of Food Science and Technology**, v. 39, p. 587-604, 2004.

DELLA LUCIA, S. M.; MINIM, L. A. **Redes neurais artificiais: fundamentos e aplicações**. [s.l: s.n], 2006. p.173-194.

DERRINGER, G.; SUICH, R. Simultaneous optimization of several response variables. **J. of Quality Technology**, v. 12, p. 214-219, 1980.

DIAS, M. M. Parâmetros na escolha de técnicas e ferramentas de mineração de dados. **Acta Scientiarum**, v. 24, n. 6, p. 1715-1725, 2002.

DÍAZ, C. et al. Application of multivariate analysis and artificial neural networks for the differentiation of red wines from the Canary Islands according to the Island of origin. **Journal of Agricultural and Food Chemistry**, v. 51, p. 4303-4307, 2003.

DUTCOSKY, S. D. **Análise Sensorial de Alimentos**. 2. ed., Curitiba: Champagnat, 2007, 239p.

FARAH, A. et al. Correlation between cup quality and chemical attributes of Brazilian coffee. **Food Chemistry**, v. 98, p. 373-380, 2006.

FERREYRA RAMÍREZ, E. F. **Modelo conexionista para avaliação de propostas para aquisição de equipamentos médico-hospitalares**. 2005. Tese (Doutorado) – Universidade de Campinas, Campinas, 2005.

GROSCH, W. Chemistry III: Volatile Compounds. In: CLARKE, R. J., VITZTHUM, O. G. (ed.) **Coffe: recent developments**. London: Blackwell Science, 2001. p.68- 89.

HAI, Z.; WANG, J. Electronic nose and data analysis for detection of maize oil adulteration in sesame oil. **Sens. and Actuators B**, v. 119, p. 449-455, 2006.

HAIR JR., J. F. et al. Análise Fatorial. In: _____. **Análise Multivariada de Dados**, 5. ed. Porto Alegre: Bookman, 2005. p.89-127.

HAMMOND, M. H. et al. Training set optimization methods for a probabilistic neural network. **Chemometrics and intelligent laboratory systems**, v. 71, p. 73-78, 2004.

HAYKIN, S. **Redes Neurais: princípio e práticas**. 2. ed. Porto Alegre: Bookman, 2001, 900p.

HAWKING, S. **O Universo numa Casca de Noz**. 5. ed. São Paulo: ARX, 2002, 215p.

HORIMOTO, Y. et al. Neural networks VS principal component regression for prediction of wheat flour loaf volume in baking tests. **Journal of Food Science**, v. 60, n. 3, p. 429-433, 1995.

HUANG, Y.; KANGAS, L. J.; RASCO, B. A. Applications of artificial neural networks (ANNs) in food science. **Critical Reviews in Food Science and Nutrition**, v. 47, p. 113-126, 2007.

HUNTER, A. et al. Application of neural networks and sensitivity analysis to improved prediction of trauma survival. **Computer Methods and Programs in Biomedicine**, v. 62, p. 11-19, 2000.

KOLMAN, E.; MARGALOT, M. Are artificial neural networks white boxes? **IEEE Transactions on neural networks**, v. 16, n. 4, p. 844-852, 2005.

KOVÁCS, Z. L. **Redes neurais artificiais: fundamentos e aplicações**. São Paulo: Acadêmica, 1996. 163p.

LIAO, S. Expert system methodologies and applications: a decade review from 1995 to 2004. **Expert Systems with Applications**, v. 28, p. 93-109, 2005.

LINDER, R.; PÖPPL, S. J. A new neural network approach classifies olfactory signals with high accuracy. **Food Quality and Preference**, v. 14, p. 435-440, 2003.

LOONEY, C. G. Advances in feedforward neural networks: demystifying knowledge acquiring black boxes. **IEEE Transactions on Knowledge and Data Engineering**, v. 8, n. 2, p. 211-226, 1996.

LUDERMIR, T. B.; YAMAZAKI, A.; ZANCHETIN, C. An optimization methodology for neural network weights and architectures. **IEEE Transactions on neural networks**, v. 17, n. 6, p. 1452-1459, 2006.

MARIA, C. A. B.; MOREIRA, R. F. A.; TRUGO, L. C. Componentes voláteis do café torrado. Parte I: Componentes heterocíclicos. **Química Nova**, v. 22, n. 2, p. 209-217, 1999.

MARINI, F. et al. Use of different artificial neural networks to resolve binary blends of monocultivar Italian olive oils. **Analytica Chimica Acta**, v. 599, p. 232-240, 2007.

MELLO, A. A.; TRUGO, L. C. Tipificação odorífera de compostos voláteis do café. In: FRANCO, M. R. B. (Ed.) **Aroma e Sabor de Alimentos**. São Paulo: Livraria Varela, 2003. p.169-175.

MELSSSEN, W.; WEHRENS, R.; BUYDENS, L. Supervised Kohonen networks for classification problems. **Chemometrics and intelligent laboratory systems**, v. 83, p. 99-113, 2006.

MONGE, M. E. et al. Detection of flavour release from pectin gels using electronic noses. **Sensors and Actuators B**, v. 101, p. 28-38, 2004.

MOREIRA, R. F. A.; TRUGO, L. C.; MARIA, C. A. B. Componentes voláteis do café torrado. Parte II. Compostos alifáticos, alicíclicos e aromáticos. **Química Nova**, v. 23, n. 2, p. 195-203, 2000.

MUKESH, D. Applications of Neural Computing for Process Chemists: I. Introduction to Neural Network. **Journal of Chemical Education**, v. 73, n. 5, p. 431-433, 1996.

NI, H.; GUNASEKARAN, S. Food quality prediction with neural networks. **Food Technology**, v. 52, n. 10, p. 60-65, 1998.

PANIGRAHI, S. et al. Neural-network-integrated electronic nose system for identification of spoiled beef. **LWT**, v. 39, p. 135-145, 2006.

PARDO, M.; SBERVEGLIERI, G. Coffee analysis with an electronic nose. **IEEE Transactions on Instrumentation and Measurement**, v. 51, n. 6, p. 1334-1339, 2002.

PÉREZ-MAGARIÑO, S. et al. Comparative study of artificial neural network and multivariate methods to classify Spanish DO rose wines. **Talanta**, v. 62, p. 983-990, 2004.

SABLANI, S. S. Status of observational models used in design and control of products and processes. **Comprehensive Reviews in Food Science and Food Safety**, v. 7, p. 130-136, 2008.

SCHALLER, E.; BOSSET, J. O.; ESCHER, F. 'Electronic noses' and their application to food. **LWT**, v. 31, p. 305-316, 1998.

TIGRINE-KORDJANI, N. et al. Relative characterization of rosemary samples according to their geographical origins using microwave-accelerated distillation, solid-phase microextraction and Kohonen self-organizing maps. **Analytical and Bioanalytical Chemistry**, v. 389, p. 631-641, 2007.

TRIHASS, J.; NIELSEN, P. V. Electronic nose technology in quality assessment: Monitorin the ripening process of Danish Blue Cheese. **Journal of Food Science**, v. 70, n. 1, 2005.

VIEGAS, M. C.; BASSOLI, D. G. Utilização do índice de retenção linear para caracterização de compostos voláteis em café solúvel utilizando GC-MS e coluna HP-INNOWAX. **Química Nova**, v. 30, n. 8, 2007.

ZHANG, H. et al. Optimization of sensor array and detection of stored duration of wheat by electronic nose. **J. of Food Engineering**, v. 82, p. 403-408, 2007.

ZUPAN, J.; NOVIČ, M.; RUISÁNCHEZ, I. Kohonen and counterpropagation artificial neural networks in analytical chemistry. **Chemometrics and Intelligent Laboratory Systems**, v. 38, p. 1-23, 1997.

WARNER, B.; MISRA, M. Understanding neural networks as statistical tools. **The American Statistician**, v. 50, n. 4, p. 284-293, 1996.

WINDEATT, T. Accuracy/Diversity and Ensemble MLP Classifier Design. **IEEE Transactions on Neural Networks**, v. 17, n. 5, p. 1194-1211, 2006.

CAPÍTULO II

2 EMPREGO DE MAPAS AUTO-ORGANIZÁVEIS NO RECONHECIMENTO DE PADRÕES AROMÁTICOS DE CAFÉ SOLÚVEL

Este artigo será submetido para Ciência e Tecnologia de Alimentos
(As normas para publicação neste periódico estão contidas no Anexo A)

Resumo

O nariz eletrônico é um instrumento que tem sido muito utilizado na análise de voláteis alimentares. Entretanto, é necessário que ao equipamento seja acoplada uma interface multivariada capaz de reconhecer padrões aromáticos. O mapa auto-organizável, por ser uma técnica robusta e capaz de trabalhar com relações não lineares, vem sendo sugerido como uma alternativa à análise de componentes principais. Neste trabalho foi empregado o algoritmo SOM (self-organizing map) para a análise de café solúvel com base nos perfis aromáticos obtidos por um nariz eletrônico. Através de testes estatísticos clássicos, foi avaliada a confiabilidade dos resultados que também foram comparados aos obtidos pela análise de componentes principais. O mapa auto-organizável empregado identificou com precisão estatística (teste de hipótese no nível bilateral de 5%) todos os cafés analisados. Os grupos formados e as relações entre eles foram muito semelhantes às obtidas pela análise de componentes principais. Além disso, a análise dos pesos da rede treinada permitiu uma inspeção gráfica da influência das variáveis de entrada sobre os grupos identificados. Os resultados obtidos confirmam a potencialidade da metodologia para a análise multivariada de dados obtidos por um nariz eletrônico. E, atestam o emprego do mapa auto-organizável como uma metodologia para o controle da qualidade aromática de café solúvel.

Palavras-chave: Redes neurais artificiais. Aprendizagem não supervisionada. Nariz eletrônico.

2.1 INTRODUÇÃO

O aroma é um dos atributos mais importantes em alimentos, principalmente, em bebidas como o café que possui um odor extremamente complexo e característico (MELLO; TRUGO, 2003). Nas últimas décadas, o nariz eletrônico tem sido usado com frequência na análise de voláteis em alimentos (SCHALLER; BOSSET; ESCHER, 1998; DEISINGH; STONE; THOMPSON, 2004) e até mesmo em café torrado e moído (PARDO; SBERVEGLIERI, 2002). Cabe ao

equipamento emular o funcionamento de um nariz humano, porém, é necessário também um sistema multivariado e “inteligente” capaz de desempenhar um papel semelhante ao córtex cerebral no reconhecimento de padrões aromáticos (CRAVEN; GARDNER; BARTLETT, 1996). Um dos métodos multivariados mais tradicionais é a análise de componentes principais (ACP). Nesta técnica os dados formam vetores de dimensão elevada são reduzidos a vetores ortogonais e de dimensão menor em um sistema de coordenadas lineares com o objetivo de facilitar a visualização de tendências ou agrupamentos (HAIR et al., 2005). Entretanto, algumas desvantagens do método são bem conhecidas (MELSSSEN; WEHRENS; BUYDENS, 2006). Primeiramente, considera-se que os dados podem ser descritos por combinações lineares. Conseqüentemente, sistemas não-lineares não serão bem representados. Um segundo ponto crítico na ACP é a qualidade do resultado que pode ser muito influenciada por amostras discrepantes. Somados a estes dois fatos, existe a possibilidade de após as transformações, o número de componentes significativas ainda ser elevado dificultando a extração de conhecimento útil dos dados. As redes neurais artificiais (RNA) são um conjunto de técnicas de mapeamento não-linear, robustas e tolerantes a discrepâncias pontuais (HUANG; KANGAS; RASCO, 2007). Dentre os vários tipos de RNA os mapas auto-organizáveis são amplamente recomendados para a mineração de dados (LU; SETIONO; LIU, 1996). Em um mapa auto-organizável, também muito conhecido como rede SOM (self-organizing map) ou mapa de características, os neurônios artificiais estão colocados em nós de uma grade que é normalmente uni ou bidimensional. Estes neurônios se tornam seletivamente sintonizados a vários padrões de entrada através de um processo de aprendizagem competitiva e não supervisionada. Um SOM é, portanto, caracterizado pela formação de um mapa topológico dos padrões de entrada no qual as localizações espaciais dos neurônios na grade são indicativas das características estatísticas intrínsecas contidas nos padrões de entrada. Por apresentar caráter inerentemente não-linear, um SOM pode ser visto como uma generalização não-linear da análise de componentes principais com a vantagem de ser uma abordagem mais simples do ponto de vista matemático (ZUPAN; NOVIČ; RUISÁNCHEZ, 1997; HAYKIN, 2001). As redes SOM foram desenvolvidas por Teuv Kohonen na década de 80 e apresentam forte inspiração neurofisiológica. Elas tem base no mapa topológico presente no córtex cerebral. Essas redes são úteis principalmente em reconhecimento de padrões, ou seja, quando não existe conhecimento sobre as

classes a que devem pertencer as amostras avaliadas (AZEVEDO; BRASIL; OLIVEIRA, 2000). Além dessa tarefa, o SOM tem sido utilizados em: problemas de caixeiro viajante, mineração de dados, diagnóstico financeiro, controle de processos químicos e controle de qualidade de água (BRAGA; CARVALHO; LUDERMIR, 2000). Na indústria de alimentos os modelos computacionais “inteligentes” tornaram-se muito importantes e têm sido usados em todas as etapas desde a produção até a segmentação de mercado (GOYACHE et al., 2001; CORNEY, 2002). Neste contexto foram encontradas várias aplicações para o SOM: estudo da maturação do queijo Argentino Port Salut (VERDINI et al., 2007); caracterização de variedades de morango (BOISHEBERT; GIRAUDEL; MONTURY, 2006); diferenciação de vinho tinto de acordo com a origem (DÍAZ et al., 2003); monitoramento do processo de fermentação com leveduras (KOLEHMAINEN; RÖNKKÖ; RAATIKAINEN, 2003); identificação de misturas binárias de azeite de oliva Italianos (MARINI et al., 2007); caracterização de amostras de alecrim de acordo com a origem geográfica (TIGRINE-KORDJANI et al., 2007); classificação de variedades de arroz (MARINI, ZUPAN; MAGRÌ, 2005); uso de um nariz eletrônico na identificação de chás que sofreram diferentes processamentos (DUTTA et al., 2003). Entretanto, os resultados obtidos através de redes neurais sempre estão sujeitos a variabilidade devido à sensibilidade às condições iniciais e convergência a um mínimo local pelo algoritmo de aprendizagem. Para um uso mais racional do SOM, mesmo sendo redes menos sensíveis a estas variabilidades, é importante o emprego de ferramentas estatísticas para avaliar tanto a preservação da topologia inerente aos dados (DEMARTINES; HÉRAULT, 1997), como a variabilidade do erro de quantização (distância euclidiana entre as amostras e os neurônios da grade) e estabilidade da vizinhança (BODT; COTTRELL; VERLEYSSEN, 2002). No último critério é verificado se os grupos formados e as relações entre eles são estatisticamente significativas ou apenas uma mera casualidade.

Tanto quanto se sabe, não existe nenhum trabalho que integre uma rede auto-organizável ao nariz eletrônico na análise de café solúvel. Portanto, o objetivo deste trabalho é aplicar o SOM no reconhecimento de padrões aromáticos de café solúvel através de dados obtidos por um nariz eletrônico. Para verificar a confiabilidade dos resultados obtidos foi avaliada a preservação da topologia, erro de quantização e estabilidade da vizinhança. Foi realizada, também, uma análise dos pesos da rede treinada na tentativa de extrair regras para os grupos formados.

Todos os resultados foram comparados com a ACP para apresentar uma alternativa não-linear de análise multivariada.

2.2 MATERIAIS E MÉTODOS

2.2.1 Café Solúvel

Foram usados sete tipos de café solúvel (Tabela 1), sendo as amostras experimentais produzidas na planta piloto do laboratório de pesquisa e desenvolvimento da Companhia Iguazu de Café Solúvel (Cornélio Procópio – Paraná). As demais amostras são cafés solúveis comerciais adquiridos no varejo do Brasil e Inglaterra (BASSOLI, 2006).

Tabela 1 – Descrição resumida das amostras de café solúvel analisadas

Amostra*	Café	Descrição
1	1	Amostra experimental de torra média. Café arábica classificado como bebida dura e proveniente da região sul de Minas Gerais.
2		
3		
4	2	Amostra experimental de torra média. Café arábica classificado como bebida dura e proveniente da região norte do Paraná.
5		
6		
7	3	Amostra experimental de torra média. Café conilon proveniente do estado de Rondônia.
8		
9		
10	4	Amostra comercial brasileira de torra escura e constituída de café conilon e arábica.
11		
12		
13	5	Amostra brasileira, comercial e de composição desconhecida. Intensidade de torra, estimada sensorialmente, escura.
14		
15		
16	6	Amostra inglesa, comercial e de composição desconhecida. Intensidade de torra, estimada sensorialmente, média.
17		
18		
19	7	Amostra inglesa, comercial e de composição desconhecida. Intensidade de torra, estimada sensorialmente, média/clara.
20		
21		

* Amostras em triplicata.

2.2.2 Nariz Eletrônico

Todas as análises foram realizadas nos laboratórios de pesquisa e desenvolvimento da Companhia Iguazu de Café Solúvel. Foi utilizado o equipamento Portable Electronic Nose Type PEN2 (Airsense Analytics, Alemanha), que monitora a variação da condutância em cada um de seus sensores (Tabela 2) conforme a passagem do fluxo gasoso com a amostra (HAI; WANG, 2006; ZHANG et al., 2007). Essa variação é, então, comparada com àquela obtida pelo branco (ar ambiente) gerando, assim, um sinal adimensional.

Tabela 2 – Especificidade dos sensores utilizados no nariz eletrônico PEN2

Sensor	Especificidade
S1	Compostos aromáticos
S2	Ampla espectro, muito sensível, reage com óxidos de nitrogênio
S3	Amônia, usado para compostos aromáticos
S4	Hidrogênio
S5	Alcanos e compostos aromáticos menos polares
S6	Metano, amplo espectro
S7	Compostos sulfurados, terpenos, limonenos e pirazinas
S8	Álcoois, amplo espectro
S9	Compostos orgânicos sulfurados
S10	Metanos alifáticos em altas concentrações (> 100 ppm)

No preparo das amostras 2,8g de café solúvel foram colocados em frascos específicos para análise de headspace. Cada café (Tabela 1) foi avaliado em três repetições genuínas. Os parâmetros operacionais do equipamento foram (BASSOLI, 2006): 1,0 s de intervalo entre as amostras; 60 s de tempo de medição; 120s para limpeza ou purga dos sensores; 400 mL/min de fluxo de injeção.

2.2.3 Integração e Autoescalamento

Após a aquisição dos dados, a área abaixo da curva gerada por cada sensor foi numericamente integrada (SPERANDIO; MENDES; SILVA, 2003), formando um vetor de entrada com 10 valores para cada amostra. Antes de serem

alimentados na rede neural os vetores de entrada foram autoescalados (média igual a zero e variância unitária) para garantir a mesma métrica para todas as variáveis (HAYKIN, 2001).

2.2.4 Mapa Auto-Organizável (SOM)

Neste trabalho foi implementado o algoritmo SOM (em inglês self-organizing map – SOM) bidimensional proposto e discutido em maiores detalhes por HAYKIN (2001), que serviu de suporte metodológico. O principal objetivo do mapa auto-organizável é transformar um padrão de sinal incidente de dimensão arbitrária em um mapa discreto bidimensional e realizar esta transformação adaptativamente de uma maneira topologicamente ordenada. No modelo de Kohonen (Figura 1) cada neurônio da grade está totalmente conectado com todos os nós de fonte da camada de entrada. Esta grade representa uma estrutura alimentada adiante com uma única camada computacional consistindo de neurônios arranjados em linhas e colunas. Os neurônios apresentam a mesma quantidade de pesos, que por sua vez é igual ao número de variáveis analisadas. Cada variável pode ser representada por um plano de pesos (Figura 1), que após a fase de treinamento pode ser utilizado na tentativa de extrair regras sobre os grupos formados. O algoritmo responsável pela formação do SOM começa, primeiramente, inicializando os pesos sinápticos da grade. Isto é feito atribuindo-lhes valores pequenos (entre 0 e 1) tomados de um gerador de números aleatórios. Fazendo dessa forma, nenhuma organização prévia é imposta ao mapa de características. Após a inicialização, há três processos essenciais: competição, cooperação e adaptação sináptica.

No processo competitivo, para cada padrão de entrada, os neurônios da grade calculam seus respectivos valores de uma função discriminante (distância euclidiana). O neurônio que apresentar a menor distância euclidiana é chamado de neurônio vencedor, ou centróide, para o padrão de entrada considerado. O neurônio vencedor localiza o centro de uma função que representa a vizinhança topológica de neurônios cooperativos. A razão do processo cooperativo provém da evidência neurobiológica de interação lateral entre os neurônios.

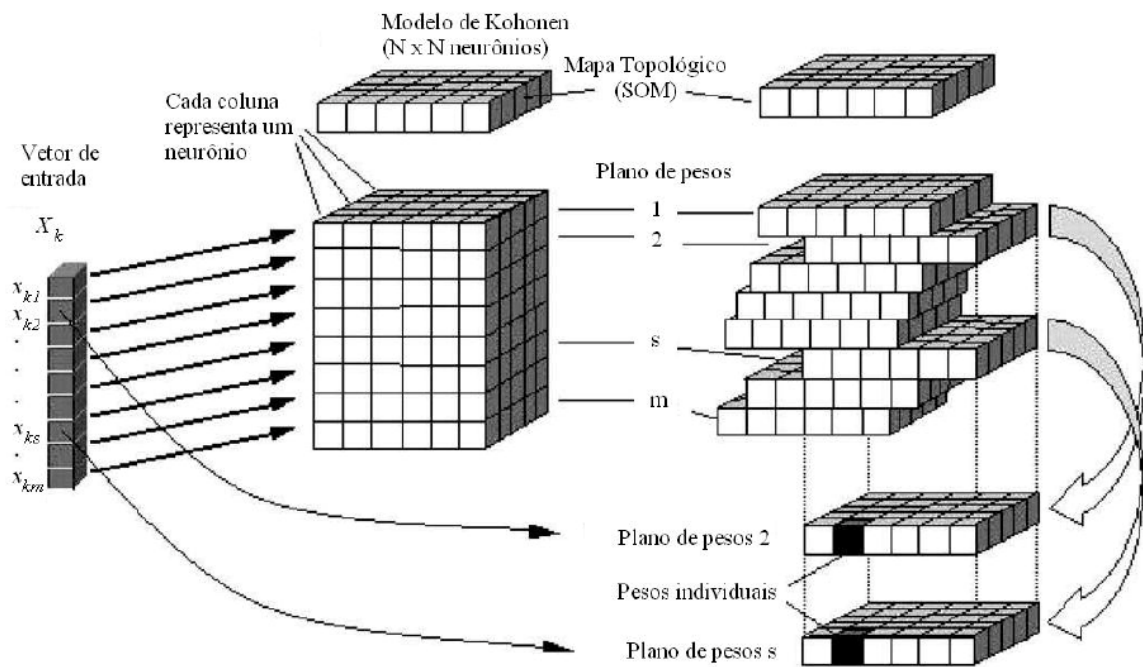


Figura 1 – Mapa auto-organizável segundo o modelo de Kohonen. Na figura k representa o número de padrões de entrada, m é a quantidade variáveis de entrada e N é o número de neurônios em cada dimensão (MARINI; ZUPAN; MAGRÌ, 2005).

A função escolhida para representar a vizinhança topológica deve satisfazer as seguintes propriedades: (i) ser simétrica em relação ao ponto de máximo definido pela posição do neurônio vencedor; (ii) sua amplitude deve ser monotonamente decrescente com o aumento da distância em relação ao neurônio vencedor. Uma escolha típica que satisfaz estas exigências é a função gaussiana apresentada na equação (1).

$$h_{j,i} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right) \quad (1)$$

Na equação acima o parâmetro σ é o raio efetivo da vizinhança topológica (Figura 2a). Ele mede o grau com o qual os neurônios vizinhos ao vencedor participam do processo de aprendizagem. Em um sentido qualitativo, a vizinhança topológica gaussiana é mais biologicamente apropriada que uma vizinhança quadrada (Figura 2b). Seu uso também faz com que o algoritmo SOM convirja mais rapidamente que com uma vizinhança topológica retangular. Ainda na equação (1), $d_{j,i}$ representa a distância lateral entre o neurônio vencedor i e o

neurônio excitado j . No caso de uma grade bidimensional esta distância é definida através da distância euclidiana entre as posições de i e j , ambas determinadas no espaço de saída discreto e definidas pelos vetores \mathbf{r}_i e \mathbf{r}_j respectivamente.

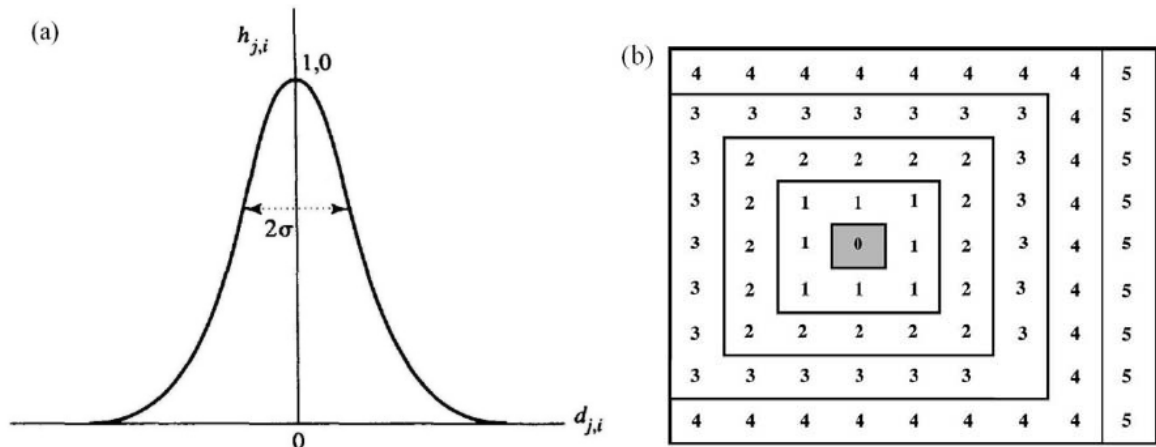


Figura 2 – Representação de vizinhança topológica: (a) gráfico unidimensional da função gaussiana; (b) representação de uma vizinhança quadrada ilustrando a distância de cada neurônio (raio) para aquele marcado em cinza.

Um ponto importante do algoritmo SOM é a redução do tamanho da vizinhança ao longo do treinamento. No início a função deve ser suficientemente larga para cobrir todo o mapa, porém a medida que o treinamento prossegue a largura deve ser diminuída para garantir a formação de regiões topológicas especializadas. Uma maneira de implementar a redução é através de um decaimento exponencial do tipo representado pela equação (2).

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_1}\right), \quad n = 0, 1, 2, \dots \quad (2)$$

Onde σ_0 é o raio efetivo de inicialização do algoritmo e τ_1 é uma constante de tempo, sendo recomendado o uso de $\tau_1 = 1000/\log \sigma_0$. O tempo discreto n representa a época de treinamento, ou seja, uma apresentação completa de todos os padrões de entrada.

Durante o processo adaptativo é necessário que o vetor de peso sináptico \mathbf{w}_j do neurônio j da grade se modifique em relação ao vetor de entrada \mathbf{x} . O

processo de correção ocorre através de uma modificação no postulado de aprendizagem de Hebb. Matematicamente e usando o formalismo de tempo discreto, o processo adaptativo é realizado por:

$$\mathbf{w}_j(n+1) = \mathbf{w}_j + \eta(n)h_{j,i}(\mathbf{x})(\mathbf{x} - \mathbf{w}_j(n)) \quad n = 0, 1, 2, \dots \quad (3)$$

Onde $\eta(n)$ é a taxa de aprendizagem que, como representado, deve ser variável e decrescente ao longo das n épocas de treinamento. O processo de adaptação pode ser dividido em duas etapas: ordenação e convergência. Na etapa de ordenação a taxa de aprendizagem deve ser maior para garantir que o mapa topológico, inicialmente desordenado, sofra maiores ajustes em seus pesos sinápticos. Na fase de convergência é realizada uma sintonia fina do mapa usando uma taxa de aprendizagem menor. O decréscimo da taxa de aprendizagem também pode ser realizado através de um decaimento exponencial, como apresentado na equação (4). Nesta equação, η_0 é a taxa de aprendizagem inicial e τ_2 é uma outra constante de tempo, sendo os valores recomendados, respectivamente, 0,1 e 1000.

$$\eta(n) = \eta_0 \exp\left(-\frac{n}{\tau_2}\right), \quad n = 0, 1, 2, \dots \quad (4)$$

Após o treinamento, o resultado pode ser visualizado através de um mapa topológico e mapas de peso. No primeiro, cada padrão de entrada é posicionado sobre o mapa de características de acordo com a posição do neurônio vencedor da amostra considerada (Figura 1). A função do mapa topológico é facilitar a visualização de agrupamentos e também a relação de vizinhança entre os grupos formados. Grupos próximos partilham de alguma similaridade, assim como, quanto maior a distância maior será a diferença de comportamento. O mapa de pesos (ZUPAN; NOVIČ; RUISÁNCHEZ, 1997) é um gráfico de contorno para cada plano de pesos (Figura 1). Em conjunto com o mapa topológico é possível extrair regras de comportamento para cada grupo formado e inferir a influência de cada variável sobre o resultado obtido.

2.2.5 Preservação da Topologia

Como um SOM faz uma representação de um espaço multi-dimensional em apenas duas dimensões, deve ser avaliado se as propriedades topológicas dos dados foram preservadas durante a transformação. Para uma avaliação qualitativa da preservação da topologia, Demartines e Héroult (1997) propuseram a utilização do gráfico $dy-dx$. Neste diagrama de dispersão representa-se no eixo x (dx) a distância euclidiana entre as posições definidas pelos vetores de entrada para todos os pares possíveis entre as amostras existentes. E, no eixo y (dy) a distância euclidiana entre os centróides de cada um pares possíveis. Se a topologia for preservada, os pontos devem se distribuir ao longo de uma linha. Além do gráfico, neste trabalho, também foi calculado o coeficiente de correlação (SPIEGEL, 1974) para avaliar o grau de linearidade entre os pontos.

2.2.6 Confiabilidade do Mapa Auto-Organizável

O objetivo desta etapa é verificar, através de testes estatísticos clássicos, se o agrupamento e as relações de vizinhança entre os grupos devem-se às propriedades intrínsecas dos dados ou apenas ao acaso. Segundo Bodt et al. (2002), as duas principais fontes de variabilidade para o SOM são devidas à inicialização e amostragem. O treinamento de vários mapas usando sempre o mesmo conjunto de dados possibilitaria apenas a análise da variabilidade devido ao processo de geração aleatória dos pesos iniciais. Portanto, é necessária a utilização de uma técnica de reamostragem que permita um conjunto de dados diferentes a cada novo treinamento. Com a metodologia bootstrap é possível utilizar amostras geradas artificialmente com a mesma distribuição empírica (média e desvio padrão) dos dados originais. Assim, utilizando o princípio de reamostragem bootstrap, cada SOM foi treinado 100 vezes (BODT; COTTRELL; VERLEYSEN, 2002), sendo que em cada repetição os pesos iniciais e as amostras eram diferentes, permitindo assim, avaliar as duas principais causas de variabilidade. Após cada repetição de treinamento foi avaliado o erro médio de quantização (EMQ) do mapa em relação aos K padrões entrada, conforme a equação (5). Sendo que, EQ_k representa o erro

de quantização para cada amostra, que é definido como a distância euclidiana entre o vetor de entrada e o vetor de pesos do respectivo neurônio vencedor.

$$EMQ = \frac{1}{K} \sum_{k=1}^K EQ_k \quad (5)$$

Com os valores de EMQ de cada repetição, foi calculado um coeficiente de variação que permitiu avaliar a estabilidade da representação dos dados pelo SOM. Além disso, segundo Bodt et al. (2002), um gráfico do coeficiente de variação contra o número de neurônios usado fornece uma maneira empírica para selecionar a tamanho mais adequado do SOM para os dados analisados. Segundo o mesmo autor, o aumento no número de neurônios usados em cada dimensão leva a um aumento do coeficiente variação. Pois, assim sendo, quanto maior o número unidades no mapa maior serão as possibilidades de distribuição das amostras, levando assim, a uma instabilidade nos resultados.

Para avaliar a confiabilidade dos agrupamentos e relações de vizinhança é necessário definir uma nova função (6). Nesta equação, o raio r é definido para uma função de vizinhança quadrada conforme ilustrado na Figura 2b. Ser vizinhos significa que dois padrões de entrada \mathbf{x}^i e \mathbf{x}^j estão projetados no mapa topológico em centróides localizados dentro da região de raio r . O índice b indica que a função (6) deve ser avaliada para cada repetição de treinamento. Pode acontecer que em uma dada repetição, duas amostras sejam vizinhas e em outra não. Esta variabilidade é justamente o objetivo da avaliação de confiabilidade da vizinhança.

$$VIZ_{i,j}^b(r) = \begin{cases} 0 & \text{se } \mathbf{x}^i \text{ e } \mathbf{x}^j \text{ não são vizinhos dentro da região de raio } r \\ 1 & \text{se } \mathbf{x}^i \text{ e } \mathbf{x}^j \text{ são vizinhos dentro da região de raio } r \end{cases} \quad (6)$$

Após a avaliação das relações de vizinhança em todas as repetições, é necessária a obtenção de um valor médio para a função (6) conforme ilustrado pela equação (7).

$$EVIZ_{i,j}(r) = \frac{\sum_{b=1}^B VIZ_{i,j}^b(r)}{B} \quad (7)$$

Na equação anterior, B é o número de repetições de treinamento e $EVIZ_{i,j}(r)$ representa a estabilidade da relação de vizinhança entre um par de amostras x_i e x_j dentro do raio r considerado. Se o par considerado for vizinho em todas as repetições $EVIZ_{i,j}(r) = 1$, caso contrário, $EVIZ_{i,j}(r) = 0$.

O próximo passo é realizar um teste de significância estatística do valor de $EVIZ_{i,j}(r)$. Antes, entretanto, é necessário incluir mais algumas considerações sobre a vizinhança quadrada. A quantidade de vizinhos em uma região de raio r de uma mapa bidimensional é igual a $v = (2r + 1)^2$, como mostra a Figura 2b. Sendo N o número total de neurônios no mapa, a probabilidade de um determinado par de observações ser vizinho ao acaso é de v/N . Por exemplo, para N um mapa bidimensional com 7 neurônios em cada dimensão e um raio de vizinhança igual a 1, a probabilidade de vizinhança ao acaso seria de 18,4%; para $r = 2$, 51,0% e assim por diante. Porém, os neurônios localizados nos extremos do mapa não apresentam o mesmo número de vizinhos daqueles localizados no centro. Em Bodt et al. (2002) foi recomendado o uso da quantidade média de vizinhos, principalmente em mapas pequenos e de raios maiores. Para o exemplo citado acima, se for considerado um número médio de vizinhos as probabilidades ficam, respectivamente, iguais a 15,0% e 35,0%.

Para construir um teste de hipótese foi usada uma distribuição binomial com probabilidade de sucesso (x_i e x_j são vizinhos) igual a v/N em B repetições independentes. A hipótese nula (H_0) considera que a relação de vizinhança de x_i e x_j é devida ao acaso. A hipótese alternativa (H_1) supõe que o par pode ser estatisticamente considerado vizinho ou não dentro de um raio pré-determinado. Para facilitar o cálculo do teste de hipótese é preciso aproximar a distribuição binomial usando uma distribuição gaussiana. Para a aproximação ser válida é necessário cumprir dois critérios: (i) a quantidade de repetições (B) de ser suficientemente grande (> 30); (ii) as probabilidades de sucesso ou insucesso não devem ser próximas a zero ou, neste caso, e

$$\frac{Bv}{N} > 10 \text{ e } B\left(1 - \frac{v}{N}\right) > 10$$

(SPIEGEL, 1974). Usando a segunda condição e fixando a quantidade de repetições, $B = 100$ neste trabalho, foi possível construir um intervalo de tamanho válido para a vizinhança em função do número total de neurônios do

mapa (8), assim como, um intervalo de raio ideal para avaliar a estabilidade das relações de vizinhança.

$$\frac{N}{10} < \nu < \frac{9N}{10} \quad (8)$$

Após as considerações acima é possível construir o intervalo para aceitação de H_0 com 95% de confiança (9). Caso o valor de $EVIZ_{i,j}(r)$ seja maior que o limite superior deste intervalo, H_1 é aceita e o par de amostras é considerado vizinhos no nível de 2,5%. Se o valor for menor que o limite inferior, H_1 é igualmente aceita e o par é considerado não vizinho no mesmo nível de significância.

$$B \frac{\nu}{N} \pm 1,96 \sqrt{B \frac{\nu}{N} - \left(1 - \frac{\nu}{N}\right)} \quad (9)$$

O teste de hipótese no valor de $EVIZ_{i,j}(r)$ permite avaliar a estabilidade da relação de vizinhança entre um par específico de observações. Para uma análise que considere todos os pares possíveis, dois critérios podem ser examinados. Primeiro, na distribuição utilizada para o teste de hipótese a probabilidade ao acaso de aceitar H_1 é de 5%. Portanto, se o percentual de vizinhos ou não vizinhos, estatisticamente válidos, para um determinado r for menor do que 5%, a análise não pode ser considerada significativa. A segunda alternativa é a construção de gráficos de probabilidade acumulada (BODT et al., 2002) que permitem uma análise qualitativa de significância. Uma discrepância entre o gráfico de probabilidade acumulada do mapa auto-organizável e da distribuição binomial é um indicativo de que as relações de vizinhança construídas não são casuais.

2.2.7 Análise de Componentes Principais (ACP)

É um dos métodos multivariados mais utilizado e consiste na combinação linear das variáveis originais produzindo componentes principais (CP) ortogonais entre si. Cada CP é responsável pela representação de um percentual da variância dos dados, sendo ideal que no resultado final quase toda a variabilidade seja representada por no máximo três CP. Segundo Hair et al. (2005), para escolha do número adequado de CPs existem vários critérios, entre eles: (i) critério do

autovalor – são consideradas apenas as CPs com autovalor superior a 1; (ii) critério de percentagem de variância – a quantidade de CPs escolhidas deve ser suficiente para explicar pelo menos 95% da variância dos dados; (iii) critério do teste scree – o ponto no qual o gráfico de scree começa a ficar horizontal é considerado indicativo do número máximo de componentes a serem extraídos. Após a escolha da quantidade de CPs a serem utilizadas, é possível fazer uma análise do gráfico de escores fatoriais, o qual permite a visualização de agrupamentos nos dados. Já no gráfico das cargas fatoriais é possível visualizar a influência de cada variável original nos grupos formados (VERDINI et al., 2007).

2.2.8 Implementação Computacional

Para execução do algoritmo SOM, reamostragem bootstrap, testes preservação da topologia e confiabilidade das relações de vizinhança, foi construído um programa em linguagem FORTRAN 90. Por uma questão de qualidade de apresentação, todos os gráficos dos resultados obtidos pelo programa desenvolvido foram gerados no software STATISTICA 7.1 (StatSoft Inc., Tulsa, OK, USA). A análise de componentes principais (usando a matriz de correlação) e os testes estatísticos básicos (correlação, ANOVA e teste Tukey) também foram realizados no STATISTICA 7.1. Todos os cálculos foram realizados em um computador com processador Intel Celeron D[®] de 2,8 GHz e 1,0 Gb de memória RAM.

2.3 RESULTADOS E DISCUSSÃO

2.3.1 Pré-Processamento

As áreas médias de cada sensor para cada café (Tabela 1) estão apresentadas e estatisticamente comparadas na Tabela 3. Foi observado que para os dez sensores houve diferença significativa ($p < 0,05$) entre os cafés. Assim sendo, para a análise dos dados usando o SOM e ACPs todos os sensores foram considerados como variáveis de entrada.

Tabela 3 – Comparação entre as áreas médias obtidas pelo nariz eletrônico para cada sensor e café *

Sensor	Café						
	1	2	3	4	5	6	7
S1	36,03 ^d	46,70 ^b	54,84 ^a	52,78 ^a	48,98 ^b	41,71 ^c	34,84 ^d
S2	427,08 ^a	172,18 ^d	165,00 ^d	114,96 ^e	174,08 ^{c,d}	214,93 ^c	358,12 ^b
S3	32,86 ^d	45,27 ^b	52,39 ^a	51,04 ^a	47,03 ^b	40,47 ^c	33,60 ^d
S4	59,37 ^{a,b}	57,00 ^c	57,37 ^c	59,00 ^{a,b}	58,77 ^b	59,20 ^{a,b}	60,10 ^a
S5	37,88 ^d	48,05 ^b	55,74 ^a	53,70 ^a	50,27 ^b	42,98 ^c	36,54 ^d
S6	109,88 ^b	84,03 ^c	62,56 ^e	70,40 ^{d,e}	77,35 ^{c,d}	109,86 ^b	137,56 ^a
S7	141,15 ^a	70,39 ^b	65,07 ^b	61,31 ^b	70,27 ^b	68,22 ^b	134,53 ^a
S8	85,41 ^c	67,87 ^d	52,80 ^e	65,35 ^d	69,31 ^d	93,30 ^b	114,51 ^a
S9	314,45 ^a	117,74 ^d	96,18 ^{e,f}	79,86 ^f	107,14 ^{d,e}	149,10 ^c	232,04 ^b
S10	59,61 ^{a,b}	54,14 ^c	53,68 ^c	60,06 ^{a,b}	58,75 ^{a,b}	59,74 ^{a,b}	62,22 ^a

As médias pertencentes a mesma linha e acompanhadas da mesma letra não apresentam diferença significativa em nível de 5% pelo teste Tukey.

2.3.2 Análise de Componentes Principais

Como mencionado anteriormente, a ACP foi executada usando a matriz de correlação dos sinais integrados e autoescalonados de cada um dos 10 sensores utilizados pelo nariz eletrônico (Tabela 2). De acordo com os critérios propostos por Hair et al. (2005), a quantidade de CPs extraídas deveria ser: duas segundo o critério do autovalor, três para o critério da variância explicada e quatro para o teste scree. Na Figura 3, é possível verificar através do gráfico como é definido o número de componentes para cada um dos critérios citados no item 2.2.7. Como o objetivo deste trabalho não é fazer uma análise detalhada dos dados usando ACP, mas sim, usar estes resultados para comparar e validar o uso do SOM foram escolhidas apenas duas componentes (critério do autovalor) totalizando 93,4% da variância total dos dados. Além disso, como será discutido adiante, duas componentes foram suficientes para discriminar os tipos de café e comparar os resultados obtidos com os do SOM.

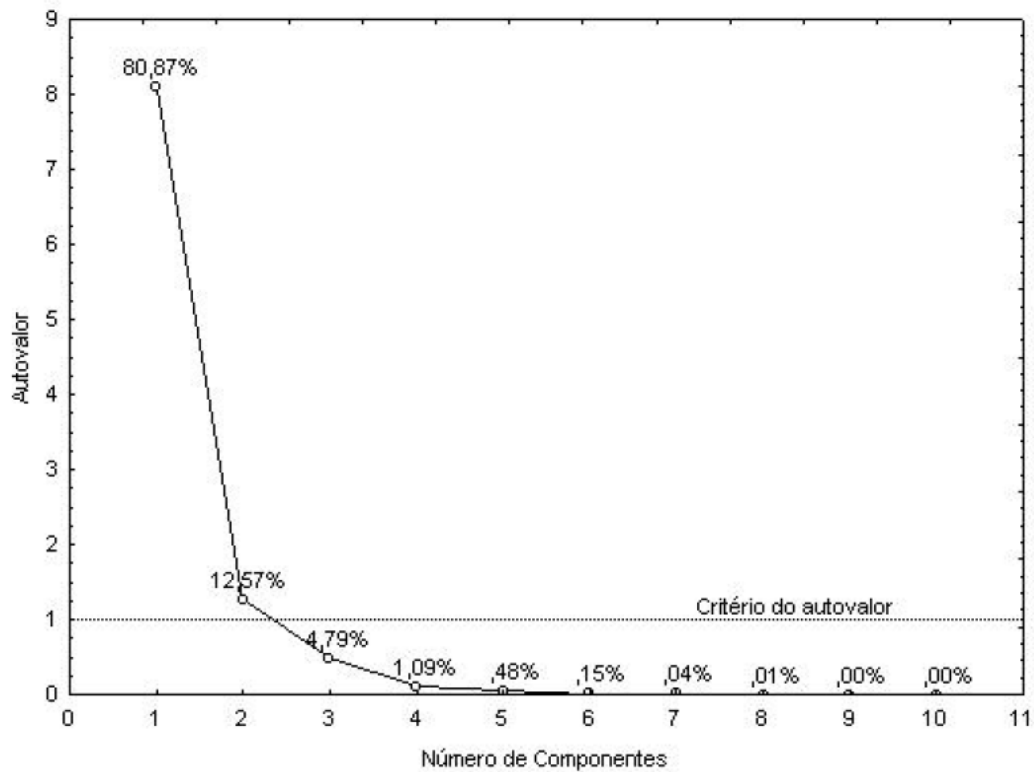


Figura 3 – Gráfico de autovalor para o critério do teste scree. No mesmo gráfico foi representado o limite do critério do autovalor.

Fazendo a projeção dos escores fatoriais (Figura 4a) no sistema de coordenadas formado pelas duas primeiras componentes selecionadas, é possível observar que houve uma separação das amostras de acordo com o tipo de café. As amostras 6 e 8 se afastaram um pouco mais das outras repetições para os respectivos cafés, entretanto, não existiu uma sobreposição entre os grupos.

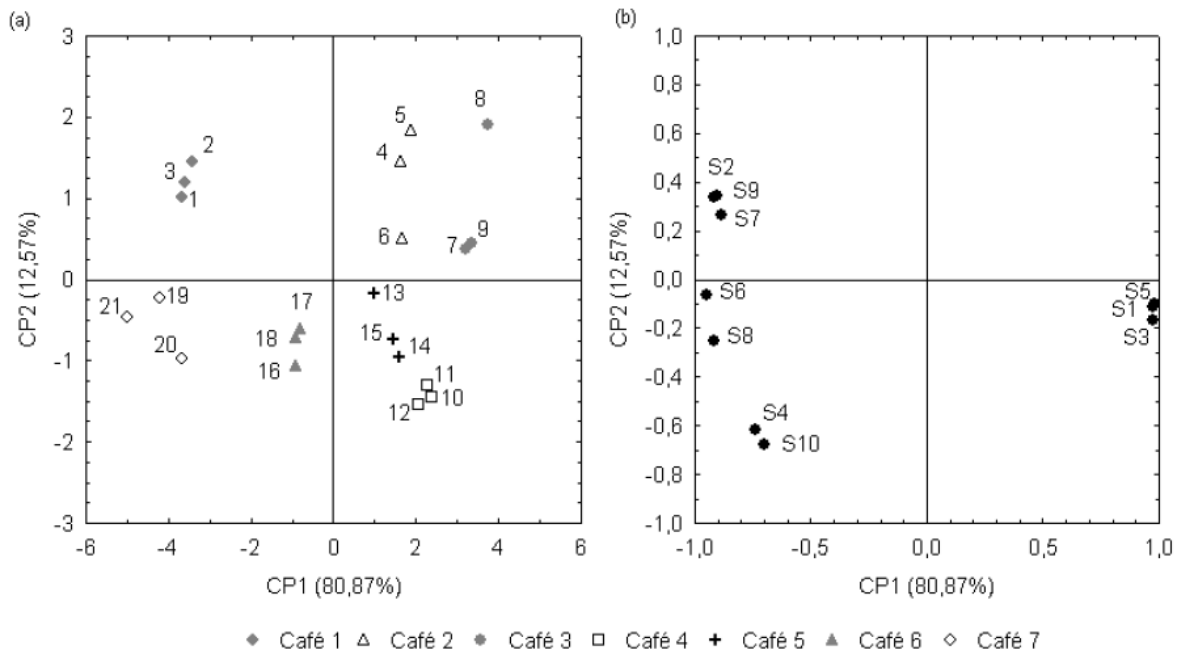


Figura 4 – Projeção dos escores (a) e das cargas (b) fatoriais nas duas componentes principais selecionadas.

A partir do gráfico das cargas fatoriais (Figura 4b) foi verificado que as variáveis de entrada podem ser divididas em quatro grupos: grupo 1 (S1, S3 e S5), grupo 2 (S2, S7 e S9), grupo 3 (S4 e S10) e grupo 4 (S6 e S8). Uma distribuição semelhante foi obtida por Zhang et al. (2007), quando um nariz eletrônico com os mesmos sensores, marca e modelo foi utilizado para trigo. Os sensores pertencentes ao mesmo grupo possuem alta correlação positiva entre si (Tabela 4). O mesmo perfil de correlação foi apresentado por Hai et al. (2006), o qual também utilizou o mesmo equipamento para a identificação de adulteração em óleo de gergelim. O padrão de correlação é resultado da especificidade de cada sensor (Tabela 2). Outras discussões referentes a importância das variáveis originais no agrupamento das amostras serão realizadas em conjunto com a análise do SOM.

Tabela 4 – Correlação entre as áreas de cada sensor*.

Sensor	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	1	-0,90	1,00	-0,61	1,00	-0,96	-0,84	-0,90	-0,89	-0,59
S2	-0,90	1	-0,93	0,52	-0,89	0,80	0,95	0,70	0,99	0,44
S3	1,00	-0,93	1	-0,60	0,99	-0,93	-0,87	-0,86	-0,93	-0,57
S4	-0,61	0,52	-0,60	1	-0,62	0,67	0,55	0,76	0,52	0,96
S5	1,00	-0,89	0,99	-0,62	1	-0,96	-0,84	-0,91	-0,89	-0,60
S6	-0,96	0,80	-0,93	0,67	-0,96	1	0,76	0,98	0,77	0,66
S7	-0,84	0,95	-0,87	0,55	-0,84	0,76	1	0,69	0,92	0,49
S8	-0,90	0,70	-0,86	0,76	-0,91	0,98	0,69	1	0,66	0,76
S9	-0,89	0,99	-0,93	0,52	-0,89	0,77	0,92	0,66	1	0,44
S10	-0,59	0,44	-0,57	0,96	-0,60	0,66	0,49	0,76	0,44	1

* Todos os coeficientes de correlação estatisticamente significativos em nível de 5%.

2.3.3 Mapa Auto-Organizável

Foram testados seis SOM quadrados contendo de 4 a 9 neurônios em cada dimensão (Tabela 5). Cada mapa foi treinado ao longo de 5000 épocas, sendo que, o treinamento foi repetido 100 vezes usando para cada repetição pesos de inicialização e amostragem diferentes com objetivo de avaliar a estabilidade e confiabilidade dos resultados obtidos.

Tabela 5 – Resumo dos mapas auto-organizáveis treinados.

Neurônios		Quantidade de pesos	Largura inicial da vizinhança (σ_0)	Erro médio de quantização (EMQ)		Correlação obtida pelo gráfico $dy-dx^*$
Por dimensão	Total			Média*	Desvio padrão	
4	16	160	1,5	0,6046	0,0546	0,8039
5	25	250	2,0	0,3377	0,0421	0,7683
6	36	360	2,5	0,1846	0,0314	0,7612
7	49	490	3,0	0,0858	0,0223	0,7847
8	64	640	3,5	0,0381	0,0172	0,7777
9	81	810	4,0	0,0162	0,0133	0,7808

* Média para as 100 repetições.

Em todos os mapas, a largura inicial da função de vizinhança, equação (2), foi ajustada para garantir que no início do treinamento grande parte dos neurônios fosse incluída e ao término desse mesmo período, convergisse para um valor próximo a zero (HAYKIN, 2001). É notório, Tabela 5, que o aumento do

número de neurônios em cada dimensão provoca uma redução no EMQ, equação (5). Entretanto, este comportamento pode ser indesejável pois, a expansão do mapa permite um centróide para cada amostra, ou seja, a rede acaba “memorizando” os padrões de entrada e não “aprende” as características de cada grupo. Quando este fato ocorre, diz-se que a rede foi superajustada e perdeu a sua capacidade de generalização para padrões de entrada diferentes dos usados durante o treinamento (BASHEER; HAJMEER, 2000). Segundo Bodt et al. (2002), um critério melhor para escolher a quantidade de neurônios seria a análise do coeficiente de variação do EMQ. Um aumento muito elevado deste coeficiente indicaria um acréscimo na instabilidade dos agrupamentos formados. Na Figura 5, é possível observar que quanto mais neurônios, maior foi a taxa de crescimento do coeficiente de variação nas redes treinadas. Indicando que, um mapa com quatro unidades em cada dimensão (16 ao todo) seria mais adequado para os dados analisados. Outro fato que corrobora é a análise do coeficiente de correlação do gráfico $dy-dx$ (Tabela 5 e Figura 5). A rede com quatro neurônios foi a de maior correlação. Assim sendo, houve uma maior preservação da topologia dos dados (DEMARTINES; HÉRAULT, 1997). Além disso, foi observado que um aumento no mapa provoca apenas uma dispersão maior entre as amostras, contudo as propriedades de vizinhança são mantidas. Boishebert et al. (2006) também relataram este efeito, ou seja, a aplicação de mapas maiores seria um desperdício de tempo computacional.

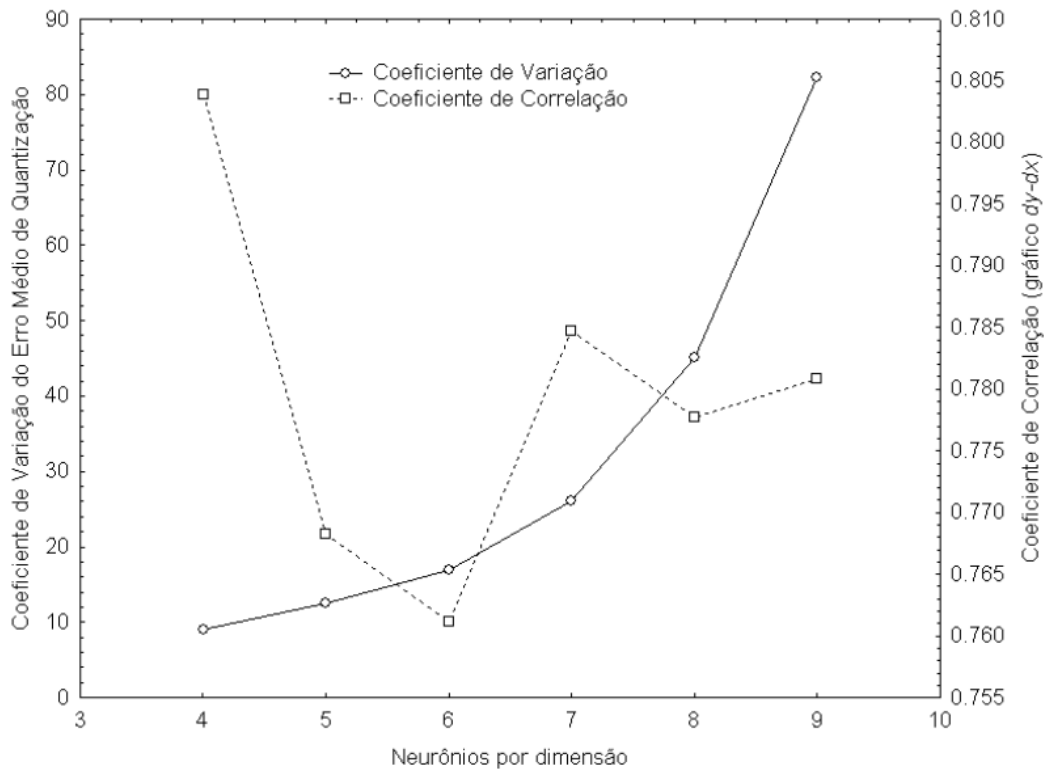


Figura 5 – Comportamento do coeficiente de variação do EMQ e coeficiente de correlação do gráfico dy-dx em função do número de neurônios em cada dimensão do mapa auto-organizável.

Na Figura 6 estão representados alguns parâmetros de treinamento e o gráfico dy-dx (Figura 6d) para o mapa auto-organizável de dimensões 4 x 4 unidades. Como recomendado, tanto a taxa de aprendizagem (Figura 6a) como a largura da função de vizinhança (Figura 6b) tiveram o valor reduzido próximo a zero ao final do treinamento (5000 épocas). No gráfico do erro médio de quantização foi observado um decréscimo muito rápido do valor nas primeiras 100 épocas (não está representado na Figura 6c devido à escala utilizada). Após este período a taxa de redução permaneceu praticamente constante.

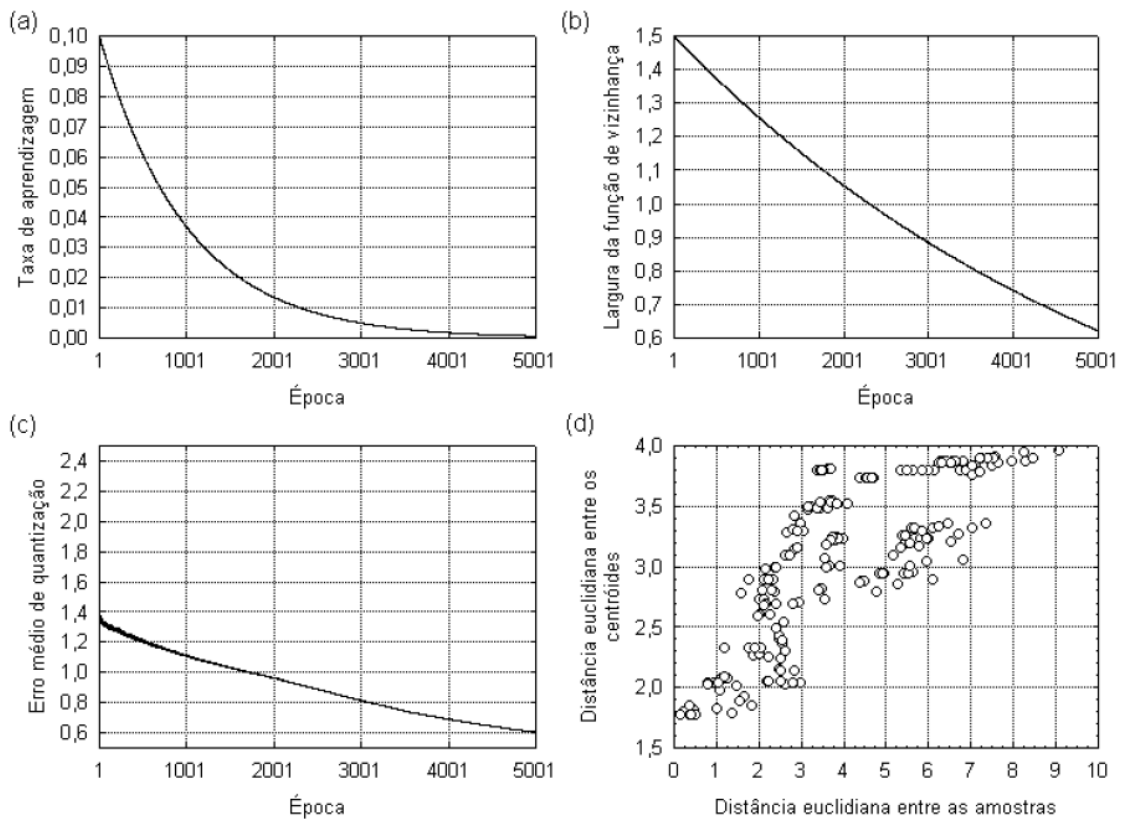


Figura 6 – Mapa auto-organizável 4 x 4: variação ao longo do treinamento da taxa de aprendizagem (a), da largura da função de vizinhança (b) e do erro médio de quantização (c); gráfico dy-dx (d).

A avaliação da confiabilidade das relações de vizinhança, equação (6), foi realizada usando regiões quadradas (Figura 2) com raios iguais a 0 e 1. Estes valores estão dentro do intervalo ideal de vizinhos ($1,6 < v < 14,4$) que obedece as condições necessárias para realizar o teste de hipótese em um mapa 4 x 4, equação (8). Um resumo das propriedades dos raios de vizinhança escolhidos está apresentado na Tabela 6.

Tabela 6 – Propriedades dos raios de vizinhança quadrada testados no mapa 4 x 4.

Raio (r)	v corrigida ^a	v/N ^b	Intervalo de aceitação de H_0 ^c	Vizinhança ^d		
				Vizinhos	Não vizinhos	% Total ^e
0	1,00	6,25	$0,0151 < EVIZ_{i,j}(r) < 0,1099$	20	182	96,19
1	6,25	39,06	$0,2950 < EVIZ_{i,j}(r) < 0,4863$	51	155	98,10

^a Quantidade corrigida de vizinhos dentro do raio de vizinhança considerado.

^b Probabilidade de vizinhança ao acaso usando a quantidade corrigida de vizinhos.

^c A hipótese H_0 supõe que a relação de vizinhança deve-se ao acaso, equação (9).

^d Quantidade de vizinhos ou não, estatisticamente significativos ao nível de 5%.

^e Percentual em relação ao total de pares possíveis (210) para 21 amostras.

Segundo os critérios (item 2.2.6) estabelecidos por Bodt et al. (2002), as análises de estabilidade das relações de vizinhança para $r = 0$ e $r = 1$ podem ser consideradas estatisticamente válidas. O percentual de vizinhos ou não vizinhos estatisticamente significativos (Tabela 6) é muito maior que 5%. Além disso, o gráfico de probabilidade acumulada do SOM 4×4 difere em ambos os raios daquele obtido para uma distribuição binomial (Figura 7).

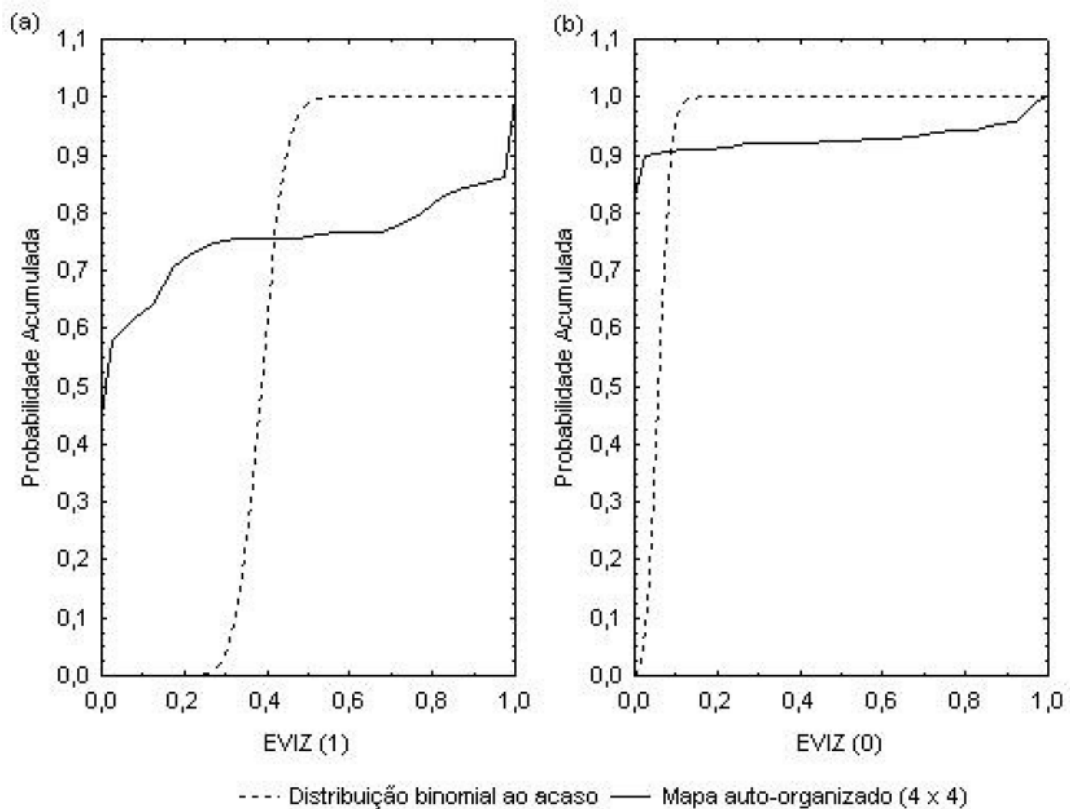


Figura 7 – Gráficos de probabilidade acumulada para um raio de vizinhança igual a 1 (a) e 0 (b).

Quando r for igual 0 significa que as amostras devem ocupar o mesmo centróide para serem consideradas vizinhas, ou seja, elas não apresentam diferença de comportamento e são consideradas repetições de um padrão. Para $r = 1$ as amostras são consideradas vizinhas se estiverem em centróides adjacentes. Logo, são amostras parecidas, mas não podem ser consideradas repetições de um mesmo padrão de comportamento. Na Figura 8a foram representadas as relações de vizinhança para a amostra 1 (café 1). Como esperado, observou-se que as

amostras 2 e 3 estão posicionadas dentro do mesmo centróide, pois são repetições do mesmo café. Já as amostras 19, 20 e 21, todas pertencentes ao café 7, estão em um centróide adjacente indicando uma semelhança de perfil aromático. A análise do gráfico dos escores fatoriais (Figura 4a) confirma a proximidade entre os grupos e justifica também o menor valor de $EVIZ_{1,20}(1)$, pois é visível que a amostra 20 está mais afastada, em relação as amostras 19 e 21, do grupo de amostras do café 1. Como o mesmo comportamento descrito para a amostra 1 foi observado para as repetições 2 e 3, os respectivos gráficos foram omitidos. As relações de vizinhança do café 7, representado pela amostra 19 e ilustradas na Figura 8b, confirmam a similaridade aromática com o café 1. Além disso, é evidente que o SOM agrupou as repetições 20 e 21 no mesmo centróide e que a amostra 20 está um pouco mais afastada do grupo pois, o valor de $EVIZ_{19,20}(0)$ é menor que $EVIZ_{19,21}(0)$. Para o café 6, amostra 16 na Figura 8c, foi observado apenas o agrupamento das repetições dentro do mesmo centróide. Portanto, não existe nenhum padrão aromático estatisticamente semelhante ao do café 6 dentro das amostras analisadas. As afirmações acima são corroboradas pela análise cuidadosa (as escalas de cada componente são diferentes) do gráfico dos escores fatoriais (Figura 4a).

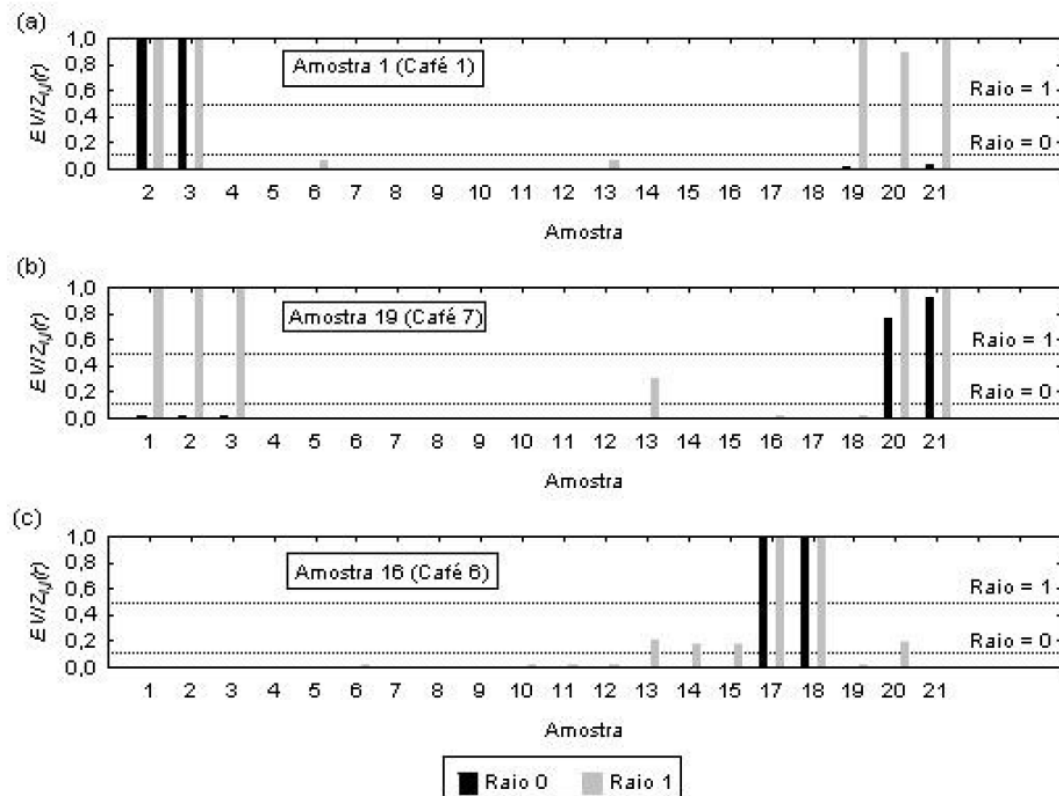


Figura 8 – Gráficos com os valores de $EVIZ_{ij}(r)$ para as amostras 1 (a), 19 (b) e 16 (c). As linhas pontilhadas horizontais representam o limite acima do qual uma amostra é estatisticamente considerada vizinha em nível de 2,5% para um determinado raio.

Para o café 2 (representado pela amostra 4 – Figura 9a) foi constatado o correto agrupamento das repetições em um mesmo centróide e uma vizinhança significativa ($p < 0,025$) com o café 3 quando o raio é igual a 1. A análise da Figura 9b permite inferir que a amostra 6 está mais afastada das outras do grupo (4 e 5), porém, as relações de vizinha são as mesmas com exceção à amostra 13 (café 5). A Figura 4a confirma e ilustra claramente o comportamento descrito para a amostra 6. A Figura 9c (amostra 7) sustenta as relações de vizinhança entre os cafés 2 e 3. Para amostra 8 também foi observado um distanciamento do grupo, semelhante ao da amostra 6, e ao contrário das repetições 7 e 9 não foi identificada uma vizinhança significativa com a amostra 13 para $r = 1$. A observação da Figura 4a, mais uma vez, sustenta as das proposições estabelecidas.

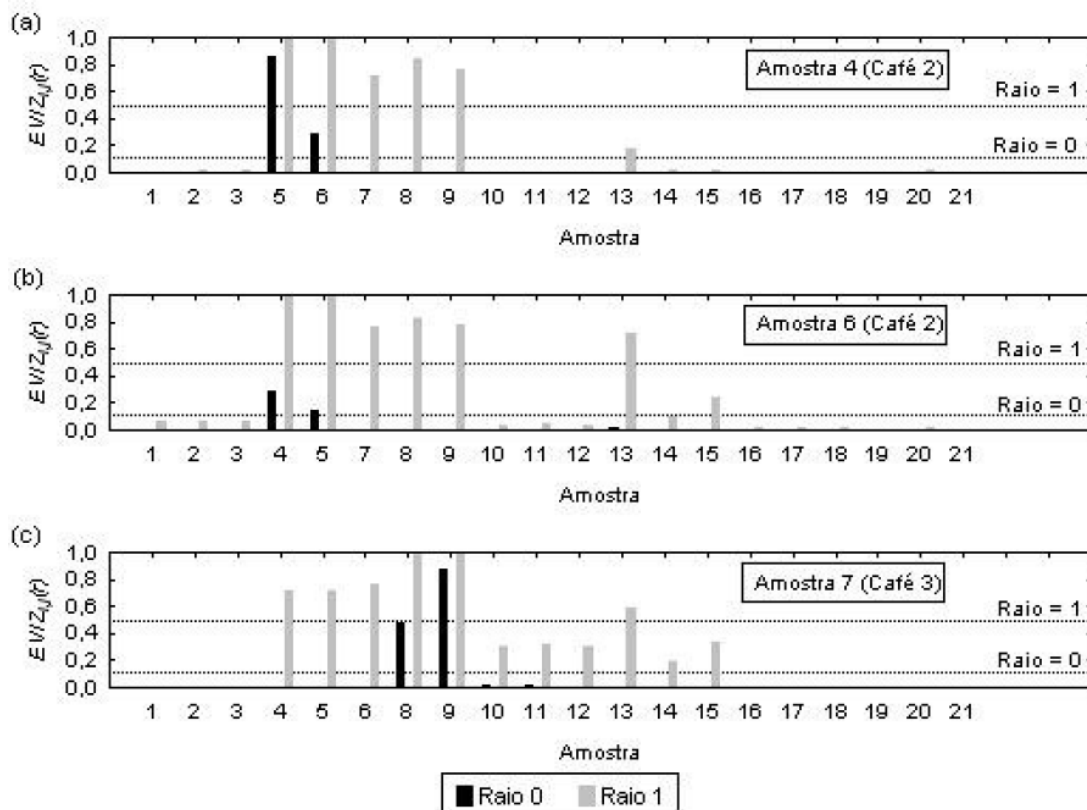


Figura 9 – Gráficos com os valores de $EVIZ_{ij}(r)$ para as amostras 4 (a), 6 (b) e 7 (c). As linhas pontilhadas horizontais representam o limite acima do qual uma amostra é estatisticamente considerada vizinha em nível de 2,5% para um determinado raio.

A amostra 10 (café 4 – Figura 10a) foi posicionada corretamente dentro do mesmo centróide com as repetições 11 e 12. É possível ver que existe uma vizinhança significativa ($p < 0,025$) com as amostras do café 5 confirmada também, pela inspeção da Figura 10c. O comportamento ligeiramente diferenciado da amostra 13 é representado na Figura 10b, que ratifica as já mencionadas relações de vizinhança com as amostras 6, 7 e 9. Foi verificado que os resultados obtidos pelo SOM são muito semelhantes àqueles da ACP (Figura 4a). Em Verdini et al. (2007) também foi relatada uma similaridade entre as metodologias. Segundo os autores, apenas pequenas variações foram observadas devido aos diferentes pesos que cada método dá para as variáveis de entrada. Já em Díaz et al. (2003), foi descrito que o SOM melhorou a separação das amostras em relação a ACP, por ser um método que permite a representação de relações não-lineares presentes nos dados. Vale a pena destacar que independentemente do tamanho e da

complexidade da matriz de dados o resultado do SOM sempre será bidimensional (HAYKIN, 2001).

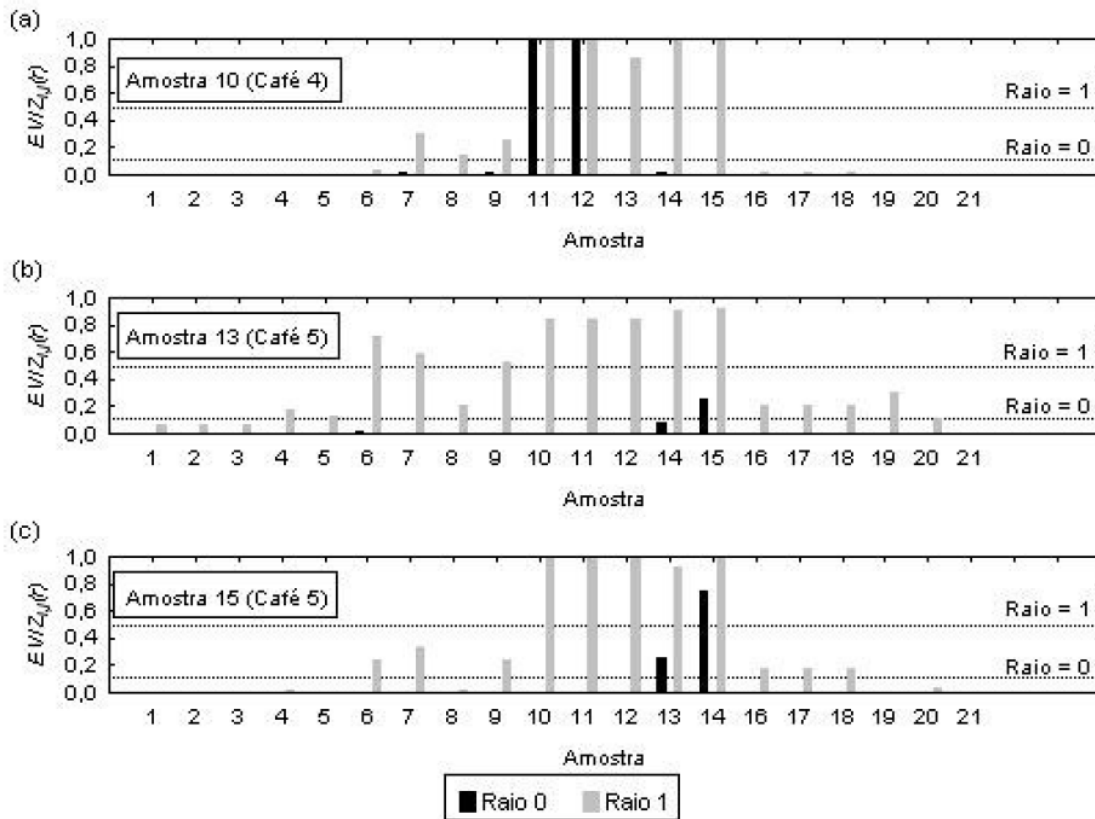


Figura 10 – Gráficos com os valores de $EVIZ_{i,j}(r)$ para as amostras 10 (a), 13 (b) e 15 (c). As linhas pontilhadas horizontais representam o limite acima do qual uma amostra é estatisticamente considerada vizinha em nível de 2,5% para um determinado raio.

É possível pesquisar entre as repetições de treinamento realizadas, uma que apresente o comportamento topológico inferido pela análise da estabilidade das relações de vizinhança (Figuras 8-10). Na Figura 11, foram representados os mapas topológicos das repetições de treinamento 20 e 98. É observada uma diferença de posicionamento dos grupos, entretanto as relações de vizinhança são mantidas.

para 2 e 5; menor 3 e 4. Todas as observações descritas acima estão de acordo com a ACP (Figura 4) e o teste de comparação de médias (Tabela 3). É possível, também, notar mais uma semelhança entre os mapas de peso e o gráfico das cargas fatoriais (Figura 4b). Para o grupo de sensores 1 (S1, S3 e S5) e 4 (S6 e S8) ocorre uma variação maior dos pesos ao longo da dimensão vertical do mapa (Figuras 12a e 12d), na ACP estes dois grupos praticamente contribuem apenas para a CP1. Nos grupos 2 (S2, S7 e S9) e, principalmente, 3 (S4 e S10) existe uma variação bidimensional dos pesos, assim como na ACP, tanto CP1 e CP2 estão correlacionadas com estes sensores.

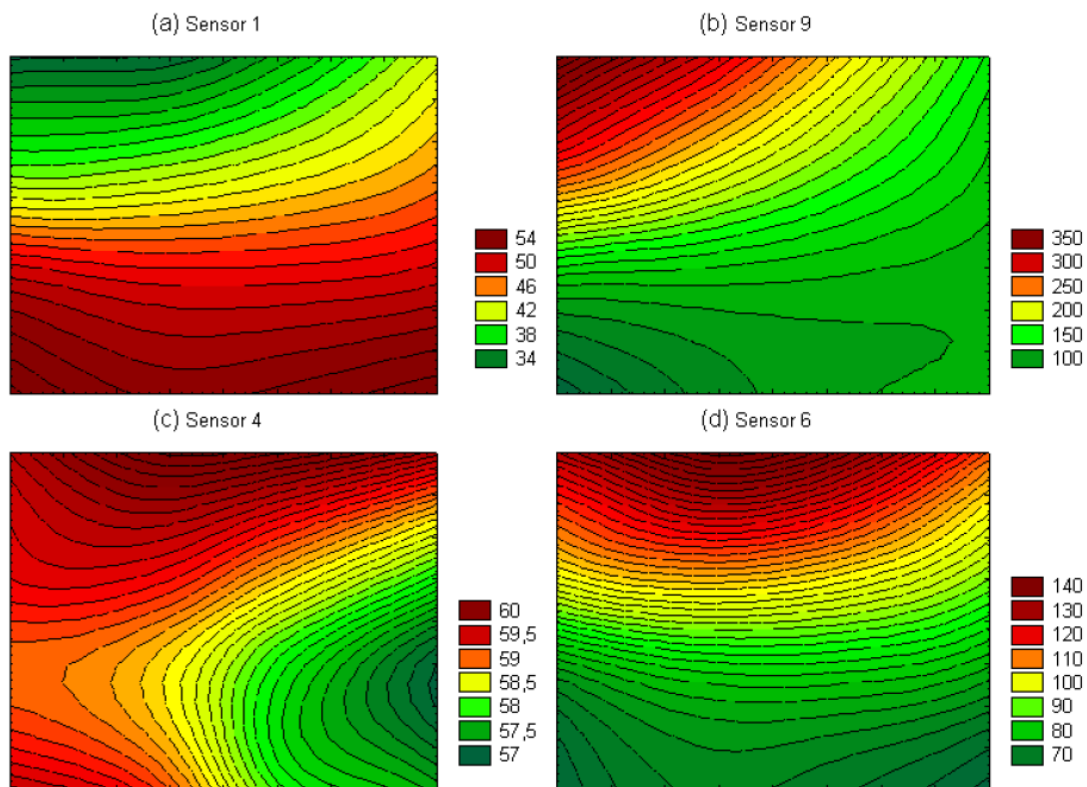


Figura 12 – Mapas de peso de alguns sensores selecionados (20ª repetição de treinamento).

Através dos mapas de peso e topológico foi observada uma semelhança muito grande entre os resultados da ACP e do SOM. Este fato sinaliza que, o comportamento aromático obtido pelo nariz eletrônico para os sete cafés analisados pode ser descrito através de relações lineares. Entre as vantagens do

uso do SOM foram verificadas a simplicidade matemática, a quantidade e qualidade dos resultados na forma de gráficos, os quais, facilitam a extração de regras sobre os dados sem a necessidade de um conhecimento profundo da metodologia. Além disso, o SOM permite a representação de relações não lineares em mapas bidimensionais sem aumento algum da complexidade do método. Assim, quando os dados apresentam comportamentos não lineares o uso do SOM é mais recomendado que a ACP (DÍAZ et al., 2003). Por outro lado, o treinamento de uma rede neural pode ser computacionalmente mais oneroso que as metodologias multivariadas clássicas.

2.4 CONCLUSÕES

As redes neurais por vezes são criticadas e designadas “caixas pretas” pela variabilidade dos resultados. Neste trabalho, porém, o uso de uma técnica de reamostragem (bootstrap) aliada a um método estatístico tradicional (teste de hipótese) permitiu avaliar a confiabilidade dos resultados. Além disso, a inspeção dos mapas de pesos, que permitiu a extração de regras sobre os grupos formados, deixa claro que as críticas são infundadas ou podem ser contornadas.

Ficou comprovado que, tanto a ACP como o SOM foram capazes de diferenciar os sete tipos de café. As relações de vizinhança e a influência das variáveis originais também foram, de modo análogo, representadas por ambas as metodologias.

A combinação do nariz eletrônico com a metodologia de análise aplicada neste trabalho possibilita a utilização confiável do SOM no controle da qualidade aromático de café solúvel. Além disso, a análise dos resultados e a bibliografia consultada leva a crer que a combinação do SOM aos dados de cromatografia gasosa pode ser uma alternativa promissora na identificação de compostos de impacto aromático.

2.5 REFERÊNCIAS

- AZEVEDO, F. M.; BRASIL, L. M.; OLIVEIRA, R. C. L. **Redes neurais com aplicações em controle e em sistemas especialistas**. Florianópolis: Bookstore, 2000, 401p.
- BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of Microbiological Methods**, v. 43, p. 3-31, 2000.
- BASSOLI, D. G. **Impacto aromático dos componentes voláteis do café solúvel: uma abordagem analítica e sensorial**. 2006. 214 f. Tese (Doutorado em Ciência de Alimentos), Universidade Estadual de Londrina, Londrina, 2006.
- BODT, E.; COTTRELL, M.; VERLEYSEN, M. Statistical tools to assess the reliability of self-organizing maps. **Neural Networks**, v. 15, p. 967-978, 2002.
- BOISHEBERT, V.; GIRAUDEL, J. L.; MONTURY, M. Characterization of strawberry varieties by SPME-CG-MS and Kohonen self-organizing map. **Chemometrics and Intelligent Laboratory Systems**, v. 80, p. 13-23, 2006.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC, 2000. 262 p.
- CORNEY D. Food bytes: intelligent systems in food industry. **British Food Journal**, v. 104, n. 10, p. 787-805, 2002.
- CRAVEN, M. A.; GARDNER, J. W.; BARTLETT, P. N. Electronic noses: development and future prospects. **Trends in Analytical Chemistry**, v. 15, n. 9, p. 486-493, 1996.
- DEISINGH, A. K.; STONE, D. C.; THOMPSON, M. Applications of electronic noses and tongues in food analysis. **International Journal of Food Science and Technology**, v. 39, p. 587-604, 2004.
- DEMARTINES, P.; HÉRAULT, J. Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets. **IEEE Transactions on Neural networks**, v. 8, n. 1, 1997.
- DÍAZ, C. et al. Application of multivariate analysis and artificial neural networks for the differentiation of red wines from the Canary Islands according to the Island of origin. **Journal of Agricultural and Food Chemistry**, v. 51, p. 4303-4307, 2003.
- DUTTA, R. et al. Electronic nose based tea quality standardization. **Neural Networks**, v. 16, p. 847- 853, 2003.
- GOYACHE, F. et al. The usefulness of artificial intelligence techniques to assess subjective quality of products in the food industry. **Trends in Food Science & Technology**, v. 12, p. 370-381, 2001.

HAI, Z.; WANG, J. Electronic nose and data analysis for detection of maize oil adulteration in sesame oil. **Sens. and Actuators B**, v. 119, p. 449-455, 2006.

HAIR JR., J. F. et al. Análise Fatorial. In: _____. **Análise multivariada de dados**. 5. ed. Porto Alegre: Bookman, 2005. p. 89-127.

HAYKIN, S. **Redes neurais: princípio e práticas**. 2. ed. Porto Alegre: Bookman, 2001, 900p.

HUANG, Y.; KANGAS, L. J.; RASCO, B. A. Applications of artificial neural networks (ANNs) in food science. **Critical Reviews in Food Science and Nutrition**, v. 47, p. 113-126, 2007.

KOLEHMAINEN, M.; RÖNKKÖ, P.; RAATIKAINEN, O. Monitoring of yeast fermentation by ion mobility spectrometry measurement and data visualisation with self-organizing maps. **Analytica Chimica Acta**, v. 484, p. 93-100, 2003.

LU, H. SETIONO, R.; LIU, H. Effective data mining using neural networks. **IEEE Transactions on Knowledge and Data Engineering**, v. 8, n. 6, p. 957-961, 1996.

MARINI, F. et al. Use of different artificial neural networks to resolve binary blends of monocultivar Italian olive oils. **Analytica Chimica Acta**, v. 599, p. 232-240, 2007.

MARINI, F.; ZUPAN, J.; MAGRÌ, A. L. Class-modeling using Kohonen artificial neural networks. **Analytica Chimica Acta**, v. 544, p. 306-314, 2005.

MELLO, A. A.; TRUGO, L. C. Tipificação odorífera de compostos voláteis do café. In: Franco, M. R. B. (Ed.) **Aroma e sabor de alimentos**. São Paulo: Livraria Varela, 2003. p. 169-175.

MELSSSEN, W.; WEHRENS, R.; BUYDENS, L. Supervised Kohonen networks for classification problems. **Chemometrics and intelligent laboratory systems**, v. 83, p. 99-113, 2006.

PARDO, M.; SBERVEGLIERI, G. Coffee analysis with an electronic nose. **IEEE Transactions on Instrumentation and Measurement**, v. 51, n. 6, p. 1334-1339, 2002.

SCHALLER, E.; BOSSET, J. O.; ESCHER, F. 'Eletronic noses' and their application to food. **LWT**, v. 31, p. 305-316, 1998.

SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. Integração numérica. In: _____. **Cálculo numérico: características matemáticas e computacionais dos métodos numéricos**. São Paulo: Prentice Hall, 2003. p. 183-230.

SPIEGEL, M. R. **Estatística**. São Paulo: McGraw-Hill, 1974. 580p.

TIGRINE-KORDJANI, N. et al. Relative characterization of rosemary samples according to their geographical origins using microwave-accelerated distillation, solid-

phase microextraction and Kohonen self-organizing maps. **Analytical and Bioanalytical Chemistry**, v. 389, p. 631-641, 2007.

VERDINI, R. A. et al. Multivariate statistical methods for Port Salut Argentino cheese analysis base don ripening time conditions, and sampling sites. **Chemometrics and intelligent laboratory systems**, v. 86, p. 60-67, 2007.

ZHANG, H. et al. Optimization of sensor array and detection of stored duration of wheat by electronic nose. **J. of Food Engineering**, v. 82, p. 403-408, 2007.

ZUPAN, J.; NOVIČ, M.; RUISÁNCHEZ, I. Kohonen and counterpropagation artificial neural networks in analytical chemistry. **Chemometrics and Intelligent Laboratory Systems**, v. 38, p. 1-23, 1997.

2.6 AGRADECIMENTOS

Agradeço à Fundação Araucária pelo suporte financeiro, à CAPES pela bolsa de doutorado e à Companhia Iguaçu de Café Solúvel que, representada pelo Dr. Denisley Gentil Bassoli, forneceu todo o suporte experimental e analítico necessário para este trabalho.

CAPÍTULO III

3 REDE NEURAL OTIMIZADA PARA A CLASSIFICAÇÃO DE CAFÉ SOLÚVEL ATRAVÉS DE UM NARIZ ELETRÔNICO: I) DESENVOLVIMENTO PARA ALGUMAS AMOSTRAS DE CAFÉ SOLÚVEL

Este artigo será submetido para Sensors and Actuators B: Chemical (As normas para publicação neste periódico estão contidas no Anexo B)

Resumo

O aroma é uma das características mais importantes para alimentos em particular para o café. A avaliação desta característica sensorial é complexa, porém, primordial no controle de qualidade de cafés solúveis. Neste trabalho, foi desenvolvida uma rede neural artificial para classificar cafés solúveis com base no perfil aromático captado por um nariz eletrônico. Para este fim, foi construído um algoritmo híbrido contendo uma metodologia de reamostragem (bootstrap), uma heurística de escolha dos parâmetros de rede (planejamento fatorial e otimização simplex), uma rede neural (perceptron de múltiplas camadas treinado por retropropagação), uma média de ensemble (combinação das respostas de 100 redes treinadas) para aumentar a confiabilidade da classificação e, um método de extração de conhecimento (índice de casualidade). O emprego deste algoritmo híbrido produziu um classificador neural confiável que obteve êxito de 100% no reconhecimento dos cafés estudados. O emprego do índice de casualidade permitiu inferir algumas regras de como os cafés foram separados de acordo com os sensores disponíveis no nariz eletrônico. Os resultados obtidos indicam que a metodologia aplicada é promissora como uma ferramenta para o controle de qualidade.

Palavras-chave: Perceptron de múltiplas camadas. Bootstrap. Média de ensemble. Simplex. Funções de desejabilidade.

3.1 INTRODUÇÃO

O significado do aroma, como uma das características determinantes da qualidade de alimentos, torna sua avaliação objetiva uma das tarefas mais importantes na indústria de alimentos. Para o café, o aroma é o critério mais importante na avaliação de qualidade e um dos parâmetros determinantes na escolha pelo consumidor [1]. O aroma do café é formado por inúmeros componentes com variados grupos funcionais e essa composição depende de fatores como

espécie e variedade, condições de crescimento e colheita, armazenagem, intensidade de torra e tipo de torrador, além de outras condições de processo [2 - 4]. A busca pela aproximação do aroma do café solúvel ao do café de coador tem sido constante, e em geral, meta importante dos processos de aumento da intensidade dos aromas presentes no café solúvel ou mesmo da sua aromatização [5]. A percepção dos compostos voláteis do café solúvel pelo nariz de provadores especializados é uma técnica usada habitualmente na indústria de alimentos e bebidas [6]. Entretanto, nas últimas décadas tem sido grande o esforço no sentido de substituir, ou pelo menos auxiliar, a decisão humana por instrumentos que operem de maneira similar ao nariz de um especialista [7, 8]. O nariz eletrônico é um instrumento contendo um grupo de sensores para compostos gasosos, com especificidade parcial, acoplado a um sistema apropriado de reconhecimento de padrões capaz de identificar odores simples e complexos [9]. Existem trabalhos recentes de revisão [8 - 10] que relatam a aplicação dos narizes eletrônicos nas mais variadas áreas da indústria de alimentos. Além disso, trabalhos atuais vêm confirmando a versatilidade do instrumento na avaliação de aromas de alimentos: monitoramento da maturação de queijo [11]; identificação de carne bovina contaminada por microrganismos [12]; reconhecimento de tipos diferentes de óleos comestíveis e méis [13]; identificação de adulteração em óleo de gergelim [14]; estudo da liberação de aromas em géis de pectina [15]; classificação de trigo em função do tempo de estocagem [16]; avaliação de vinho [17]; padronização da qualidade de chá [18] e até mesmo na análise café torrado [19].

Durante a avaliação de um café por um provador treinado, seu nariz capta os odores liberados pela amostra e transforma em impulsos eletroquímicos que serão interpretados por neurônios especializados no córtex cerebral (Figura 1). Do mesmo modo, o nariz eletrônico deve estar acoplado a um sistema que possa emular o comportamento cerebral humano e reconhecer padrões associados ao perfil aromático da amostra [20]. A análise de componentes principais (ACP) e as redes neurais artificiais (RNA) estão entre as técnicas mais utilizadas para o reconhecimento de padrões a partir dos dados fornecidos pelo nariz eletrônico [9, 10]. A maior vantagem das RNA sobre a ACP é a possibilidade de trabalhar com problemas mais complexos e não-lineares [21, 22]. As RNA tem sido alvo de recentes pesquisas nas mais diversas áreas, sendo que Haykin (2001) apresenta um amplo texto sobre o assunto abordando tanto a implementação como as áreas

de aplicação. Entretanto, essa metodologia é criticada por ser essencialmente do tipo “caixa preta” e pela dificuldade na escolha dos melhores parâmetros da rede que vão influenciar diretamente na qualidade do classificador obtido [24]. Com objetivo de superar o problema da escolha de parâmetros tem sido empregado metodologias de otimização [25] tais como planejamentos de experimentos [26], algoritmos de busca direta como o simplex seqüencial [27] e o algoritmo genético [28].

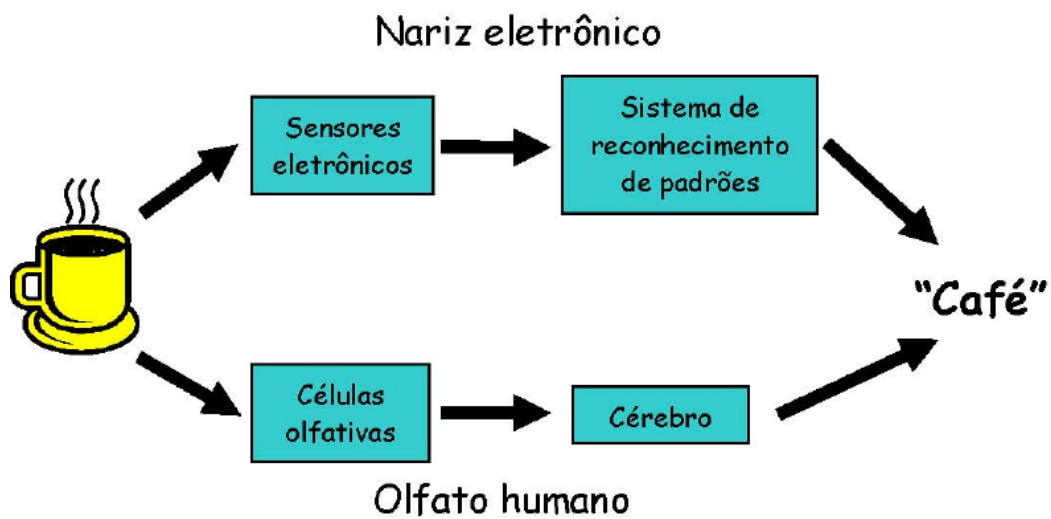


Figura 1 – Representação esquemática do olfato humano e nariz eletrônico, adaptado de Craven et al. (1996).

O objetivo deste trabalho foi construir e utilizar uma rede neural artificial para classificar café solúvel a partir dos perfis aromáticos obtidos por um nariz eletrônico. Além disso, foi aplicada uma heurística de busca dos parâmetros ótimos da rede e outra para aumentar a confiabilidade e repetibilidade do classificador neural produzido.

3.2 MATERIAIS E MÉTODOS

3.2.1 Café Solúvel

Foram usados sete tipos de café solúvel (Tabela 1), sendo três amostras experimentais produzidas na planta piloto do laboratório de pesquisa e desenvolvimento da Companhia Iguaçu de Café Solúvel (Cornélio Procópio – Paraná). As demais amostras são cafés solúveis comerciais adquiridos no varejo do Brasil e Inglaterra [5].

Tabela 1 – Descrição simplificada das amostras de cafés solúveis analisadas.

Amostra	Origem	Classificação	Composição	Intensidade de torração
1 ^a	Brasil	Experimental	Arábica	Média
2 ^b	Brasil	Experimental	Arábica	Média
3 ^c	Brasil	Experimental	Conilon	Média
4	Brasil	Comercial	Conilon/arábica	Escura
5	Brasil	Comercial	Desconhecida	Escura ^d
6	Inglaterra	Comercial	Desconhecida	Média ^d
7	Inglaterra	Comercial	Desconhecida	Média/clara ^d

^a Café, classificado como bebida dura, proveniente da região sul de Minas Gerais.

^b Café, classificado como bebida dura, proveniente da região norte do Paraná.

^c Café proveniente do estado de Rondônia.

^d Intensidade de torração estimada por avaliação sensorial.

3.2.2 Nariz Eletrônico

Todas as análises foram realizadas no laboratório de pesquisa e desenvolvimento da Companhia Iguaçu de Café Solúvel. Foi utilizado o equipamento Portable Electronic Nose Type PEN2 (Airsense Analytics, Alemanha), que monitora a variação da condutância em cada um de seus sensores (Tabela 2) conforme a passagem do fluxo gasoso com a amostra [14, 16]. Essa variação é então comparada com aquela obtida pelo branco (ar ambiente) gerando, assim, um sinal adimensional.

No preparo das amostras 2,8g de café solúvel foram colocados em frascos específicos para análise de headspace. Cada café (Tabela 1) foi avaliado em

três repetições genuínas. Os parâmetros operacionais do equipamento foram [5]: 1,0 s de intervalo entre as amostras; 60 s de tempo de medição; 120s para limpeza ou purga dos sensores; 400 mL/min de fluxo de injeção.

Tabela 2 – Especificidade dos sensores utilizados no nariz eletrônico PEN2.

Sensor	Especificidade
S1	Compostos aromáticos
S2	Amplo espectro, muito sensível, reage com óxidos de nitrogênio
S3	Amônia, usado para compostos aromáticos
S4	Hidrogênio
S5	Alcanos e compostos aromáticos menos polares
S6	Metano, amplo espectro
S7	Compostos sulfurados, terpenos, limonenos e pirazinas
S8	Álcoois, amplo espectro
S9	Compostos orgânicos sulfurados
S10	Metanos alifáticos em altas concentrações > 100 ppm

3.2.3 Normalização

Após a aquisição dos dados, a área abaixo da curva gerada por cada sensor foi numericamente integrada, formando um vetor de entrada com 10 valores médios do sinal analítico para cada amostra. Antes de ser alimentado na rede neural os vetores de entrada foram pré-processados para que as funções de ativação dos neurônios artificiais não sejam facilmente saturadas ou ocorra um erro de overflow [23]. Os métodos de pré-processamento usados foram: máximo e mínimo (minimax), transformação para uma escala entre 0 e 1; autoescalamento, vetor de entrada com média zero e variância unitária; vetor unitário, norma euclidiana igual a 1 [21].

3.2.4 Rede Neural Artificial

Foi empregada uma rede neural artificial do tipo perceptron de múltiplas camadas (MLP – multi-layer perceptron) que é muito utilizada para a classificação de padrões em conjunto com o nariz eletrônico [9 -15]. A arquitetura

deste tipo de rede (Figura 2) é composta por uma camada de entrada com 10 neurônios (um para cada sensor do nariz eletrônico), uma ou duas camadas intermediárias (responsáveis pela separação dos padrões através da formação de fronteiras de decisão) contendo uma quantidade de neurônios a ser definida e uma camada de saída com 7 neurônios, um para cada tipo de café, que constroem combinações lineares das fronteiras de decisão formadas pelos neurônios ocultos [23].

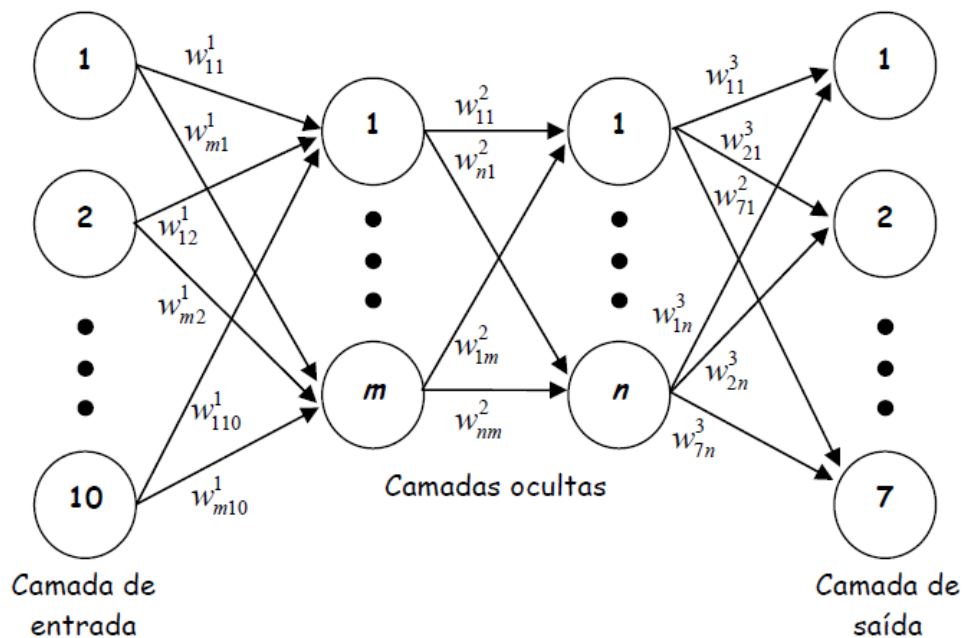


Figura 2 – Representação gráfica da rede perceptron de múltiplas camadas com duas camadas ocultas. Para cada conexão entre os neurônios existe um peso w_{ij} associado, sendo i o índice do neurônio de entrada do sinal, j o neurônio de saída do sinal e l camada onde esta localizado o neurônio de entrada.

Em cada neurônio da rede MLP é realizado um somatório ponderado pelos pesos sinápticos dos sinais provenientes dos neurônios da camada anterior. Esta soma, chamada de campo local induzido, é aplicada a uma função de ativação não linear que irá produzir a saída do neurônio [23]. Nas redes MLP as funções de ativação mais utilizadas são as sigmóides do tipo logística e tangente hiperbólica [29].

Nos algoritmos computacionais tradicionais, o conhecimento é construído de forma explícita sob a forma de regras definidas a priori. Já as redes neurais são sistemas capazes de fazer generalizações, ou seja, geram suas próprias regras, para associar as variáveis de entrada às de saída, a partir do aprendizado com dados de treinamento [13]. Isto é possível através de uma regra, algoritmo de aprendizagem, que muda os pesos das conexões entre os neurônios em função das respostas obtidas a partir das variáveis de entrada. Para a rede construída foi utilizado o algoritmo de retropropagação [23] que consiste de dois passos através das diferentes camadas da rede: um passo a frente, a propagação onde é calculada a resposta da rede para um dado padrão de entrada, e um passo atrás, a retropropagação cujo objetivo é ajustar os pesos em função do erro quadrático médio em relação à resposta esperada. Conforme descrito anteriormente (item 3.2.3.), o vetor de entrada é formado por 10 valores, já o vetor de resposta desejada é constituído por um vetor de dimensão igual a 7 sendo que, para uma amostra pertencente a classe k o k -ésimo valor é igual a 1 e todos os outros são zerados. A correção dos pesos, além do erro cometido pela rede, depende de uma constante de momento e de uma taxa de aprendizagem. Os valores destes parâmetros devem ser otimizados para cada problema e em geral a escolha é feita através de tentativa e erro [28 - 30].

Uma apresentação completa do conjunto de treinamento é denominada época e o processo de aprendizagem é mantido até os pesos sinápticos e os níveis de bias se estabilizarem e o erro quadrático médio convergir para um valor mínimo [31]. Para a rede empregada, foi utilizado o modo seqüencial de treinamento, uma vez que, os pesos são atualizados a cada novo exemplo apresentado a rede. Para todas as redes utilizadas neste estudo foram utilizadas 100 épocas para o treinamento.

Tendo como objetivo uma boa generalização, foi utilizado o método de treinamento com parada antecipada para evitar que a rede acabe sendo excessivamente ajustada aos dados de treinamento. É possível identificar o início do excesso de treinamento através do uso de amostras de validação, onde além dos dados de treinamento é usado um conjunto de teste que após cada época é submetido à rede para verificar a capacidade de generalização desta com dados não utilizados para o ajuste dos pesos. Tipicamente, o modelo não funciona tão bem sobre o conjunto de teste quanto sobre o de treinamento, em relação ao qual foi

baseado. A curva do erro quadrático médio, também chamada de curva de aprendizagem, para o treinamento decresce monotonamente para um número crescente de épocas. Já para o conjunto de teste, esta curva decresce para um mínimo e então começa a crescer conforme o treinamento continua, ou seja, o que a rede aprende após este ponto é essencialmente o ruído contido nos dados de treinamento. Nesta heurística, o ponto mínimo na curva de aprendizagem de teste é usado como o critério de parada ou armazenamento do conjunto de pesos [23].

3.2.5 Bootstrap

O bootstrap (reamostragem) consiste na obtenção de medidas de dispersão, (como a média, variância ou desvio-padrão) através da aplicação de métodos tradicionais, de inferência e estimação, a diversos conjuntos de dados obtidos a partir de um único conjunto original mediante a geração de dados aleatórios com a mesma média e desvio padrão dos originais [32]. Como mencionado no item 3.2.2, foram analisados sete tipos de café em triplicatas genuínas, totalizando 21 amostras, que é uma quantidade pequena para ser subdividida em um conjunto de treinamento e outro de teste. Usando a metodologia de bootstrap foram criadas mais 63 amostras a partir do conjunto original, sendo estas utilizadas para o treinamento e as 21 originais para o teste.

3.2.6 Média de Ensemble

Uma maneira de melhorar a capacidade de generalização das redes neurais é a combinação das respostas de várias RNAs, também chamada de máquinas de comitê, as quais foram treinadas de maneira semelhante de modo que o resultado final seja melhor do que aquele obtido ao ser utilizada apenas uma única RNA. Uma maneira fácil de implementar o comitê de máquinas é a média de ensemble, que consiste em combinar as respostas das redes treinadas através de uma média. O uso de um ensemble diminui a probabilidade de erro final, pois será coberto um espaço maior de hipóteses dentro das possíveis soluções para o problema. A desvantagem deste método reside no aumento do tempo

computacional, porém, em problemas de pequeno porte como este, o acréscimo não é significativo [23, 32]. Neste trabalho foi utilizada uma média de ensemble com 100 redes, sendo que para cada rede um novo conjunto de pesos iniciais e de exemplos de treinamento era gerado aleatoriamente.

3.2.7 Otimização dos Parâmetros de Rede

Foi construído um planejamento fatorial 2 x 3 [33] com 5 repetições de cada experimento para avaliar os parâmetros qualitativos, função de ativação e pré-processamento, os quais, supostamente, influenciam o erro quadrático médio da rede MLP, treinada com o algoritmo de retropropagação (Tabela 3).

Tabela 3 – Planejamento fatorial 2 x 3 com 5 repetições para cada experimento.

Experimento	Função de Ativação	Normalização
1	Tangente hiperbólica	Autoescalamento
2	Tangente hiperbólica	Minimax
3	Tangente hiperbólica	Vetor unitário
4	Logística	Autoescalamento
5	Logística	Minimax
6	Logística	Vetor unitário

Antes de serem realizados os experimentos apresentados na Tabela 3, cada uma das 6 redes propostas tiveram os outros parâmetros otimizados através do simplex seqüencial [34]. Os parâmetros escolhidos para otimização foram: a quantidade de neurônios na primeira e segunda camada ocultas, taxa de aprendizagem, constante de momento e a inclinação da função de ativação logística. A otimização foi realizada visando maximizar a porcentagem de classificação correta e minimizar o erro quadrático médio para o conjunto de teste. Além disso, estas metas deveriam ser alcançadas com a menor rede possível, entende-se por isso aquela que apresentar a menor dimensão para o vetor de pesos [23]. Para combinar estas três exigências em uma única resposta global, que será maximizada pelo método simplex seqüencial, foram utilizadas as funções de desejabilidade [35].

3.2.8 Índice de Casualidade (IC)

As redes neurais artificiais tem sido questionadas por exibirem uma metodologia do tipo caixa preta. Mais recentemente, alguns autores estão sugerindo metodologias para extrair informação de uma rede treinada e tornar as RNA mais confiáveis e com maiores aplicações [28, 36 - 38]. Neste artigo foi utilizado o método do índice de casualidade [39] para avaliar a influência de cada sensor na classificação das amostras da rede treinada. O índice de casualidade (IC) é definido como a multiplicação de todos os pesos que ligam uma determinada entrada a uma determinada saída. Usando notação matricial o IC é definido como:

$$\mathbf{IC} = \prod_{k=1}^n \mathbf{W}^{(k)} \mathbf{W}^{(k+1)} \quad (1)$$

onde n é o número de camadas ocultas; $\mathbf{W}^{(k)}$ é a matriz de pesos para a k -ésima camada e \mathbf{IC} é uma matriz de dimensões $i \times o$, sendo i o número de variáveis de entrada e o a quantidade de variáveis de saída. Para uma rede com duas entradas, uma camada oculta com dois neurônios e duas saídas (Figura 3) teríamos:

$$\begin{aligned} \mathbf{IC} &= \mathbf{W}^{(1)} \mathbf{W}^{(2)} = \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} w_{11}^{(2)} & w_{12}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} \end{bmatrix} \\ \mathbf{IC} &= \begin{bmatrix} w_{11}^{(1)} w_{11}^{(2)} + w_{12}^{(1)} w_{21}^{(2)} & w_{11}^{(1)} w_{12}^{(2)} + w_{12}^{(1)} w_{22}^{(2)} \\ w_{21}^{(1)} w_{11}^{(2)} + w_{22}^{(1)} w_{21}^{(2)} & w_{21}^{(1)} w_{12}^{(2)} + w_{22}^{(1)} w_{22}^{(2)} \end{bmatrix}. \end{aligned}$$

O procedimento pode ser facilmente generalizado para redes maiores. Índices de casualidade positivos indicam que a variável de entrada aumenta a resposta da rede, já os valores negativos indicam uma redução. Além disso, quanto maior for o índice de casualidade maior será a influência de uma variável de entrada para a classificação de uma amostra em um determinado grupo [21].

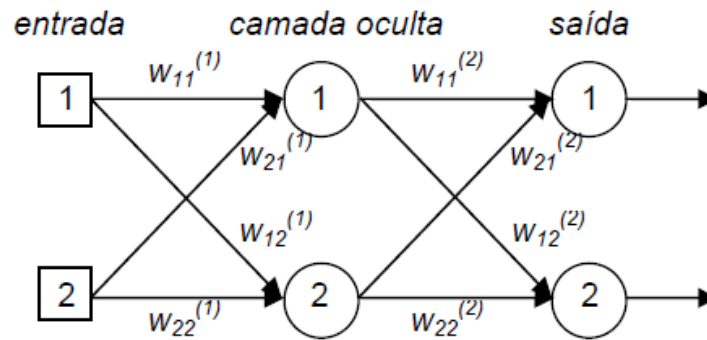


Figura 3 – Representação de uma RNA com duas entradas, duas saídas e uma camada oculta de dois neurônios.

3.2.9 Implementação Computacional

A integração numérica dos dados obtido pelo nariz eletrônico, o pré-processamento, todas as redes e o procedimento de otimização usando o simplex seqüencial acoplado às funções de desejabilidade foram desenvolvidos em linguagem FORTRAN para Windows. As análises estatísticas ANOVA, comparação de médias (Tukey) e análise de agrupamentos (distâncias euclidianas – algoritmo complete linkage) foram realizadas no software Statistica 7.1. Todos os cálculos foram conduzidos em um computador com processador Intel Celeron D[®] de 2,8 GHz e 1,0 Gb de memória RAM.

3.3 RESULTADOS E DISCUSSÃO

Os resultados para as redes propostas (Tabela 3) com seus parâmetros otimizados usando o método simplex seqüencial encontram-se na Tabela 4. Os parâmetros ótimos obtidos para cada rede não podem ser diretamente comparados com outros valores da literatura uma vez que são extremamente influenciados pelas condições do problema [26]. Porém, é possível fazer algumas importantes considerações.

Como se observa na Tabela 4, a época média onde ocorre um mínimo no erro médio quadrático de validação é inferior a 100 indicando que 100 épocas de treinamento são suficientes. Com relação à arquitetura, todas as redes

otimizadas possuem duas camadas ocultas, sendo que, a quantidade total de neurônios variou de 7 a 12. É notório também que, as maiores redes (10, 11 e 12 neurônios) foram aquelas que usavam a tangente hiperbólica como função de ativação. Em Basheer e Hajmeer (2000) é listada uma revisão sobre trabalhos que apresentam critérios empíricos sobre a quantidade de neurônios necessária. Para o presente problema, com 10 variáveis de entrada e 7 de saída, a quantidade suficiente, de acordo com esses critérios, seria de 8 a 11 neurônios. Como os valores obtidos estão próximos do intervalo sugerido e as redes (experimentos de 1 a 4) apresentaram 100% de classificação correta, fica evidente que a metodologia simplex foi eficaz na otimização dos parâmetros de arquitetura e na minimização do tamanho da rede mantendo um bom desempenho. É conhecido que é menos provável que uma rede neural com tamanho mínimo aprenda as idiossincrasias ou ruído dos dados de treinamento, podendo assim generalizar melhor sobre novos dados [23, 25, 40].

Tabela 4 – Parâmetros ótimos obtidos para as diferentes redes propostas na Tabela 3

Função de Ativação	Pré-processamento					
	Minimax		Autoescalonamento		Vetor unitário	
	Tangente Hiperbólica	Função Logística	Tangente Hiperbólica	Função Logística	Tangente Hiperbólica	Função Logística
Experimento	1	2	3	4	5	6
Parâmetros otimizados usando o Simplex seqüencial						
Neurônios na camada 1	2	3	4	4	4	3
Neurônios na Camada 2	9	4	6	4	8	4
Taxa de Aprendizagem	0,0426	0,9500	0,0476	0,7554	0,0558	0,5159
Constante de Momento	0,2850	0,5995	0,3343	0,3168	0,3354	0,3285
Inclinação	-	1,1270	-	1,4131	-	1,8436
Desempenho e tamanho da rede otimizada						
Erro Quadrático	0,0701 ^d	0,0145 ^b	0,0345 ^c	0,0090 ^a	0,1724 ^f	0,0968 ^e
Médio(ensemble) [*]	± 0,0047	± 0,0004	± 0,0015	± 0,0005	± 0,0037	± 0,0026
Erro Quadrático	0,1689 ^d	0,0355 ^a	0,1070 ^b	0,0249 ^a	0,2107 ^e	0,1310 ^c
Médio (média) ^{**}	± 0,1791	± 0,0271	± 0,0980	± 0,0177	± 0,0393	± 0,0460
%Classificação correta(ensemble) [*]	100,00 ^c	100,00 ^c	100,00 ^c	100,00 ^c	90,47 ^a	94,29 ^b
% Classificação correta (média) ^{**}	± 0,00	± 0,00	± 0,00	± 0,00	± 3,37	± 2,13
Pesos sinápticos	82,49 ^c	95,64 ^e	87,46 ^d	97,20 ^e	77,57 ^b	71,80 ^a
Época mínima ^{***}	± 14,34	± 7,46	± 10,98	± 5,30	± 8,75	± 15,22
	119	84	123	99	147	84
	93,50	92,45	91,55	86,33	93,04	94,70
	± 7,24	± 9,95	± 9,33	± 19,01	± 6,66	± 7,25

^{*} Média e desvio padrão para 5 repetições do ensemble.

^{**} Média e desvio padrão das 500 redes que compõem as 5 repetições do ensemble.

^{***} Época onde ocorreu um mínimo no erro quadrático médio para as amostras de teste. Média e desvio padrão das 500 redes que compõem as 5 repetições do ensemble.

Médias acompanhadas com a mesma letra não apresentam diferença significativa em nível de 5%.

Para os parâmetros do algoritmo de retropropagação, taxa de aprendizagem e constante de momento, observou-se que nas redes com tangente hiperbólica, é necessária uma taxa de aprendizagem sensivelmente menor, uma ordem de grandeza, do que àquelas com função logística (Tabela 4). A constante de momento foi pouco influenciada pela escolha da função de ativação ou metodologia de normalização visto que seu valor ficou em torno de 0,3 para todas as redes com exceção de uma. Em relação à inclinação da função logística não foi encontrada nenhuma referência sobre a utilização de valores diferentes de 1,0. O que se pode afirmar é que um aumento da inclinação provoca um aumento de sensibilidade da resposta do neurônio em relação ao seu campo local induzido [23].

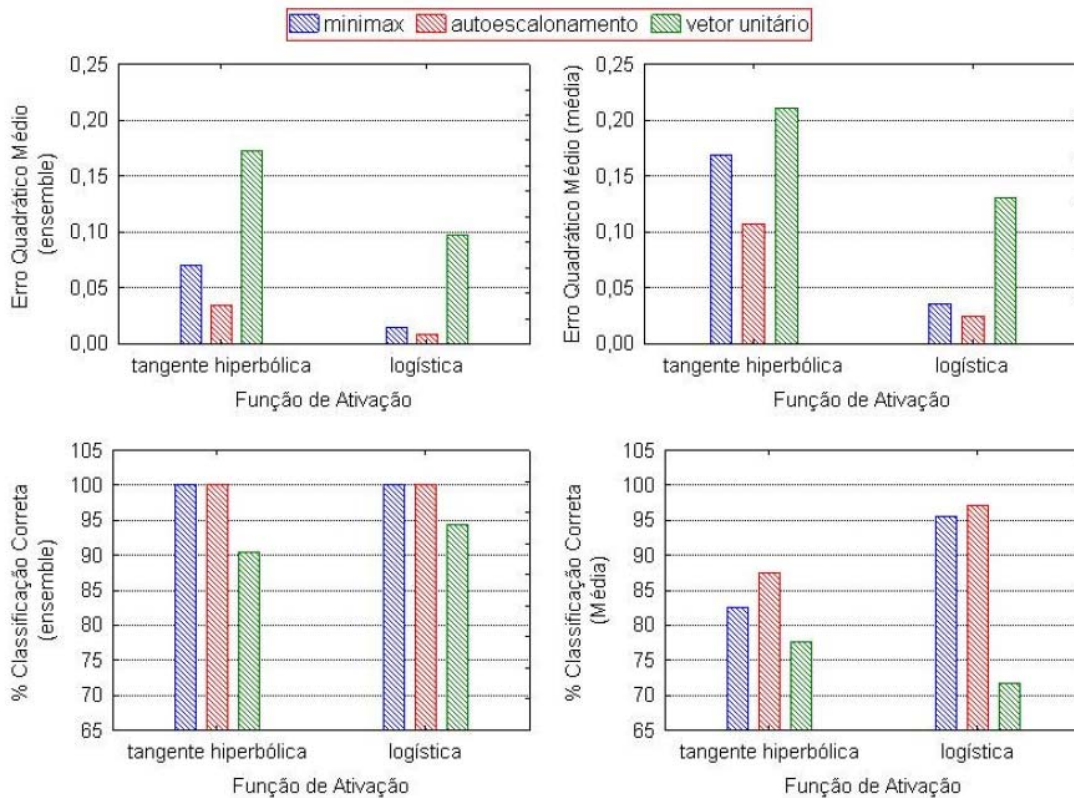


Figura 4 – Erro quadrático médio e porcentagem de classificação das redes otimizadas.

A partir da análise da Tabela 4 e Figura 4 observa-se uma redução no erro quadrático médio e um aumento na porcentagem de classificação correta quando são comparadas as respostas médias e do ensemble, além disso, é evidente a redução na dispersão dos resultados. Existem, portanto, fortes evidências

de que o uso da resposta do ensemble melhora desempenho e aumenta a precisão da rede que se torna estável e convergente, duas qualidades importantes em um modelo neural [40]. Destaca-se, também, que as redes com função logística apresentaram um erro quadrático médio menor e uma porcentagem de classificação igual ou superior em relação àquelas com tangente hiperbólica. A Tabela 4 e Figura 4 mostram seguramente que a função logística é a melhor escolha como função de ativação para este problema. Em Torrecilla et al. (2007), também, foi determinado que a função logística apresenta um melhor desempenho em uma rede usada para prever a concentração de compostos poliaromáticos. Entretanto, em Sukthomya e Tannock (2005) foi determinado que a função de ativação do tipo tangente hiperbólica apresenta um desempenho melhor. A observação dos resultados deste trabalho e dos estudos citados anteriormente permite reforçar que, os melhores parâmetros de uma rede são dependentes do problema. Portanto, a utilização de uma metodologia de otimização [25] é fundamental para encontrar o melhor conjunto de parâmetros que irá permitir a construção de uma rede neural eficiente e confiável.

Com relação à normalização, ficou evidente que o método minimax e de autoescalamento são superiores ao vetor unitário. Ou seja, as melhores redes seriam as representadas pelos experimentos 2 ou 4. A rede 4 apresenta um desempenho ligeiramente superior, entretanto, a rede 2 possui menos neurônios originando um modelo mais simples e com maior capacidade de generalização [42]. Para realizar uma melhor escolha entre estas duas redes foram analisadas as respostas de cada uma delas para verificar qual classifica as amostras com maior certeza (Figura 5).

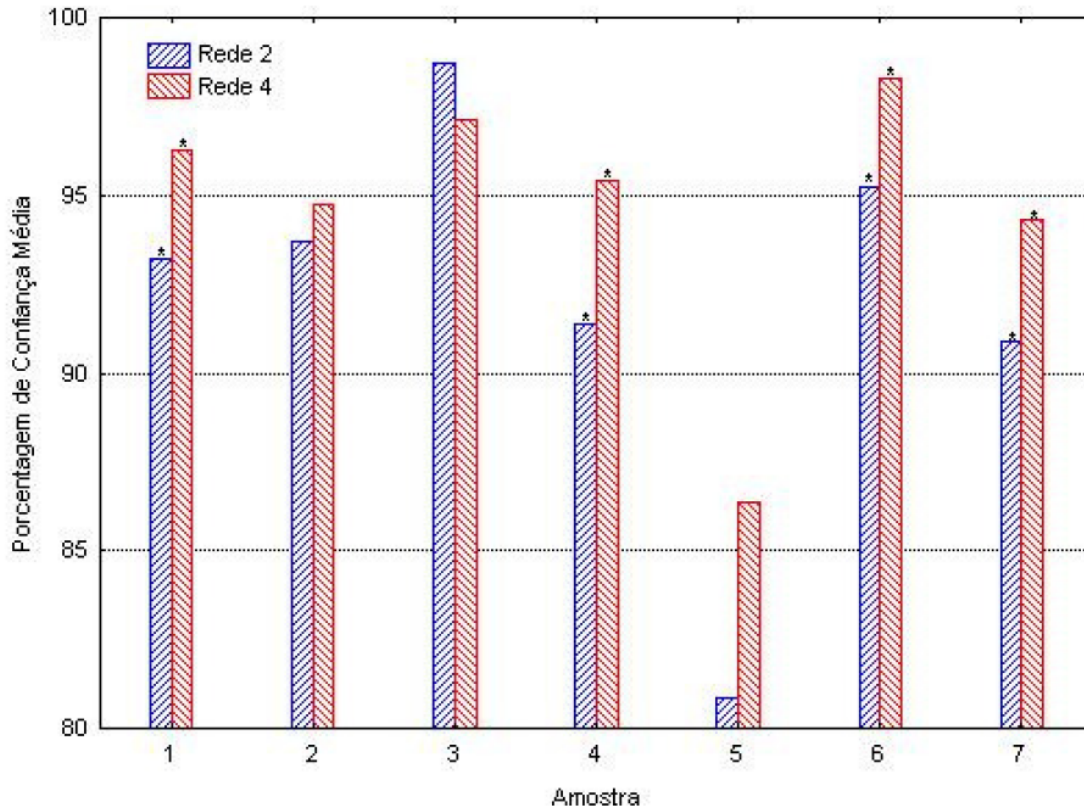


Figura 5 – Porcentagem de confiança média na classificação realizada pelas redes 2 e 4. A média foi retirada sobre 15 observações (3 repetições de cada amostra X 5 repetições do ensemble de cada rede). Existe diferença significativa ($p < 0,05$) entre as redes nas colunas marcadas com asterisco.

Na Figura 5 é possível observar que para todas as amostras de café (exceto a 5) a porcentagem média de confiança na classificação realizada foi superior a 90%, confirmando a qualidade dos classificadores neurais obtidos. Destaca-se também que, exceto para a amostra 3, a confiança média na classificação é superior para a rede 4, principalmente para a amostra 5 que foi a mais difícil de ser classificada. Além disso, a confiança média entre todas as amostras foi de 92,00% para a rede 2 e de 94,64% para a 4, sendo estas médias estatisticamente diferentes ($p < 0,05$). Portanto, apesar da rede 4 ser ligeiramente maior que a 2, foi escolhida como melhor classificador neural a primeira por apresentar uma confiança média maior.

Após a definição da rede 4 como a melhor, foi calculado o índice de casualidade médio (ICM) para as 5 repetições de ensemble desta rede. Os resultados obtidos encontram-se na Tabela 5. Para a amostra 1 os sensores que mais contribuem são S2 e S9. Fazendo um gráfico (Figura 6 – Gráfico D) de

dispersão das áreas médias (Tabela 6) para estes sensores é possível observar uma nítida separação da amostra em questão. Além disso, a Tabela 6 mostra que as áreas destes sensores para a amostra 1 são estatisticamente maiores que para as outras redes. Ou seja, a rede “aprendeu” que a amostra 1 destaca-se das outras por apresentar um sinal maior para os sensores S2 e S9 do nariz eletrônico. Um reflexo do fácil agrupamento da amostra 1 está implícito na alta porcentagem de confiança de classificação (Figura 5) apresentada pela rede 4.

Tabela 5 – Índice de casualidade médio para a rede 4 .

Sensor	Amostra						
	1	2	3	4	5	6	7
S1	-11,12 ^{d,e}	-23,75 ^e	27,33 ^a	30,96 ^b	5,69 ^{b,c}	-23,78 ^f	-9,13 ^f
S2	40,55 ^a	-2,15 ^b	18,82 ^c	-39,69 ^g	-0,07 ^c	-23,82 ^f	2,80 ^e
S3	-16,85 ^{e,f}	-16,75 ^{c,d,e}	19,01 ^c	32,40 ^b	1,69 ^c	-14,63 ^e	-5,42 ^f
S4	-3,98 ^c	-38,11 ^f	-11,19 ^e	20,45 ^c	17,85 ^a	6,18 ^d	11,02 ^d
S5	-9,42 ^{c,d}	-21,85 ^{d,e}	26,43 ^{a,b}	26,98 ^{b,c}	9,38 ^b	-25,01 ^f	-9,55 ^f
S6	-10,62 ^{c,d}	8,34 ^a	-20,25 ^f	-2,72 ^d	-30,73 ^f	33,57 ^b	21,99 ^c
S7	29,13 ^b	-13,03 ^{c,d}	6,92 ^d	0,42 ^d	-16,40 ^e	-40,32 ^g	25,30 ^{b,c}
S8	-22,99 ^f	-12,41 ^c	-37,68 ^g	10,69 ^d	-7,79 ^d	44,88 ^a	30,34 ^{a,b}
S9	44,70 ^a	0,73 ^{a,b}	20,84 ^{b,c}	-17,73 ^f	-25,91 ^f	-24,53 ^f	-3,64 ^{e,f}
S10	-18,34 ^f	-78,21 ^g	-43,00 ^g	72,39 ^a	21,52 ^a	15,19 ^c	34,05 ^a

* Média referente as 5 repetições de ensemble da rede 4. As médias pertencentes a mesma coluna e acompanhadas da mesma letra não apresentam diferença significativa ao nível de 5%.

Para a amostra 2 os ICMs foram menores em relação a amostra 1 e, os sensores de maior destaque foram S2, S6 e S9. Observa-se na Figura 6 (Gráficos B e C) que a separação não é tão eficaz quanto para a amostra 1, principalmente em relação a amostra 5 que localiza-se muito próxima. Na Figura 5 é demonstrado que a confiança na classificação é menor, corroborando com a dificuldade de separação entre as amostras 2 e 5. Mais uma vez fica claro que o “aprendizado” da rede conseguiu captar as características das amostras e com auxílio do índice de casualidade é possível detectar uma certa semelhança de perfil aromático, para todos sensores exceto S4 e S10 (ver Tabela 6), entre as amostras 2 e 5.

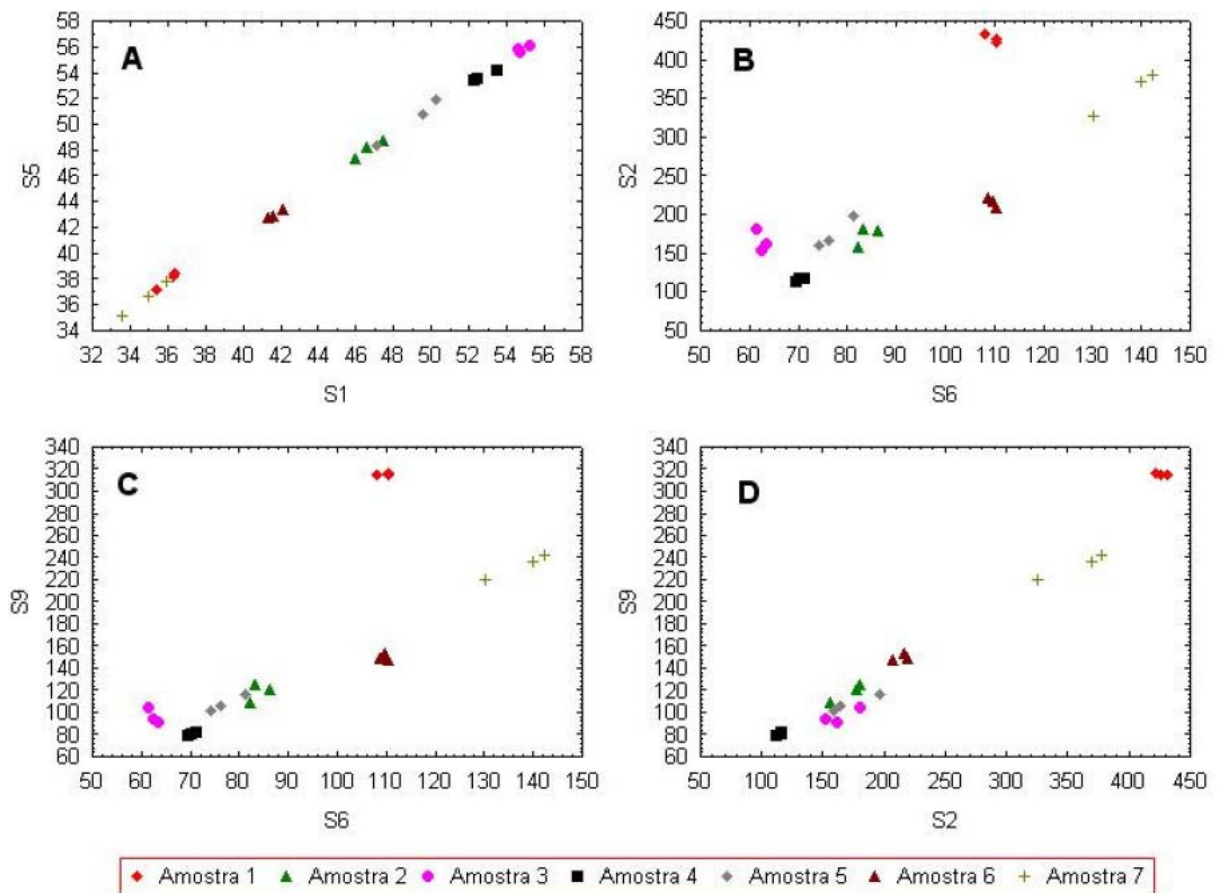


Figura 6 – Gráficos de dispersão das áreas médias de alguns sensores (Tabela 6).

Para a amostra 3 os sensores S1 e S5 foram os de maior ICM. O sinal destes (Tabela 6) não difere da amostra 4 e é superior ($p < 0,05$) as outras. Esta observação é reforçada pela inspeção no gráfico de dispersão (Figura 6 – Gráfico A) que mostra uma proximidade entre as amostras 3 e 4. Além disso é possível ver uma forte correlação entre os sinais dos sensores S1 e S5. Este comportamento já foi observado em outro trabalho onde foi usado um nariz eletrônico similar [14].

Tabela 6 – Comparação entre as áreas médias de cada sensor para cada amostra .

Sensor	Amostra						
	1	2	3	4	5	6	7
S1	36,03 ^d	46,70 ^b	54,84 ^a	52,78 ^a	48,98 ^b	41,71 ^c	34,84 ^d
S2	427,08 ^a	172,18 ^d	165,00 ^d	114,96 ^e	174,08 ^{c,d}	214,93 ^c	358,12 ^b
S3	32,86 ^d	45,27 ^b	52,39 ^a	51,04 ^a	47,03 ^b	40,47 ^c	33,60 ^d
S4	59,37 ^{a,b}	57,00 ^c	57,37 ^c	59,00 ^{a,b}	58,77 ^b	59,20 ^{a,b}	60,10 ^a
S5	37,88 ^d	48,05 ^b	55,74 ^a	53,70 ^a	50,27 ^b	42,98 ^c	36,54 ^d
S6	109,88 ^b	84,03 ^c	62,56 ^e	70,40 ^{d,e}	77,35 ^{c,d}	109,86 ^b	137,56 ^a
S7	141,15 ^a	70,39 ^b	65,07 ^b	61,31 ^b	70,27 ^b	68,22 ^b	134,53 ^a
S8	85,41 ^c	67,87 ^d	52,80 ^e	65,35 ^d	69,31 ^d	93,30 ^b	114,51 ^a
S9	314,45 ^a	117,74 ^d	96,18 ^{e,f}	79,86 ^f	107,14 ^{d,e}	149,10 ^c	232,04 ^b
S10	59,61 ^{a,b}	54,14 ^c	53,68 ^c	60,06 ^{a,b}	58,75 ^{a,b}	59,74 ^{a,b}	62,22 ^a

As médias pertencentes a mesma linha e acompanhadas da mesma letra não apresentam diferença significativa ao nível de 5%.

Na amostra 4 os sensores de maior ICM foram S10 e S3. Na Figura 7 (Gráfico D), está representado o diagrama de dispersão da área média para estes sensores. É possível observar uma separação de todas as amostras.

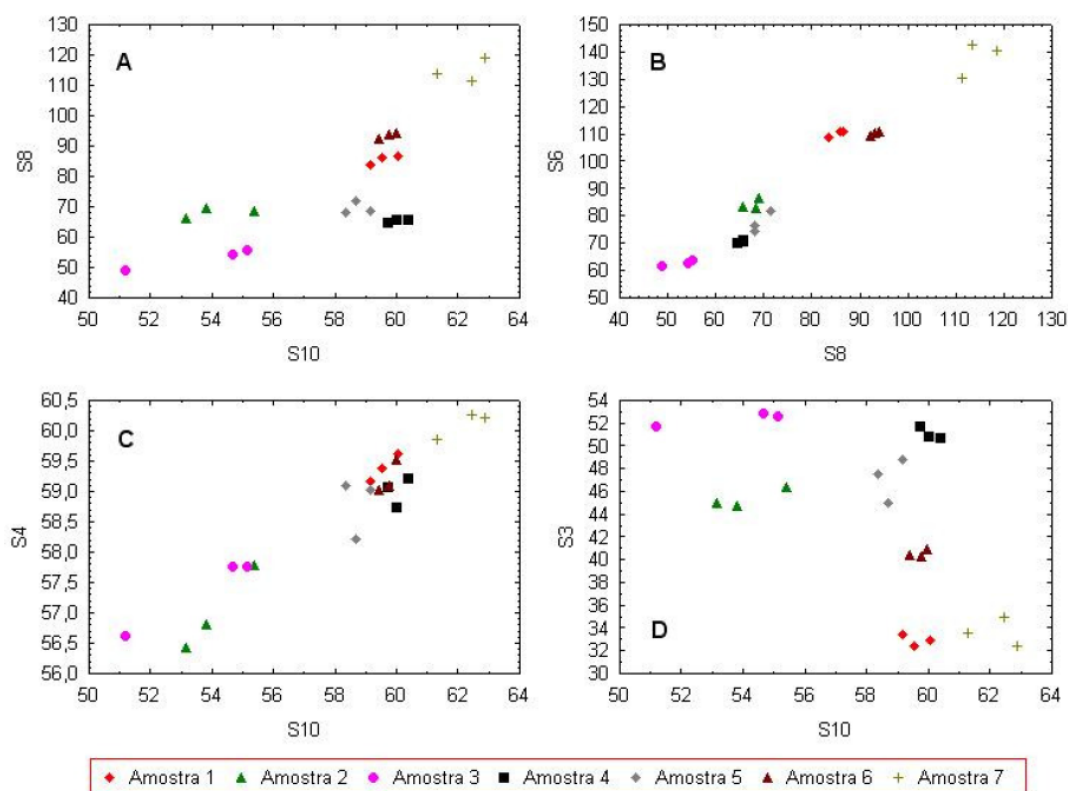


Figura 7 – Gráficos de dispersão das áreas médias de alguns sensores (Tabela 6).

Para a amostra 5 já foi destacada uma semelhança com a amostra 2, tanto que, a porcentagem de confiança na classificação (Figura 5) foi a menor entre todas as amostras. Os sensores de maior ICM (Tabela 5) foram S10 e S4. Comparando os sinais de cada sensor (Tabela 6) para as amostras 2 e 5 é possível verificar que os únicos sensores com sinais estatisticamente diferentes ($p < 0,05$) também são S10 e S4. Novamente é notório que a rede conseguiu “reconhecer” uma regra para diferenciar as duas amostras. Na Figura 7 (Gráfico C) nota-se a separação entre as amostras. Na mesma Figura (Gráficos A e D), fica claro que o sensor S10 é o maior responsável pela diferenciação entre as amostras 2 e 5. Para demonstrar a complexidade de separação entre as amostras 2 e 5 foi construído um dendrograma usando as distâncias euclidianas (Figura 8). Através da análise das distâncias euclidianas a amostra 13 (café 5) é erroneamente classificada como café 2. Assim como, a amostra 6 (café 2) é classificada como café 5. Esta evidência comprova que a separação entre as amostras do tipo 2 e 5 não é tão trivial de ser realizada.

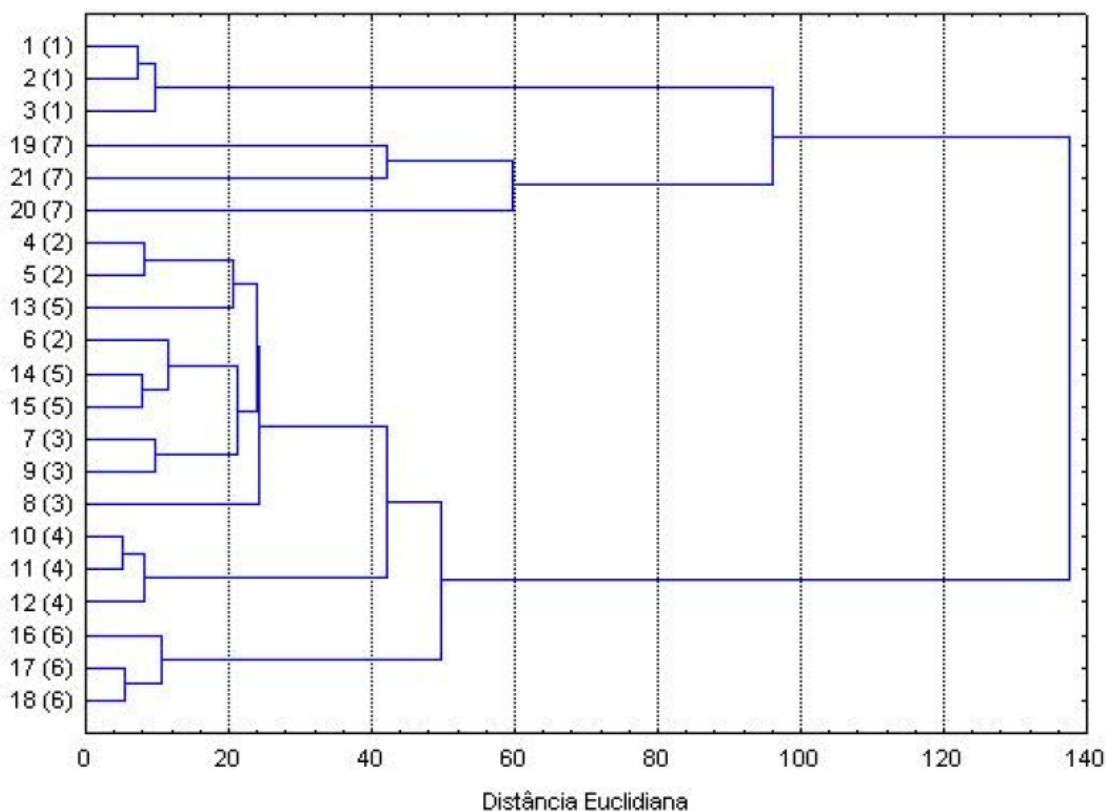


Figura 8 – Dendrograma das amostras de café. Os números entre parênteses representam a amostra a qual a repetição pertence.

Os sensores de maior ICM para a amostra 6 foram S6 e S8. A área média do sinal do sensor S8 (Tabela 6) é menor apenas do que àquela obtida para a amostra 7. Já para o sensor S6 o comportamento é repetido em relação a amostra 7, contudo o sinal não difere estatisticamente ($p > 0,05$) do obtido para a amostra 1. Na Figura 7 (Gráfico B) pode ser observada a separação da amostra 6 e a relação entre os sinais para as amostra 1 e 7 conforme descrito acima.

Para a amostra 7 os sensores de maior ICM foram S8 e S10. Como já comentado anteriormente, o sensor S8 apresentou uma maior área média para a amostra 7 em relação a todas as outras. Na Figura 7 (Gráfico A) é possível observar a separação da amostra 7.

O sensor S10 foi o mais citado, sendo três vezes reconhecido como entre as áreas médias (Tabela 6), destaca-se que, o sensor S7 diferentemente dos outros sensores, separa as amostras em apenas dois grupos. Tal fato pode ser uma possível explicação para a rede dar pouca importância aos dados advindos deste sensor. O mais importante, seguindo o critério do ICM (Tabela 5). Através da análise da Figura 7 (Gráficos A, C e D) e Tabela 6 fica claro que o sensor S10 é responsável pela separação das amostras em dois grupos. Sendo o primeiro formado pelas amostras 2 e 3, e o segundo pelas amostras 1, 4, 5, 6 e 7. Por outro lado, o sensor S7 foi o único que não esteve entre os mais importantes. Verificando a comparação

3.4 CONCLUSÕES

O método do bootstrap para reamostragem permitiu a produção de um banco de dados apropriado para o treinamento da rede, que somado ao aumento da capacidade de generalização, via emprego da média de “ensemble” e do método simplex de otimização possibilitaram o desenvolvimento de uma rede neural artificial capaz de reconhecer 100% das amostras de café solúvel avaliadas. O classificador neural ótimo obtido, acoplado ao nariz eletrônico, apresenta-se como uma ferramenta eficiente e eficaz para o controle de qualidade do café solúvel. A próxima etapa consiste na aplicação da metodologia adotada em um banco de dados contendo uma variedade maior de cafés solúveis.

O emprego do índice de casualidade permitiu extrair algumas informações sobre as regras utilizadas pela rede para realizar a classificação dos

cafés estudados. A associação destas relações preliminares com outras informações como dados cromatográficos e/ou sensoriais podem levar a um entendimento mais profundo sobre o impacto aromático dos compostos que constituem o café.

3.5 REFERÊNCIAS

- [1] A. Farah, M. C. Monteiro, V. Calado, A. S. Franca, L. C. Trugo, Correlation between cup quality and chemical attributes of Brazilian coffee, *Food Chem.* 98 (2006) 373-380.
- [2] C. A. B. de Maria, R. F. A. Moreira, L. C. Trugo, Componentes voláteis do café torrado. Parte I: Componentes heterocíclicos, *Química Nova* 22 (1999) 209-217.
- [3] R. F. A. Moreira, L. C. Trugo, C. A. B. de Maria, Componentes voláteis do café torrado. Parte II. Compostos alifáticos, alicíclicos e aromáticos, *Química Nova* 23 (2000) 195-203.
- [4] A. A. Mello e L. C. Trugo, in: M. R. B. Franco (Ed.), *Aroma e Sabor de Alimentos*, São Paulo, Livraria Varela, 2003, pp.169-175 (Chapter 12).
- [5] D. G. Bassoli, Impacto aromático dos componentes voláteis do café solúvel: uma abordagem analítica e sensorial, Tese de Doutorado, Universidade Estadual de Londrina, (2006).
- [6] S. D. Dutcosky, *Análise Sensorial de Alimentos*, 2ª ed., Curitiba, Champagnat, 2007, p.239.
- [7] I. Bardot, L. Bochereau, N. Martin, B. Palagos, Sensory-Instrumental correlations by combining data analysis and neural network techniques, *Food Quality and Preference* 5 (1994) 159-166.
- [8] E. Schaller, J. O. Bosset, F. Escher, 'Electronic noses' and their application to food, *LWT* 31 (1998) 305-316.
- [9] A. K. Deisingh, D. C. Stone, M. Thompson, Applications of electronic noses and tongues in food analysis, *International J. of Food Sci. and Technol.* 39 (2004) 587-604.
- [10] Y. Huang, L. J. Kangas, B. A. Rasco, Applications of artificial neural networks (ANNs) in food science, *Critical Reviews in Food Sci. and Nutrition* 47 (2007) 113-126.
- [11] J. Trihass e P. V. Nielsen, Electronic nose technology in quality assessment: Monitoring the ripening process of Danish Blue Cheese, *J. of Food Sci.* 70 (2005) 44-49.

- [12] S. Panigrahi, S. Balasubramanian, H. Gu, C. Logue, M. Marchello, Neuralnetwork- integrated electronic nose system for identification of spoiled beef, *LWT* 39 (2006) 135-145.
- [13] R. Linder e S. J. Pöppel, A new neural network approach classifies olfactory signals with high accuracy, *Food Quality and Preference* 14 (2003) 435-440.
- [14] Z. Hai e J. Wang, Electronic nose and data analysis for detection of maize oil adulteration in sesame oil, *Sens. and Actuators B* 119 (2006) 449-455.
- [15] M. E. Monge, D. Bulone, D. Giancomazza, D. L. Bernik, R. M. Negri, Detection of flavour release from pectin gels using electronic noses, *Sens. and Actuators B* 101 (2004) 28-38.
- [16] H. Zhang, J. Wang, X. Tian, H. Yu, Y. Yu, Optimization of sensor array and detection of stored duration of wheat by electronic nose, *J. of Food Engineering* 82 (2007) 403-408.
- [17] S. Buratti, D. Ballabio, S. Benedetti, M. S. Cosio, Prediction of Italian red wine sensorial descriptors from electronic nose, electronic tongue and spectrophotometric measurements by means of genetic algorithm regression models, *Food Chem.* 100 (2007) 211-218.
- [18] R. Dutta, K. R. Kashwan, M. Bhuyan, E. L. Hines, , J. W. Gardner, Electronic nose based tea quality standardization, *Neural Networks* 16 (2003) 847-853.
- [19] M. Pardo e G. Sberveglieri, Coffee analysis with an electronic nose, *IEEE Transactions on Instrumentation and Measurement* 51 (2002) 1334-1339.
- [20] M. A. Craven, J. W. Gardner, P. N. Bartlett, Electronic noses – development and future prospects, *Trends in Anal. Chem.* 15 (1996) 486-493.
- [21] S. Pérez-Magariño, M. Ortega-Heras, M. L. González San José, Z. Boger, Comparative study of artificial neural network and multivariate methods to classify Spanish DO rose wines, *Talanta* 62 (2004) 983-990.
- [22] T. Windeatt, Accuracy/Diversity and Ensemble MLP Classifier Design, *IEEE Transactions on Neural Networks* 17 (2006) 1194-1211.
- [23] S. Haykin, *Redes Neurais: Princípio e Práticas*, 2ª ed., Porto Alegre, Bookman, 2001, p.900.
- [24] B. Curry e P. H. Morgan, Model selection in Neural Networks: Some difficulties, *European J. of Operational Research* 170 (2006) 567-577.
- [25] T. B. Ludermir, A. Yamazaki, C. Zanchetin, An optimization methodology for neural network weights and architectures, *IEEE Transactions on Neural Networks* 17 (2006) 1452-1459.

- [26] W. Sukthomya e J. Tannock, The optimization of neural network parameters using Taguchi's design of experiments approach: an application in manufacturing process modeling. *Neural Computation & Application* 14 (2005) 337-344.
- [27] S. Dragović e A. Onjia, Prediction of peak-to-background ratio in gamma-ray spectrometry using simplex optimized artificial neural network, *Applied Radiation and Isotopes* 63 (2005) 363-366.
- [28] A. P. Braga, A. C. P. L. F. Carvalho, T. B. Ludermir, *Redes Neurais Artificiais: Teoria e Aplicações*, Rio de Janeiro, LTC, 2000, p.262.
- [29] I. A. Basheer e M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, *J. of Microbiological Methods* 43 (2000) 3-31.
- [30] B. Warner e M. Misra, Understanding neural networks as statistical tools, *The American Statistician* 50 (1996) 284-293.
- [31] F. M. Azevedo, L. M. Brasil, R. C. L. Oliveira, *Redes Neurais com Aplicações em Controle e em Sistemas Especialistas*, Florianópolis, Bookstore, 2000, p401.
- [32] E. F. Ferreyra Ramírez, Modelo conexionista para avaliação de propostas para aquisição de equipamentos médico-hospitalares, Tese de Doutorado, Universidade de Campinas (2005).
- [33] D. C. Montgomery, *Design and Analysis of Experiments*, 6ª ed, New York, John Wiley & Sons, 2005, p.643.
- [34] E. Bona, D. Borsato, R. S. S. F. Silva, R. P. Herrera, Aplicativo para otimização empregando o método simplex seqüencial, *Acta Scientiarum* 22 (2000) 1201-1206.
- [35] G. Derringer e R. Suich, Simultaneous optimization of several response variables. *J. of Quality Technology* 12 (1980) 214-219.
- [36] E. Kolman e M. Margaliot, Are artificial neural networks white boxes? *IEEE Transactions on Neural Networks* 16 (2005) 844-852.
- [37] J. J. Montañó e A. Palmer, Numeric sensitivity analysis applied to feedforward neural networks, *Neural Computing & Applications* 12 (2003) 119-125.
- [38] S. Papadokonstantakis, A. Lygeros, S. P. Jacobsson, Comparison of recent methods for inference of variable influence in neural networks, *Neural Networks* 19 (2006) 500-513.
- [39] Z. Boger, Selection of quasi-optimal inputs in chemometrics modeling by artificial neural network analysis, *Anal. Chim. Acta* 490 (2003) 31-40.
- [40] S. M. Della Lucia e L. A. Minim, in: S. M. Della Lucia (Ed.), *Análise Sensorial: Estudos com Consumidores*, Viçosa, UFV, 2006, pp.173-194, (Chapter 7).

[41] J. S. Torrecilla, M. L. Mena, P. Yáñez-Sedeño, J. García, Application of artificial neural network to the determination of phenolic compounds in olive oil mill wastewater, J. of Food Engineering 81 (2007) 544-552.

[42] C. G. Looney, Advances in feedforward neural networks: demystifying knowledge acquiring black boxes, IEEE Transactions on Knowledge and Data Engineering 8 (1996) 211-226.

3.6 AGRADECIMENTOS

Agradeço à Fundação Araucária pelo suporte financeiro, à CAPES pela bolsa de doutorado e à Companhia Iguaçu de Café Solúvel que, representada pelo Dr. Denisley Gentil Bassoli, forneceu todo o suporte experimental e analítico necessário para este trabalho.

CAPÍTULO IV

4 REDE NEURAL OTIMIZADA PARA A CLASSIFICAÇÃO DE CAFÉ SOLÚVEL ATRAVÉS DE UM NARIZ ELETRÔNICO: II) GENERALIZAÇÃO PARA UM BANCO DE DADOS INDUSTRIAL

Este artigo será submetido para Journal of Food Engineering (As normas para publicação neste periódico estão contidas no Anexo C)

Resumo

A avaliação aromática de bebidas, como o café, é uma tarefa de alta complexidade e importância para a indústria. O emprego do nariz eletrônico associado às redes neurais tem sido uma alternativa promissora. Neste trabalho foram utilizadas duas redes neurais, o perceptron de múltiplas (aprendizagem supervisionada) e o mapa auto-organizável (aprendizagem não supervisionada), para analisar os padrões aromáticos de café solúvel obtidos através de um nariz eletrônico. A rede perceptron obteve um desempenho de classificação correta de 90% para as amostras de validação e de 75% para as de teste. Esta performance só foi alcançada graças a utilização da média de ensemble e da codificação binária das classes. O mapa auto-organizável confirmou que os erros de classificação ocorreram entre produtos que apresentam similaridade aromática, de acordo com o nariz eletrônico. Além disso, o mapa permite testar estatisticamente similaridades entre produtos e associar a distribuição das amostras com a dispersão dos valores observados para cada sensor ou qualquer outra variável disponível para descrever os produtos. A combinação do nariz eletrônico com as redes neurais artificiais mostrou ser uma alternativa promissora para o desenvolvimento de novos produtos, comparação de similaridades entre produtos concorrentes e para o controle da qualidade aromática de café solúvel.

Palavras-chave: Perceptron de múltiplas camadas. Mapas auto-organizáveis. Reamostragem Bootstrap. Classificação binária.

4.1 INTRODUÇÃO

O reconhecimento e a classificação de padrões é uma tarefa naturalmente realizada pelos humanos graças à evolução e adaptação, durante milhares de anos, de nosso sistema nervoso central, mais especificamente o córtex cerebral. Entretanto, a solução de problemas de classificação através de sistemas

automatizados, na maioria dos casos, é de extrema complexidade, principalmente se os padrões forem descritos por um grande número de variáveis independentes (BISHOP, 1995; HAYKIN, 2001). Existem várias metodologias estatísticas tradicionais e consagradas, tanto paramétricas quanto não paramétricas, usadas para a classificação e reconhecimento de padrões (HAIR et al., 2005). As redes neurais artificiais (RNA) são um conjunto de técnicas baseadas em princípios estatísticos não tão convencionais que vem crescentemente ganhando espaço no reconhecimento e classificação de padrões (BISHOP, 1995; WARNER; MISRA, 1996). As RNA são extremamente versáteis para realizar o mapeamento de relações complexas e não-lineares entre múltiplas variáveis de entrada e saída. A construção de regras de classificação baseadas apenas nos dados disponíveis, sem a imposição de um modelo a priori, é outro atrativo da técnica (SABLANI, 2008). Entre as desvantagens da metodologia citam-se: a necessidade de uma grande quantidade de dados de treinamento (BISHOP, 1995); dificuldade na escolha dos parâmetros de treinamento e do tipo de rede mais adequada ao problema (CURRY; MORGAN, 2006); variabilidade dos resultados devido à inicialização e amostragem (BODT et al., 2002). Porém, vários trabalhos vêm apresentando propostas promissoras para solucionar ou amenizar as desvantagens associadas às RNA (LOONEY, 1996; KOLMAN; MARGALOT, 2005; LUDERMIR et al., 2006; WINDEATT, 2006). Como pode ser observado nas revisões realizadas por Liao (2005) e Huang et al. (2007), as RNA tem sido empregadas com sucesso dentro da área de ciência, tecnologia e engenharia de alimentos. Em vários trabalhos (DÍAZ et al., 2003; LINDER; PÖPPL, 2003; PANIGRAHI et al., 2006; CHANDRARATNE et al., 2007; MARINI et al., 2007; TIGRINE-KORDJANI et al., 2007) foi obtido um desempenho superior ou similar dos classificadores neurais quando comparados com técnicas estatísticas tradicionais. Além disso, os modelos neurais estão sendo apontados como ferramentas promissoras para o controle de qualidade (NI; GUNASEKARAN, 1998) e análise sensorial (BOMIO, 1998; DELLA LUCIA; MINIM, 2006). Um dos desafios desta última área é a avaliação objetiva do aroma que é uma das características mais complexas e determinantes da qualidade de alimentos, principalmente de bebidas como o café (DUTCOSKY, 2007). A torrefação promove drásticas transformações químicas do grão verde de café. A extração com água quente do produto torrado e moído origina então a bebida que, quando concentrada, pode ser desidratada dando origem ao café solúvel. O somatório de diferentes

atributos específicos de determinados compostos voláteis irá compor o aroma típico do café solúvel. Furanos, pirróis, pirazinas, piridinas e fenóis são alguns grupos relevantes, entre os mais de mil componentes já detectados, para a composição desse aroma que, ao serem avaliados separadamente, nem sempre corresponderão ao aroma agradável observado no produto (MARIA et al., 1999; MOREIRA et al., 2000; MELLO; TRUGO, 2003). Portanto, o reconhecimento e a classificação de padrões aromáticos de café solúvel é uma tarefa complexa e de grande interesse. Mais recentemente, o nariz eletrônico, que é constituído por um arranjo de sensores parcialmente específicos, associado às RNA tem sido utilizado com êxito na avaliação aromática de alimentos e bebidas (DUTTA et al., 2003; DEISINGH et al., 2004; MONGE et al., 2004; TRIHASS; NIELSEN, 2005; HAI; WANG, 2006; ZHANG et al., 2007) e até em café torrado (PARDO; SBERVEGLIERI, 2002). O objetivo deste trabalho foi utilizar redes neurais artificiais, com aprendizagem supervisionada e não supervisionada, para avaliar padrões aromáticos de café solúvel obtidos através de um nariz eletrônico, usando um banco de dados industrial.

4.2 MATERIAIS E MÉTODOS

4.2.1 Café Solúvel

Foram analisados 40 cafés solúveis produzidos pela Companhia Iguaçu de Café Solúvel (Cornélio Procópio – Paraná – Brasil). Cada produto foi selecionado e identificado pela própria indústria usando uma letra maiúscula e dois algarismos arábicos. No total foram avaliados 71 lotes conforme apresentado na Tabela 1.

Tabela 1 – Identificação dos produtos e quantidade de lotes avaliados.

Produto	Lotes	Produto	Lotes	Produto	Lotes	Produto	Lotes
A01	1	A02	2	A03	2	A04	3
B01	1	B02	4	C02	1	C04	1
C05	1	C09	1	C10	1	C11	1
C13	1	D01	1	E01	1	E02	3
E03	1	E04	1	F01	1	H01	1
K01	1	K02	4	N01	4	N02	3
N03	3	N05	1	N06	1	N07	1
N08	3	N09	1	R01	3	S01	2
S02	1	S03	1	X02	1	X03	3
X04	1	X05	2	X07	3	X08	3

4.2.2 Nariz Eletrônico

Todas as análises foram realizadas no laboratório de pesquisa e desenvolvimento da Companhia Iguaçu de Café Solúvel. Foi utilizado o equipamento Portable Electronic Nose Type PEN2 (Airsense Analytics, Alemanha), que monitora a variação da condutância em cada um de seus sensores (Tabela 2) conforme a passagem do fluxo gasoso com a amostra (HAI; WANG, 2006; ZHANG et al., 2007). Essa variação é então comparada com aquela obtida pelo branco (ar ambiente) gerando, assim, um sinal adimensional.

No preparo das amostras 2,8g de café solúvel foram colocados em frascos específicos para análise de headspace. Cada lote de café solúvel (Tabela 1) foi avaliado em três repetições genuínas totalizando 213 amostras. Os parâmetros operacionais do equipamento foram (BASSOLI, 2006): 1,0 s de intervalo entre as amostras; 60 s de tempo de medição; 120s para limpeza ou purga dos sensores; 400 mL/min de fluxo de injeção.

Tabela 2 – Especificidade dos sensores utilizados no nariz eletrônico PEN2.

Sensor	Especificidade
S1	Compostos aromáticos
S2	Amplo espectro, muito sensível, reage com óxidos de nitrogênio
S3	Amônia, usado para compostos aromáticos
S4	Hidrogênio
S5	Alcanos e compostos aromáticos menos polares
S6	Metano, amplo espectro
S7	Compostos sulfurados, terpenos, limonenos e pirazinas
S8	Álcoois, amplo espectro
S9	Compostos orgânicos sulfurados
S10	Metanos alifáticos em altas concentrações > 100 ppm

4.2.3 Variáveis de Entrada: Integração e Autoescalamento

Após a aquisição dos dados, a área abaixo da curva gerada por cada sensor foi numericamente integrada (SPERANDIO et al., 2003), formando um vetor de entrada com 10 valores para cada amostra. Antes de serem alimentados na rede neural os vetores de entrada foram autoescalados (média igual a zero e

variância unitária) para garantir a mesma métrica para todas as variáveis (HAYKIN, 2001).

4.2.4 Variáveis de Saída: Representação dos produtos

Para o problema de classificação de 40 cafés solúveis, em teoria (HAYKIN, 2001), seriam necessários 40 neurônios de saída para uma rede do tipo perceptron de múltiplas camadas (MLP – multi-layer perceptron). Portanto, cada um dos k produtos seria representado por um vetor de saída d_k onde é associado o valor 1 para a k -ésima posição e 0 para as outras. Assim, o classificador a ser construído teria um tamanho demasiadamente grande sendo necessário também uma elevada quantidade de exemplos de treinamento da rede (BISHOP, 1995). Uma alternativa, denominada classificação ordinal (CO), é mapear as classes em um único neurônio de saída (STATSOFT, 2005). Os algarismos ordinais que representam as classes são convertidos para uma escala contínua entre 0 e 1, ou seja, o mesmo intervalo de saída da função de ativação logística. Além da CO é possível transformar os números inteiros que representam as classes em números binários (BASHEER; HAJMEER, 2000). Para implementar a classificação binária (CB) no presente problema são necessários 6 neurônios, os quais permitem a representação de até 64 classes (SPERANDIO et al., 2003). Neste trabalho, tanto a CO quanto a CB foram aplicadas e comparadas com o objetivo de avaliar qual a melhor regra de representação das saídas.

4.2.5 Redes Neurais Artificiais

Foi empregada uma rede neural artificial do tipo perceptron de múltiplas camadas (MLP) que é amplamente utilizada para a classificação de padrões em conjunto com o nariz eletrônico (DEISINGH et al., 2004). A arquitetura deste tipo de rede (Figura 1) é composta por uma camada de entrada com 10 neurônios (um para cada sensor do nariz eletrônico), uma ou duas camadas intermediárias (responsáveis pela separação dos padrões através da formação de fronteiras de decisão) contendo uma quantidade de neurônios a ser definida e uma

camada de saída com 1 neurônio caso seja utilizada uma CO para representar as classes ou 6 neurônios para a CB (HAYKIN, 2001).

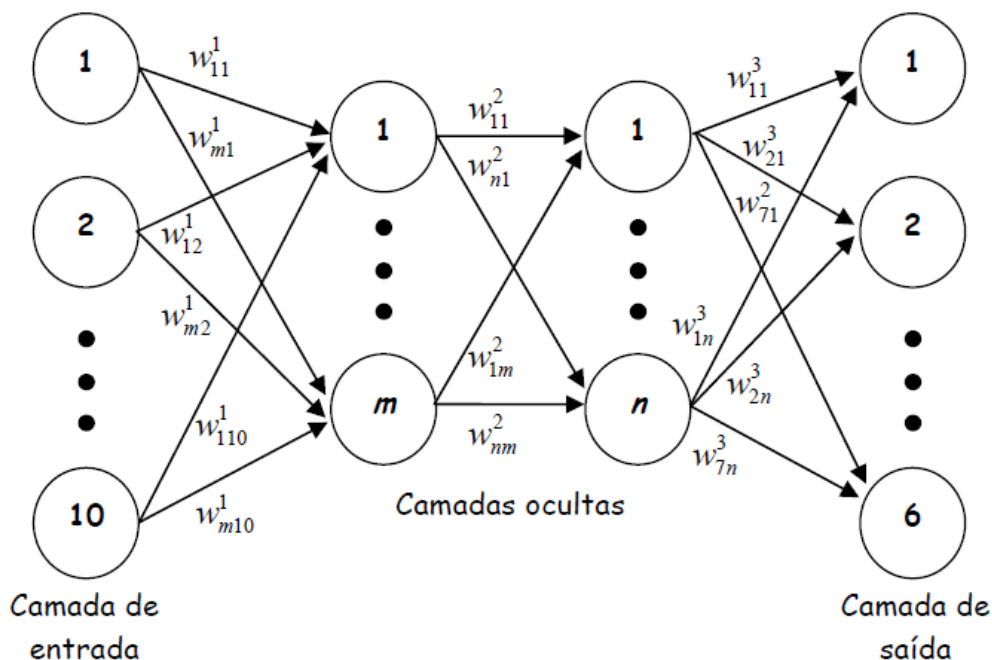


Figura 1 – Representação gráfica da rede perceptron de múltiplas camadas com duas camadas ocultas. Para cada conexão entre os neurônios existe um peso w_{ij}^l associado, sendo i o índice do neurônio de entrada do sinal, j o neurônio de saída do sinal e l camada onde esta localizado o neurônio de entrada. A camada de saída pode ter 1 ou 6 neurônios dependendo do tipo de representação adotado para as classes.

Em cada neurônio da rede MLP foi realizado um somatório, ponderado pelos pesos sinápticos, dos sinais provenientes dos neurônios da camada anterior. Esta soma foi aplicada a uma função de ativação do tipo sigmóide logística, sendo a inclinação desta função um parâmetro a ser otimizado (BISHOP, 1995). Para a etapa de aprendizagem da rede foi utilizado o algoritmo de retropropagação, que consiste de dois passos através das diferentes camadas da rede: um passo a frente, a propagação (HAYKIN, 2001), e um passo atrás, a retropropagação. No passo a frente, um padrão de atividade (vetor de entrada) é aplicado aos neurônios de entrada da rede e seu efeito se propaga através da rede, camada por camada. Finalmente, um conjunto de saídas é produzido como resposta real da rede. Durante o passo de propagação, os pesos sinápticos da rede são todos fixos. No passo atrás, por outro lado, os pesos são ajustados de acordo com uma

regra de correção de erro. Como função de erro ou de custo, foi adotado o erro quadrático médio (EQM). Tanto a taxa de aprendizagem quanto a constante de momento, que são parâmetros inerentes ao algoritmo de retropropagação, foram otimizados para o problema proposto. Foi utilizado o modo seqüencial de treinamento, ou seja, os pesos são atualizados a cada novo exemplo apresentado à rede. Em cada época os exemplos foram apresentados à rede de maneira aleatória tornando estocástica a busca no espaço de pesos. Por sua vez, isto torna menos provável que o algoritmo de retropropagação fique preso em um mínimo local. Em todas as redes utilizadas neste estudo foram utilizadas 250 épocas de treinamento. Tendo como objetivo uma boa generalização, foi utilizado o método de treinamento com parada antecipada para evitar que a rede acabe sendo excessivamente ajustada aos dados de treinamento (HAYKIN, 2001).

Além da rede MLP foi utilizado um mapa auto-organizável (em inglês self-organizing map – SOM) bidimensional proposto e discutido em maiores detalhes por Haykin (2001). O principal objetivo do mapa auto-organizável (SOM – self-organizing map) é transformar um padrão de sinal incidente de dimensão arbitrária em um mapa discreto bidimensional e realizar esta transformação adaptativamente de uma maneira topologicamente ordenada. No modelo de Kohonen (Figura 2) cada neurônio da grade está totalmente conectado com todos os nós de fonte da camada de entrada. Esta grade representa uma estrutura alimentada adiante com uma única camada computacional consistindo de neurônios arranjados em linhas e colunas. Os neurônios apresentam a mesma quantidade de pesos, que por sua vez é igual ao número de variáveis analisadas. Cada variável pode ser representada por um plano de pesos (Figura 2), que após a fase de treinamento pode ser utilizado na tentativa de extrair regras sobre os grupos formados.

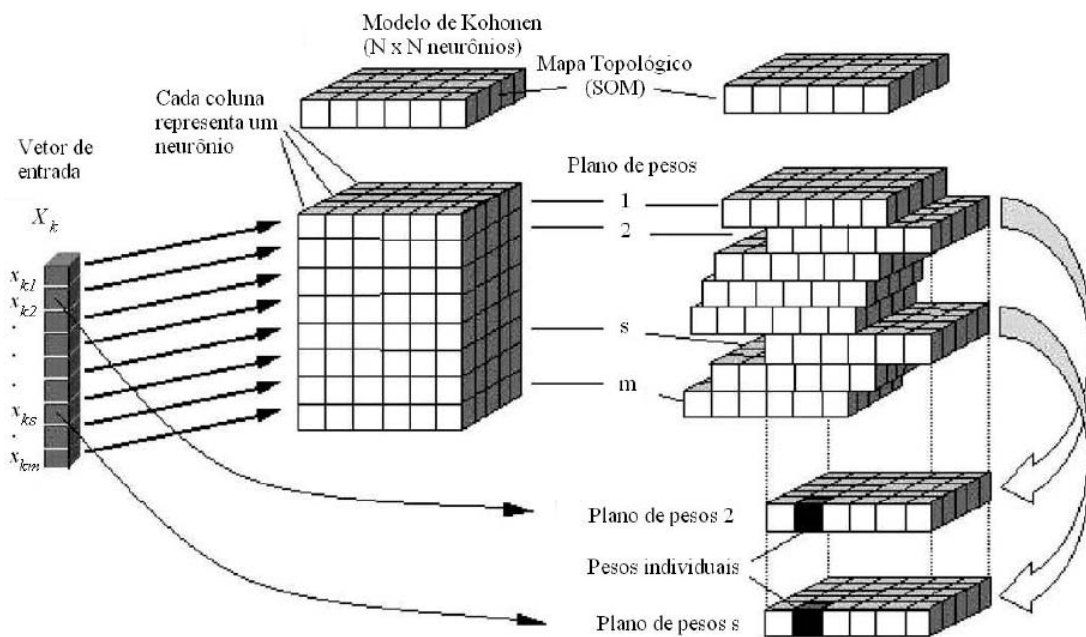


Figura 2 – Mapa auto-organizável segundo o modelo de Kohonen. Na figura k representa o número de padrões de entrada, m é a quantidade variáveis de entrada e N é o número de neurônios em cada dimensão (MARINI et al., 2005).

Após o treinamento, o resultado pode ser visualizado através de um mapa topológico e mapas de peso. No primeiro, cada padrão de entrada é posicionado sobre o mapa de características de acordo com a posição do neurônio vencedor da amostra considerada (Figura 2). A função do mapa topológico é facilitar a visualização de agrupamentos e também a relação de vizinhança entre os grupos formados. Grupos próximos partilham de certa similaridade, assim como, quanto maior a distância maior será a diferença de comportamento. O mapa de pesos (ZUPAN et al., 1997) é um gráfico de contorno para cada plano de pesos (Figura 2). Em conjunto com o mapa topológico é possível extrair regras de comportamento para cada agrupamento formado e inferir a influência de cada variável sobre o resultado obtido. Para verificar se o agrupamento e as relações de vizinhança entre as classes devem-se às propriedades intrínsecas dos dados ou apenas ao acaso foi utilizada uma metodologia baseada em métodos estatísticos clássicos e proposta por Bodt et al. (2002). Segundo o autor, as duas principais fontes de variabilidade para o SOM são devidas à inicialização e amostragem. O treinamento de vários mapas usando sempre o mesmo conjunto de dados, possibilitaria apenas a análise da variabilidade devido ao processo de geração aleatória dos pesos iniciais.

Portanto, é necessária a utilização de uma técnica de reamostragem (item 4.2.6) que permita um conjunto de dados diferentes a cada novo treinamento. Após o treinamento de várias redes foi avaliada a variabilidade do erro médio de quantização (EMQ) do mapa em relação aos k padrões entrada. De acordo com Bodt et al. (2002), um gráfico do coeficiente de variação contra o número de neurônios usados no mapa fornece uma maneira empírica para selecionar a tamanho mais adequado do SOM para os dados analisados. Segundo o mesmo autor, o aumento no número de neurônios usados em cada dimensão leva a um aumento do coeficiente variação. Pois, assim sendo, quanto maior o número unidades no mapa maior serão as possibilidades de distribuição das amostras, levando assim, a uma instabilidade nos resultados. Além disso, um teste de hipótese (usando uma distribuição de probabilidade binomial) do valor de uma função de vizinhança foi empregado para avaliar a estabilidade da relação de vizinhança entre um par específico de amostras. A hipótese nula (H_0) considera que a relação de vizinhança é devida ao acaso. A hipótese alternativa (H_1) supõe que o par pode ser estatisticamente considerado vizinho ou não dentro de um raio pré-determinado. É possível, também, uma análise que considere todos os pares possíveis construindo gráficos de probabilidade acumulada (BODT et al., 2002) que permitem uma análise qualitativa de significância. Uma discrepância entre o gráfico de probabilidade acumulada do mapa auto-organizável e da distribuição binomial é um indicativo de que as relações de vizinhança construídas não são casuais.

4.2.6 Amostragem e Bootstrap

O bootstrap (reamostragem) consiste na obtenção de medidas de dispersão, (como a média, variância ou desvio-padrão) através da aplicação de métodos tradicionais, de inferência e estimação, a diversos conjuntos de dados obtidos a partir de um único conjunto original mediante a geração de dados aleatórios com a mesma média e desvio padrão dos originais (RAMÍREZ, 2005).

Como mencionado nos itens 4.2.1 e 4.2.2, foram analisados 71 lotes de café solúvel em triplicatas genuínas totalizando 213 amostras. Estas foram subdivididas em dois conjuntos: um de validação composto por 166 amostras e outro de teste com 47. O conjunto de teste foi formado através do sorteio de uma das

repetições de cada lote para os produtos com mais de um lote. Pois, caso fossem retiradas amostras dos cafés com apenas um lote restariam apenas duas para empregar a metodologia bootstrap. Usando, então, o bootstrap foram criadas mais 664 amostras a partir do conjunto de validação. Em resumo, para cada rede foram usadas 75,7% das amostras para treinamento (geradas aleatoriamente), 18,9% para validação e 5,4% para teste (amostras originais).

4.2.7 Média de Ensemble para a Rede MLP

Uma maneira de melhorar a capacidade de generalização das redes neurais é a combinação das respostas de várias RNAs, também chamada de máquinas de comitê (HAYKIN, 2001), as quais foram treinadas de maneira semelhante de modo que, o resultado final seja melhor do que aquele obtido ao ser utilizado apenas uma única RNA. Uma maneira fácil de implementar o comitê de máquinas é a média de ensemble, que consiste em combinar as respostas das redes treinadas através de uma média. O uso de um ensemble diminui a probabilidade de erro final, pois será coberto um espaço maior de hipóteses dentro das possíveis soluções para o problema. A desvantagem deste método reside no aumento do tempo computacional (BISHOP, 1995). Neste trabalho foi utilizada uma média de ensemble com 50 redes, sendo que para cada rede um novo conjunto de pesos iniciais e de exemplos de treinamento foi gerado aleatoriamente.

4.2.8 Otimização dos Parâmetros da Rede MLP

Nas redes MLP alguns parâmetros de arquitetura e do algoritmo de aprendizagem foram otimizados através do método simplex seqüencial (BONA et al., 2000). Os parâmetros escolhidos para otimização foram: a quantidade de neurônios na primeira e segunda camada ocultas, taxa de aprendizagem, constante de momento e a inclinação da função de ativação logística. A otimização foi realizada visando maximizar a porcentagem de classificação correta e minimizar o erro quadrático médio para o conjunto de validação. Além disso, estas metas deveriam ser alcançadas com a menor rede possível, entende-se por isso aquela que

apresentar a menor dimensão para o vetor de pesos (HAYKIN, 2001). Para combinar estas três exigências em uma única resposta global, que será maximizada pelo método simplex seqüencial, foram utilizadas as funções de desejabilidade (DERRINGER; SUICH, 1980).

4.2.9 Implementação Computacional

A integração numérica dos dados obtidos pelo nariz eletrônico, o pré-processamento, a reamostragem bootstrap, as redes MLP o SOM e o procedimento de otimização usando o simplex seqüencial acoplado às funções de desejabilidade foram desenvolvidos em linguagem FORTRAN 90 para Windows. As estatísticas descritivas, ANOVA, comparação de médias (teste Tukey) e análise de componentes principais foram realizadas no software Statistica 7.1. Por uma questão de qualidade de apresentação, todos os gráficos dos resultados obtidos pelo programa desenvolvido também foram gerados no software Statistica 7.1. Para a CO (item 4.2.4) as redes foram implementadas tanto no módulo Neural Networks (Statistica 7.1) como no programa desenvolvido para comparar os resultados obtidos. Já para a CB não foi possível realizar a comparação, pois o módulo citado do software comercial não realiza este tipo de representação das classes.

4.3 RESULTADOS E DISCUSSÃO

4.3.1 Pré-Processamento

As áreas de cada sensor para cada uma das 213 amostras, foram integradas e avaliadas empregando-se estatística descritiva. Os resultados obtidos encontram-se na Tabela 3.

Tabela 3 – Estatísticas descritivas dos produtos analisados em função do sensor do nariz eletrônico.

Sensor	Média	Mínimo	Máximo	Desvio Padrão	Grupos*
S1	47,6314	35,2876	66,5042	4,7194	20
S2	135,9390	82,2900	374,4282	39,4974	20
S3	45,0266	33,2656	59,1386	4,2979	20
S4	58,3326	56,2317	59,2739	0,3973	10
S5	46,0312	34,6368	62,2796	4,6224	20
S6	71,5177	43,1103	93,5833	8,7596	21
S7	65,7896	59,8704	81,8271	3,5020	12
S8	67,2675	38,9638	86,0911	8,8284	20
S9	71,5900	60,9995	122,1716	8,8576	17
S10	56,6723	53,2077	59,6848	0,9518	14

* Grupos formados quando realizado o teste de Tukey, em nível de 5%, para tratamentos com diferentes quantidades de repetições.

Foi observado, através da análise de variância, que existe diferença significativa ($p < 0,01$) entre os produtos para todos os sensores utilizados pelo nariz eletrônico. Na Tabela 3, é possível confrontar a quantidade de grupos formados quando feita uma comparação das áreas médias usando um teste Tukey (em nível de 5% e para tratamentos com diferentes graus de liberdade). Destaca-se que, os sensores 4 e 6 apresentaram, respectivamente, o menor (10 grupos) e a maior (21 grupos) capacidade de discriminação entre os cafés analisados. Também, foi montada uma matriz de correlação entre as áreas dos sensores (Tabela 4). O padrão de correlação obtido é semelhante ao obtido por Hai e Wang (2006) que utilizaram um nariz eletrônico de mesma procedência, modelo e sensores para analisar adulterações em óleo de gergelim. Acredita-se, portanto, que este padrão deve-se à especificidade parcial de cada sensor. Em Bona (2008), onde foi utilizado um conjunto distinto de cafés solúveis, é possível observar que os sensores dividem-se nos mesmos quatro grupos em função da correlação: grupo 1 (S1, S3 e S5), grupo 2 (S2, S7 e S9), grupo 3 (S4 e S10) e grupo 4 (S6 e S8). Em relação ao trabalho citado, nota-se também uma redução nos valores dos coeficientes de correlação que pode estar relacionada ao aumento da quantidade de produtos analisados.

Tabela 4 – Correlação entre as áreas de cada sensor.

Sensor	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	1	-0,54*	0,98*	-0,06	0,98*	-0,88*	-0,25*	-0,85*	-0,46*	-0,21*
S2	-0,54*	1	-0,67*	-0,23*	-0,61*	0,52*	0,77*	0,42*	0,92*	-0,15*
S3	0,98*	-0,67*	1	0,01	0,98*	-0,86*	-0,38*	-0,81*	-0,59*	-0,14*
S4	-0,06	-0,23*	0,01	1	-0,01	0,17*	-0,21*	0,25*	-0,23*	0,40*
S5	0,98*	-0,61*	0,98*	-0,01	1	-0,92*	-0,27*	-0,88*	-0,50*	-0,13
S6	-0,88*	0,52*	-0,86*	0,17*	-0,92*	1	0,12	0,98*	0,35*	0,14*
S7	-0,25*	0,77*	-0,38*	-0,21*	-0,27*	0,12	1	0,04	0,82*	-0,11
S8	-0,85*	0,42*	-0,81*	0,25*	-0,88*	0,98*	0,04	1	0,24*	0,23*
S9	-0,46*	0,92*	-0,59*	-0,23*	-0,50*	0,35*	0,82*	0,24*	1	-0,13
S10	-0,21*	-0,15*	-0,14*	0,40*	-0,13	0,14*	-0,11	0,23*	-0,13	1

* Coeficientes de correlação estatisticamente significativos em nível de 5%.

Na Figura 3, está representada a análise de componentes principais usando a matriz de correlação das áreas integradas dos 10 sensores. Foram selecionadas as três primeiras componentes que apresentaram autovalores maiores que 1 (HAIR et al., 2005), totalizando 89,2% da variância total. Através do gráfico dos escores fatoriais fica evidente que, exceto para alguns produtos (A01, C04, C10, C11, C13, H01, N05), a separação das 213 amostras estudadas não é possível em um espaço de baixa dimensionalidade. No gráfico das cargas fatoriais são observados os agrupamentos de sensores conforme já identificado através da análise da matriz de correlação.

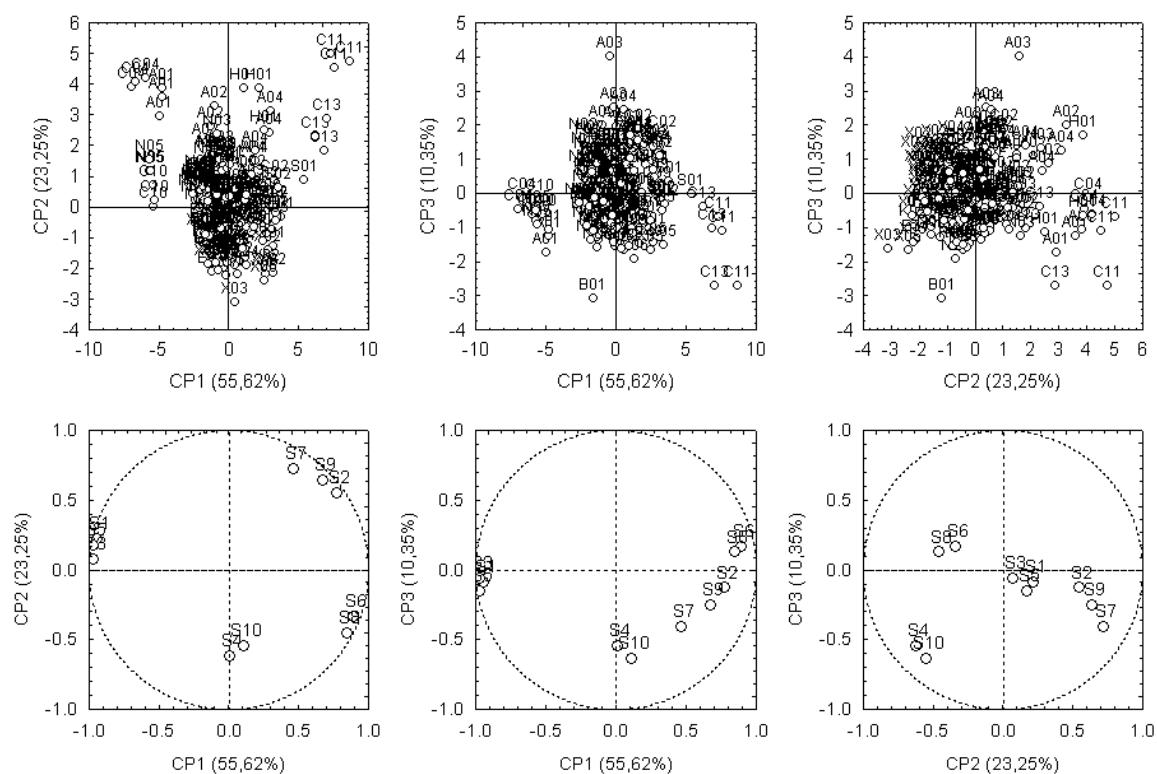


Figura 3 – Projeção dos escores e das cargas fatoriais nas três componentes principais selecionadas.

4.3.2 Redes Neurais Artificiais

As redes MLP implementadas, bem como, a comparação de desempenho estão descritas na Tabela 5. Os parâmetros otimizados não podem ser diretamente comparados com outros valores uma vez que são dependentes do problema (SUKTHOMYA; TANNOCK, 2005). Apesar da rede com CB apresentar menos neurônios ocultos (35) do que a rede com CO (37), a quantidade total de pesos da última é menor pois, nesta é usado apenas um neurônio de saída enquanto na outra são utilizados seis. Outra diferença entre os parâmetros ótimos para os dois tipos de representação das classes foi a taxa de aprendizagem, que apresentou um valor mais elevado para a CO. As 250 épocas de treinamento foram suficientes, tendo em vista que, nos três experimentos realizados a época média onde foi obtido o mínimo do EQM de validação foi inferior àquele valor. Em relação à validação, ficou bastante evidente que a CB possui um desempenho bem superior ao da CO (Tabela 5). Além disso, os resultados para a CO nas redes 2 e 3 são similares, ou seja, o programa desenvolvido apresenta uma performance similar ao

comercial com a vantagem de poder ser utilizada a CB. A diferença de desempenho de classificação entre os diferentes tipos de representação das classes pode ser devida à magnitude do erro quadrático médio, já que, no algoritmo de retropropagação a correção dos pesos é feita através do gradiente da função de erro. Mesmo a CO apresentando uma porcentagem de classificação correta muito menor do que a obtida pela CB, apresenta EQM muito menor. Assim, mesmo que um padrão de entrada seja erroneamente classificado usando a CO, o erro associado é baixo e, conseqüentemente, o ajuste dos pesos da rede será insuficiente. A Tabela 5 deixa evidente também que, o desempenho do ensemble é superior em relação à média tanto para a CB quanto para a CO. Este resultado está de acordo com Bishop (1995), que demonstra que o erro cometido pelo ensemble é sempre menor ou igual ao desempenho médio das redes. O aumento de performance vem acompanhado de uma redução da variabilidade da resposta da rede, como pode ser observado através dos desvios padrões apresentados na tabela mencionada.

Tabela 5 – Comparação de desempenho das redes MLP utilizadas^a.

Rede	Implementação		
	Fortran 90		Statística 7.1 ^b
	1	2	3
Representação das classes	Binária (CB)	Ordinal (CO)	Ordinal (CO)
Parâmetros otimizados pelo Simplex seqüencial^c			
Neurônios na camada 1	19	22	22
Neurônios na camada 2	16	15	15
Taxa de aprendizagem	0,2871	0,8881	0,8881
Constante de Momento	0,8388	0,9000	0,9000
Inclinação da função logística	0,8617	0,8327	1,0000
Tamanho da rede e épocas de treinamento			
Pesos Sinápticos	631	603	603
Época mínima (média)	173,50 ± 50,68	216,41 ± 32,18	144,80 ± 69,03
Desempenho de Validação^d			
Erro Quadrático	0,1213	0,0047	-
Médio (ensemble ^e)	± 0,0009	± 0,0001	-
Erro Quadrático	0,1787	0,0091	0,0095
Médio (média)	± 0,0201	± 0,0018	± 0,0026
% Classificação Correta (ensemble)	90,12 ± 0,33	33,13 ± 2,55	-
% Classificação Correta (média)	76,28 ± 2,77	25,53 ± 3,10	21,45 ± 3,32
Desempenho de Teste^f			
Erro Quadrático	0,1188	0,0085	-
Médio (ensemble)	± 0,0034	± 0,0005	-
Erro Quadrático	-	-	0,0193
Médio (média)	-	-	± 0,0057
% Classificação Correta (ensemble)	76,17 ± 2,33	24,26 ± 3,86	-
% Classificação Correta (média)	-	-	15,32 ± 6,27

^a Os valores médios e desvios padrões são para 250 repetições das redes, exceto no Statística 7.1 onde foram usadas 5 repetições.

^b O software Statística 7.1 não permite a representação binária das classes.

^c Com exceção da inclinação que não pode ser alterada, os parâmetros usados no Statística 7.1 foram os mesmos do programa desenvolvido em linguagem Fortran 90.

^d No software Statística 7.1 não foi possível aplicar, de maneira automatizada, a reamostragem bootstrap. Portanto, não foi avaliado o desempenho do ensemble.

^e O ensemble foi formado com 50 redes. As médias e os desvios padrões foram obtidos para 5 repetições do ensemble.

^f Por apresentar um desempenho superior, as amostras de teste foram classificadas apenas pelo ensemble no programa desenvolvido em linguagem Fortran 90.

Usando a CB (rede 1) verificou-se a influência da quantidade de repetições na resposta do ensemble. Os resultados estão apresentados na Tabela 6. Não foram observadas grandes alterações no desempenho do *ensemble* devido ao

aumento do número de repetições *bootstrap*. Apenas o erro e a porcentagem de classificação correta de validação apresentaram diferenças estatisticamente significativas em nível de 5%. A análise da porcentagem de acertos indica que 50 repetições são suficientes, uma vez que, o desempenho de classificação é semelhante ao ensemble com 70 repetições e, o tempo computacional é menor se forem usados ensembles menores.

Tabela 6 – Influência da quantidade de repetições na resposta do ensemble . *

Amostra	Repetições	40	50	60	70	80
Validação	Erro	0,1216 ^{a,b} ± 0,0014	0,1213 ^{a,b} ± 0,0009	0,1209 ^{a,b} ± 0,0018	0,1192 ^a ± 0,0031	0,1230 ^b ± 0,0009
	%Classificação Correta	89,28 ^{b,c} ± 0,66	90,12 ^{a,b} ± 0,33	89,64 ^{a,b,c} ± 0,50	90,36 ^a ± 0,42	89,04 ^c ± 0,79
	%Classificação Incorreta	10,72 ± 0,66	9,88 ± 0,33	10,36 ± 0,50	9,52 ± 0,50	10,72 ± 1,08
	%Amostras Desconhecidas	0,00 ± 0,00	0,00 ± 0,00	0,00 ± 0,00	0,12 ± 0,27	0,24 ± 0,33
	Erro	0,1170 ± 0,0040	0,1188 ± 0,0034	0,1154 ± 0,0022	0,1167 ± 0,0041	0,1161 ± 0,0016
	%Classificação Correta	74,90 ± 2,78	76,17 ± 2,33	75,32 ± 1,17	75,32 ± 1,17	75,32 ± 1,17
Teste	%Classificação Incorreta	24,25 ± 3,23	23,83 ± 2,33	24,68 ± 1,17	23,83 ± 1,78	24,25 ± 1,90
	%Amostras Desconhecidas	0,85 ± 1,17	0,00 ± 0,00	0,00 ± 0,00	0,85 ± 1,17	0,43 ± 0,95

* Médias e desvios padrões para cinco repetições do ensemble. As médias seguidas de letras diferentes apresentam diferença significativa em nível de 5%.

Todos os métodos estatísticos de classificação têm seu desempenho limitado pela variabilidade dos dados e pela sobreposição das classes no espaço multidimensional das variáveis de entrada (BISHOP, 1995). Segundo Haykin (2001), um classificador MLP, usando a função logística como não-linearidade, aproxima-se de um classificador baseado em uma estatística baesiana, desde que o tamanho do conjunto de treinamento seja suficientemente grande e que o processo de aprendizagem por retropropagação não fique preso em um mínimo local. Durante a otimização dos parâmetros através da metodologia simplex seqüencial foi observado que várias redes com arquiteturas diferentes convergiram para uma porcentagem de classificação correta em torno de 90% para a validação e

75% para o teste. A Tabela 6 sugere, também, que a capacidade máxima de classificação do banco de dados analisado está próxima dos valores já mencionados. Na Tabela 7 estão relacionados alguns dos erros de classificação cometidos pelas redes treinadas.

Tabela 7 – Matriz de confundimento das redes MLP utilizadas

	Produtos						
	A03	A04	B01	E02	K02	N01	X03
A04	X						
C02		X					
E03				X			
N08			X				
N09					X		
S03						X	
X04							X
X08				X			

Segundo informações fornecidas pela própria indústria, que disponibilizou as amostras, os produtos representados pela mesma letra são mais similares, o que justificaria alguns dos erros de classificação. Para os pares B01/N08 e K02/N09 não existem diferenças estatisticamente significativas ($p < 0,05$) quando comparados através de testes consagrados pelo uso (ANOVA; TUKEY).

Para avaliar a distribuição espacial dos dados foram treinadas, durante 5000 épocas, 50 repetições de um mapa auto-organizável (SOM) bidimensional com 13 neurônios por dimensão. A quantidade de neurônios utilizada foi estabelecida usando os critérios definidos por Bodt et al. (2002). Segundo o mapa treinado, para todos os pares de classificação incorreta apresentados na Tabela 7, existem amostras localizadas sobre o mesmo centróide (neurônio vencedor). Esta afirmação foi confirmada estatisticamente através de um teste de hipótese em nível de 5% (BODT et al., 2002). Na Figura 4 está representado o mapa topológico de uma das repetições. Cada amostra foi representada pelo símbolo do respectivo produto e posicionada de acordo com o neurônio vencedor. Os pares de confundimento (Tabela 7) estão destacados em negrito e posicionados próximos, ou seja, o mapa sugere que, segundo os dados fornecidos pelo nariz eletrônico, existe uma semelhança aromática para estes produtos. Como já mencionado, todas as

relações de vizinhança ilustradas no mapa podem ser estatisticamente avaliadas através de um teste de hipótese. Para alguns produtos, por exemplo K02 (em negrito na Figura 4), as amostras estão dispersas em vários centróides do mapa. Este comportamento sugere que existem diferenças aromáticas, segundo o nariz eletrônico, entre os quatro lotes do produto K02. Esta hipótese é corroborada pela análise de variância que indicou diferenças significativas ($p < 0,05$) entre os lotes para os sensores S1, S6 e S8.

13	H01		A02	A02	N03	N03	N03	N02	E02 F01 N02	F01 N02		A01	C04
12			A02	N03	N03	N03	F01	N02					N05
11	R01		R01	N02						X08	N02	C10	C10
10	R01	R01	R01		E02			X08	E02				
9				N01	K02	E02 E03	E02 E03	E02	X08		B01 N08	N08	B01
8		A03	A03	B02 X03		X07	X07	X03 X07	X07	E03 E04		N08	
7	A04	A03 A04	A03 A04	X04	N01		B02 X03		X07	E04 X08	D01	N08	B01
6		A03 A04		K01	N01	N01	K02 N01	X07			N07	D01	N07
5		A04 C02			N01		E01						E04
4	A04	C02	C09	K02	N09	E01 K02	N01 N06	N01	N01 S03	N01	C05	C05	X07
3	A04		S02 X05			X05	E01 K02	K02 N09	N06		K02 X03		
2	C13	C13		C09	X02					K02	B02 K02	B02 X03	
1	C11	C13	S01	S01	S01	S01 X05	X05	X05	K01	B02	X04	B02	X03
	1	2	3	4	5	6	7	8	9	10	11	12	13

Figura 4 – Mapa topológico de uma das repetições. Para as outras repetições pode ocorrer uma rotação do mapa, porém, as posições relativas são mantidas.

Para complementar a análise do mapa topológico foram construídos os mapas de peso para todos os sensores. As curvas de nível típicas da Figura 5 foram obtidas através da interpolação, por meio de splines (StatSoft Inc., Tulsa, OK, USA), do plano de pesos (Figura 1) correspondente a cada sensor. Assim como na análise de componentes principais (Figura 3), os mapas de pesos podem ser reunidos nos mesmos quatro grupos de acordo com o comportamento apresentado.

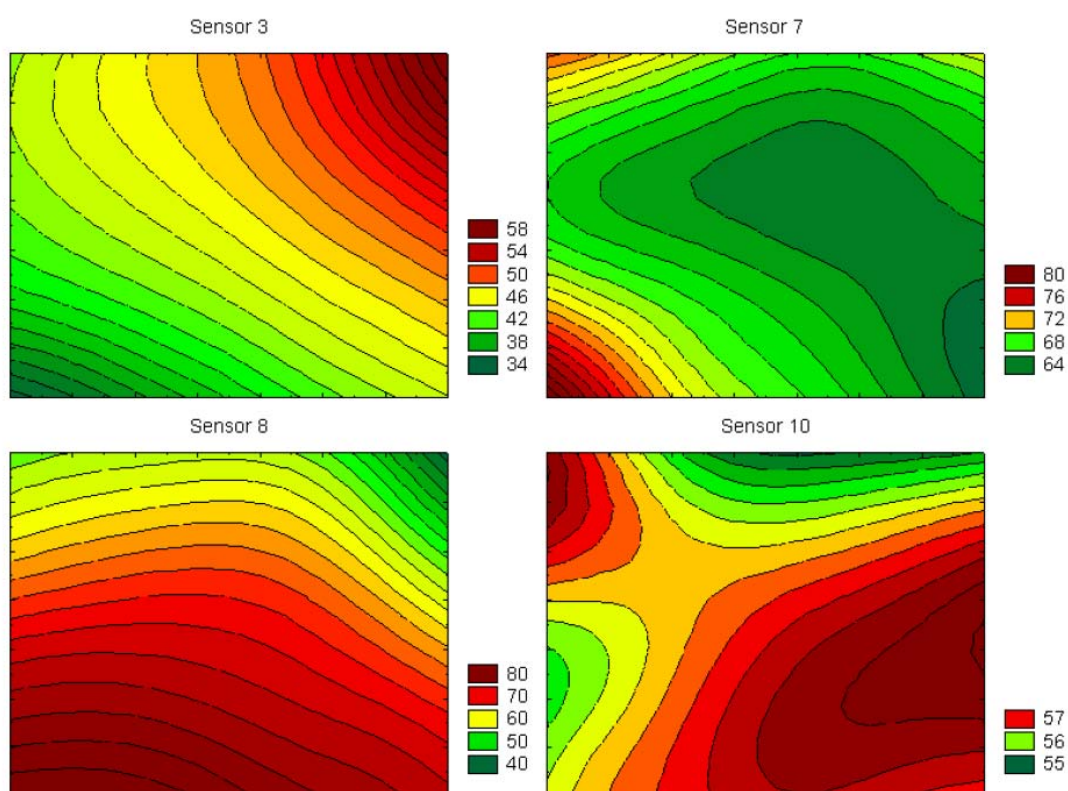


Figura 5 – Mapas de pesos de um dos sensores de cada grupo identificado.

A análise conjunta das Figuras 4 e 5 permite inferir, por exemplo, que os produtos A01, C04, C10 e N05 são os que apresentam um sinal mais intenso para os sensores S1, S3 e S5. Já os produtos C11 e C13 estão localizados na região de maior valor para os sensores S2, S7 e S9. Estas afirmações são confirmadas pelo teste Tukey, em nível de 5% e para tratamentos com diferentes graus de liberdade. Assim, o mapa pode ser uma ferramenta promissora durante a etapa de desenvolvimento ou controle de qualidade, tanto para testar estatisticamente similaridades entre produtos, como para associar a distribuição das

amostras com a dispersão dos valores observados para cada sensor ou qualquer outra variável disponível para descrever os produtos.

4.4 CONCLUSÕES

O classificador binário obteve um desempenho quase três vezes superior ao observado pelo classificador ordinal. Portanto, a maneira como é feita a representação das classes mostrou ser uma escolha muito importante para o sucesso de um classificador perceptron de múltiplas camadas. Outra opção que, comprovadamente através dos dados obtidos, aumenta o desempenho e a estabilidade das redes neurais é o uso da média de ensemble. A metodologia de otimização foi eficiente para o ajuste dos melhores parâmetros de rede. As redes otimizadas convergiram para um valor comum de porcentagem de classificação correta indicando uma sobreposição das classes dos dados utilizados. O mapa auto-organizável permitiu verificar que os produtos que foram incorretamente classificados estão distribuídos em regiões muito próximas no espaço de entrada multidimensional das variáveis de entrada. Além disso, a análise estatística das relações de vizinhança em conjunto com os mapas de peso permite propor hipóteses sobre a similaridade aromática, de acordo com o nariz eletrônico, para os produtos avaliados. A combinação do nariz eletrônico com as redes neurais artificiais mostrou ser uma alternativa promissora para o desenvolvimento de novos produtos, comparação de similaridades entre produtos concorrentes e para o controle da qualidade aromática de café solúvel.

4.5 REFERÊNCIAS

Basheer, I. A.; Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43, 3-31.

Bassoli, D. G. (2006). *Impacto aromático dos componentes voláteis do café solúvel: uma abordagem analítica e sensorial*. Londrina: Tese (Doutorado em Ciência de Alimentos), Universidade Estadual de Londrina (UEL), 214p.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford: Oxford University, 482p.

Bodt, E.; Cottrell, M.; Verleysen, M. (2002). Statistical tools to assess the reliability of self-organizing maps. *Neural Networks*, 15, 967-978.

Bomio, M. (1998). Neural networks and the future of sensory evaluation. *Food Technology*, 52, 62-63.

Bona, E. (2008). Emprego de mapas auto-organizáveis no reconhecimento de padrões aromáticos de café solúvel. *Ciência e Tecnologia de Alimentos*, (será submetido).

Bona, E.; Borsato, D.; Silva, R. S. S. F.; Herrera, R. P. (2000). Aplicativo para otimização empregando o método simplex seqüencial. *Acta Scientiarum*, 22, 1201-1206.

Chandraratne, M. R.; Kulasiri, D.; Samarasinghe, S. (2007). Classification of lamb carcass using machine vision: comparison of statistical and neural network analyses. *Journal of Food Engineering*, 82, 26-34.

Curry, B. & Morgan, P. H. (2006). Model selection in Neural Networks: Some difficulties. *European J. of Operational Research*, 170, 567-577.

Deisingh, A. K.; Stone, D. C.; Thompson, M. (2004). Applications of electronic noses and tongues in food analysis. *International J. of Food Sci. and Technol.*, 39, 587-604.

Della Lucia, S. M. & Minim, L. A. (2006). In: Della Lucia, S. M. (Ed.), *Análise Sensorial: Estudos com Consumidores*. Viçosa: UFV, pp.173-194, (Chapter 7).

Derringer, G. & Suich, R. (1980). Simultaneous optimization of several response variables. *J. of Quality Technology*, 12, 214-219.

Díaz, C.; Conde, J. E.; Estévez, D.; Olivero, S. J. P.; Trujillo, J. P. P. (2003). Application of multivariate analysis and artificial neural networks for the differentiation of red wines from the Canary Islands according to the Island of origin. *Journal of Agricultural and Food Chemistry*, 51, 4303-4307.

Dutcosky, S. D. (2007). *Análise Sensorial de Alimentos*. 2th ed., Curitiba: Champagnat, 239p.

Dutta, R.; Kashwan, K. R.; Bhuyan, M.; Hines, E. L.; Gardner, J. W. (2003). Electronic nose based tea quality standardization. *Neural Networks*, 16, 847-853.

Hair Jr., J. F.; Anderson, R. E.; Tatham, R. L.; Black, W. C. (2005). *Análise Fatorial*. In: _____. *Análise Multivariada de Dados*. 5th ed. Porto Alegre: Bookman, cap. 3, p.89-127.

Haykin, S. (2001). *Redes Neurais: Princípio e Práticas*. 2th ed., Porto Alegre: Bookman, 900p.

- Hai Z. & Wang, J. (2006). Electronic nose and data analysis for detection of maize oil adulteration in sesame oil. *Sens. and Actuators B*, 119, 449-455.
- Huang, Y.; Kangas, L. J.; Rasco, B. A. (2007). Applications of artificial neural networks (ANNs) in food science. *Critical Reviews in Food Sci. and Nutrition*, 47, 113-126.
- Kolman, E. & Margaliot, M. (2005). Are artificial neural networks white boxes? *IEEE Transactions on Neural Networks*, 16, 844-852.
- Liao, S. (2005). Expert system methodologies and applications – a decade review from 1995 to 2004. *Expert Systems with Applications*, 28, 93-109.
- Linder, R. & Pöppel, S. J. (2003). A new neural network approach classifies olfactory signals with high accuracy. *Food Quality and Preference*, 14, 435-440.
- Looney, C. G. (1996). Advances in feedforward neural networks: demystifying knowledge acquiring black boxes. *IEEE Transactions on Knowledge and Data Engineering*, 8, 211-226.
- Ludermir, T. B.; Yamazaki, A.; Zanchetin, C. (2006). An optimization methodology for neural network weights and architectures, *IEEE Transactions on Neural Networks*, 17, 1452-1459.
- Maria, C. A. B.; Moreira, R. F. A.; Trugo, L. C. (1999). Componentes voláteis do café torrado. Parte I: Componentes heterocíclicos. *Química Nova*, 22, 209-217.
- Marini, F.; Magrì, A. L.; Bucci, R.; Magrì, A. D. (2007) Use of different artificial neural networks to resolve binary blends of monocultivar Italian olive oils. *Analytica Chimica Acta*, 599, 232-240.
- Mello, A. A. & Trugo, L. C. (2003). In: M. R. B. Franco (Ed.), *Aroma e Sabor de Alimentos*. São Paulo: Livraria Varela, p.169-175, (Chapter 12).
- Monge, M. E.; Bulone, D.; Giancomazza, D.; Bernik, D. L.; Negri, R. M. (2004). Detection of flavour release from pectin gels using electronic noses. *Sens. And Actuators B*, 101, 28-38.
- Moreira, R. F. A.; Trugo, L. C.; Maria, C. A. B. (2000). Componentes voláteis do café torrado. Parte II. Compostos alifáticos, alicíclicos e aromáticos. *Química Nova*, 23, 195-203.
- Ni, H. & Gunasekaran, S. (1998). Food quality prediction with neural networks. *Food Technology*, 52, 60-65.
- Panigrahi, S.; Balasubramanian, S.; Gu, H.; Logue, C.; Marchello, M. (2006). Neuralnetwork- integrated electronic nose system for identification of spoiled beef. *LWT*, 39, 135-145.

Pardo, M. & Sberveglieri, G. (2002). Coffee analysis with an electronic nose. *IEEE Transactions on Instrumentation and Measurement*, 51, 1334-1339. 135

Ramírez, E. F. F. (2005). Modelo conexionista para avaliação de propostas para aquisição de equipamentos médico-hospitalares. Londrina: Tese de Doutorado, (Universidade de Campinas).

Sablani, S. S. (2008). Status of observational models used in design and control of products and processes. *Comprehensive Reviews in Food Science and Food Safety*, 7, 130-136.

Sperandio, D.; Mendes, J. T.; Silva, L. H. M. (2003). *Cálculo Numérico: Características Matemáticas e Computacionais dos Métodos Numéricos*. São Paulo: Prentice Hall, 354p.

StatSoft, Inc. (2005). STATISTICA (data analysis software system), version 7.1. www.statsoft.com.

Sukthomya, W. & Tannock, J. (2005). The optimization of neural network parameters using Taguchi's design of experiments approach: an application in manufacturing process modeling. *Neural Computation & Application*, 14, 337-344.

Tigrine-Kordjani, N.; Chemat, F.; Meklati, B. Y.; Tuduri, L.; Giraudel, L.; Montury, M. (2007). Relative characterization of rosemary samples according to their geographical origins using microwave-accelerated distillation, solid-phase microextraction and Kohonen self-organizing maps. *Analytical and Bioanalytical Chemistry*, 389, 631-641.

Trihass J. & Nielsen, P. V. (2005). Electronic nose technology in quality assessment: Monitoring the ripening process of Danish Blue Cheese. *J. of Food Sci.*, 70, 44-49.

Zhang, H.; Wang, J.; Tian, X.; Yu, H.; Yu, Y. (2007). Optimization of sensor array and detection of stored duration of wheat by electronic nose. *J. of Food Engineering*, 82, 403-408.

Zupan, J.; Novič, M.; Ruisánchez, I. (1997). Kohonen and counterpropagation artificial neural networks in analytical chemistry. *Chemometrics and Intelligent Laboratory Systems*, 38, 1-23.

Warner, B. & Misra, M. (1996). Understanding neural networks as statistical tools. *The American Statistician*, 50, 284-293.

Windeatt, T. (2006). Accuracy/Diversity and Ensemble MLP Classifier Design. *IEEE Transactions on Neural Networks*, 17, 1194-1211.

4.6 AGRADECIMENTOS

Agradeço à Fundação Araucária pelo suporte financeiro, à CAPES pela bolsa de doutorado e à Companhia Iguazu de Café Solúvel que, representada pelo Dr. Denisley Gentil Bassoli, forneceu todo o suporte experimental e analítico necessário para este trabalho.

5 CONCLUSÃO GERAL

5.1 CONCLUSÕES GERAIS

Neste trabalho foi observado que a associação de testes estatísticos, métodos de reamostragem e de otimização às redes neurais artificiais permite superar algumas das desvantagens associadas a esta metodologia. Com relação à grande quantidade de dados necessários para o treinamento das redes, a reamostragem bootstrap possibilitou a geração de conjuntos artificiais de dados para o treinamento, preservando as amostras originais para as etapas validação e teste. Além disso, com o procedimento de reamostragem é possível realizar o treinamento de um conjunto de redes com o objetivo de avaliar estatisticamente a confiabilidade dos resultados e de construir comitês de especialistas. Ficou comprovado que o uso da média de ensemble aumenta o desempenho e reduz a variabilidade das respostas apresentadas pelas redes. A dificuldade para escolher o melhor conjunto de parâmetros de rede foi contornada usando a otimização simplex seqüencial em conjunto com as funções de desejabilidade, as quais permitiram realizar simultaneamente a maximização de desempenho e a minimização do tamanho das redes.

As redes perceptron de múltiplas camadas apresentaram uma performance satisfatória para o apoio na tomada de decisão durante a classificação de cafés solúveis em ambos os bancos de dados utilizados. Para o banco de dados industrial, que continha vários produtos, ficou evidente que a maneira como as classes são representadas é fundamental para o sucesso do classificador.

Os mapas auto-organizáveis apresentaram resultados semelhantes à análise de componentes principais para o primeiro banco de dados. Já para o banco industrial, os mapas permitiram uma melhor separação e interpretação das relações de similaridade entre os produtos avaliados. Vale a pena salientar que a associação do teste de hipótese permite verificar a consistência estatística das regras de vizinhança propostas pelo mapa. O mapa topológico auxilia, também, na visualização da sobreposição e dispersão dos produtos no espaço multidimensional das variáveis de entrada.

Concluindo, a combinação do nariz eletrônico com as redes neurais artificiais mostrou ser uma alternativa promissora para o desenvolvimento de novos produtos, comparação de similaridades entre produtos concorrentes e para o controle da qualidade aromática de café solúvel.

5.2 SUGESTÕES DE TRABALHOS FUTUROS

De acordo com a revisão bibliográfica, observações realizadas durante a elaboração desta tese e sugestões dos membros da banca examinadora foi elaborada uma listagem de trabalhos que podem dar continuidade ao estudo deste tema.

- ✓ Avaliar o desempenho das redes de função base radial na classificação dos cafés em ambos os bancos de dados.
- ✓ Implementar outros métodos para a construção de comitês de redes e comparar o desempenho com a média de *ensemble*.
- ✓ Testar critérios para avaliar a sensibilidade das redes em relação as variáveis de entrada.
- ✓ Aplicar técnicas de poda de rede na tentativa de reduzir o tamanho das redes obtidas.
- ✓ Avaliar a influência de amostras discrepantes no resultado obtido pelos classificadores neurais.
- ✓ Ao perfil aromático obtido pelo nariz eletrônico, acrescentar outras variáveis que possam aumentar o poder de discriminação dos cafés pela rede neural artificial.
- ✓ Aplicar os métodos desenvolvidos a dados de análise sensorial dos cafés estudados para verificar se existe uma correlação com os resultados obtidos através da análise do perfil aromático fornecido pelo nariz eletrônico.

APÊNDICES

APÊNDICE A

OTIMIZAÇÃO SIMPLEX

O princípio do simplex básico, descrito por Spendley et al. (1962), foi deslocar uma figura regular, por exemplo, um triângulo equilátero sobre uma superfície, quando duas variáveis (fatores) estavam sendo consideradas (Figura 1(a)).

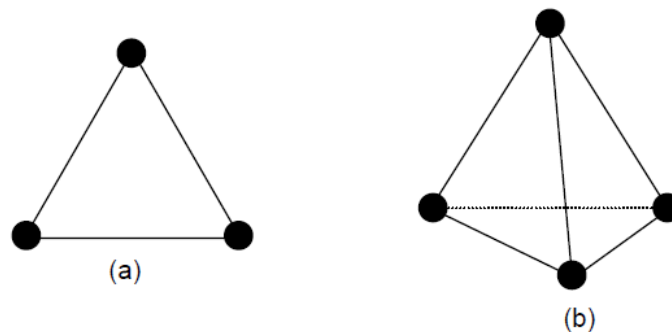


Figura 1 – Interpretação geométrica do simplex (a) para duas variáveis (b) três variáveis.

Nelder e Mead (1965) propuseram o simplex modificado, que nada mais é do que o simplex básico, alterado em seu tamanho e forma, para adaptar-se melhor à superfície de resposta. Recentemente, Pires et al. (1998) desenvolveram um aplicativo para microinformática, introduzindo uma inicialização automática para o simplex modificado.

Segundo o método descrito por Pires et al. (1998) inicia-se a otimização (maximização ou minimização) atribuindo-se limites inferiores (L_i) e superiores (U_i) para cada fator que será controlado. Os níveis inferiores são escolhidos como primeiro vértice. Spendley et al. (1962) apresentaram uma técnica que permite calcular as demais coordenadas do simplex inicial segundo as equações (1) e (2), onde n é o número de variáveis (contínuas ou qualitativas) e t a distância entre dois vértices (geralmente tomada como 1).

$$p = \frac{t}{n\sqrt{2}} (\sqrt{n+1} + n - 1) \quad (1)$$

$$q = \frac{t}{n\sqrt{2}} (\sqrt{n+1} - 1) \quad (2)$$

Nakai et al. (1984) adicionaram ao simplex, denominado supermodificado, os limites inferiores e superiores dos fatores utilizados nas equações (3) e (4).

$$m_1 = L_i + p(U_i - L_i) \quad (3)$$

$$m_2 = L_i + q(U_i - L_i) \quad (4)$$

A partir da geometria analítica pode-se determinar que as coordenadas dos vértices de um simplex regular são representadas pela matriz **M**, na qual as colunas representam os componentes dos vértices, numerados de 1 até n+1 e as linhas representam as coordenadas, i = 1 até n (HIMMELBLAU, 1972).

$$M = \begin{vmatrix} L_i & m_1 & m_2 & m_2 \\ L_i & m_2 & m_1 & m_2 \\ \dots & \dots & \dots & \dots \\ L_i & m_2 & m_2 & m_1 \end{vmatrix} \quad \text{matriz } n \times (n+1)$$

Depois de estabelecido o simplex inicial, o sistema estudado deve ser avaliado nas condições propostas pelo método. Através das repostas obtidas, os vértices são classificados em **B** (melhor), **N** (intermediários) e **W** (pior). O novo simplex será determinado rejeitando-se o vértice correspondente à pior resposta e substituindo-o pela sua reflexão (**R**) na hiperface definida pelos vértices restantes. A direção do movimento de reflexão (Figura 2) é determinada pelo centróide (**P**) formado pelos pontos remanescentes (SPENDLEY et al., 1962).

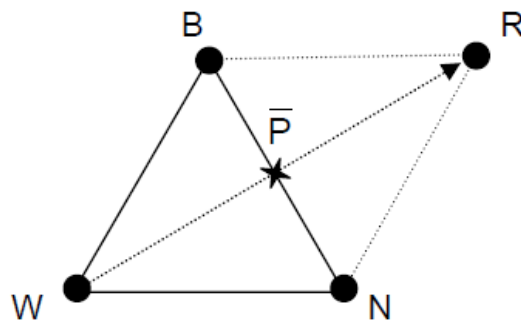


Figura 2 – Movimento de reflexão em um simplex para duas variáveis.

No exemplo ilustrado anteriormente, o centróide pode ser calculado pela equação (5). Segundo Himmelblau (1972), a equação (6), onde α é o coeficiente de reflexão, determina o vértice de reflexão para o caso de duas variáveis. Este coeficiente nada mais é do que um múltiplo ou submúltiplo do segmento de reta que une o centróide a **W**. Nelder e Mead (1965) demonstraram que são necessários um número menor de iterações para otimizar um problema quando $\alpha = 1$.

$$\bar{P} = \frac{B + N}{2} \quad (5)$$

$$R = \bar{P} + \alpha(\bar{P} - W) \quad (6)$$

Para evitar uma expansão ou contração muito rápida do simplex, que poderia afastá-lo do ponto ótimo, é gerado um vetor de movimento em direção ao ótimo (vetor movimento). Este procedimento gera um valor em um ponto intermediário ao valor da próxima reflexão. O novo simplex será formado pelos vértices **B**, **N** e **Z** conforme a equação (7). Onde β é um fator de expansão ($\beta = \alpha + 1$) e β_{opt} é o seu valor ótimo correspondente (BARROS NETO et al., 1995).

$$Z = \beta_{opt} \bar{P} + (1 - \beta_{opt})W \quad (7)$$

Este valor pode ser obtido substituindo-se g na equação (6), derivando-se posteriormente. O resultado obtido é idêntico a equação (8).

$$\beta_{opt} = \left[\frac{R_W - R_{\bar{P}}}{R_R - 2R_{\bar{P}} + R_W} \right] + 0,5 \quad (8)$$

Os valores de R_W , R e R_R são as respostas determinadas no pontos W, \bar{P} e R , respectivamente. Infinitos valores são possíveis para β_{opt} . Algumas faixas de valores, no entanto, são inconvenientes (Figura 3). Os valores de β_{opt} inferiores a -1 ou superiores a 3 representam extrapolações excessivas do simplex. Valores de β_{opt} próximos de zero devem ser evitados, porque resultariam num novo simplex muito parecido com o original. Se o valor de β_{opt} for igual a 1 o simplex perde uma dimensão. O novo ponto Z da equação (7) coincide com o ponto do centróide, e a reflexão não pode ser efetuada. Caso o valor de β_{opt} caia numa dessas faixas ele é descartado e o novo vértice passa a ser definido por uma série de regras de decisão (BARROS NETO et al., 1995).

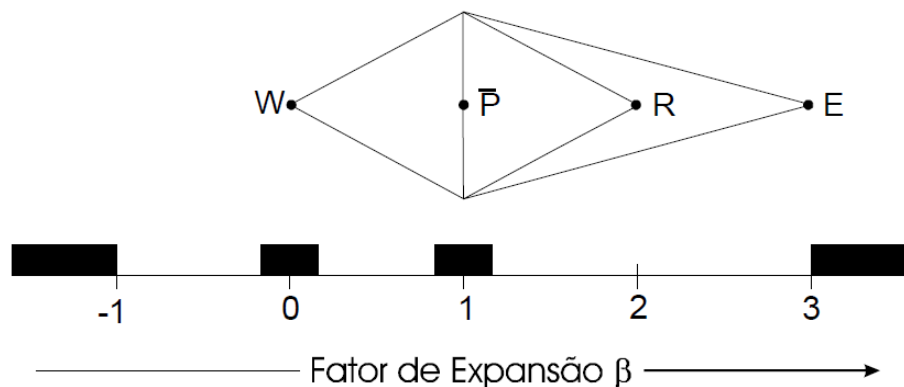


Figura 3 – Intervalos proibidos para os valores de β (BARROS NETO et al., 1995).

APÊNDICE B

FUNÇÕES DE DESEJABILIDADE

As funções gerais desenvolvidas por Derringer e Suich (1980) para a otimização simultânea de diversas variáveis de resposta são conhecidas como funções de desejabilidade ou de conveniência. Tais funções envolvem uma transformação de cada valor de resposta Y_i , para um valor d_i compreendido no intervalo $0 \leq d_i \leq 1$. A transformação é realizada através das seguintes expressões:

$$d_i = \begin{cases} 0 & Y_i \leq a \\ \left[\frac{Y_i - a}{b - a} \right]^s & a < Y_i < b \\ 1 & Y_i \geq b \end{cases} \quad (9)$$

onde a é o limite inferior e b o limite superior da variável de resposta.

O expoente s especificado pelo usuário é usado na transformação como parâmetro de subjetividade. Valores como $0,1 \leq s < 1$ ou $1 \leq s \leq 10$ alteram a importância das respostas, facilitando ou dificultando a busca do ótimo, como pode ser observado na Figura 4.

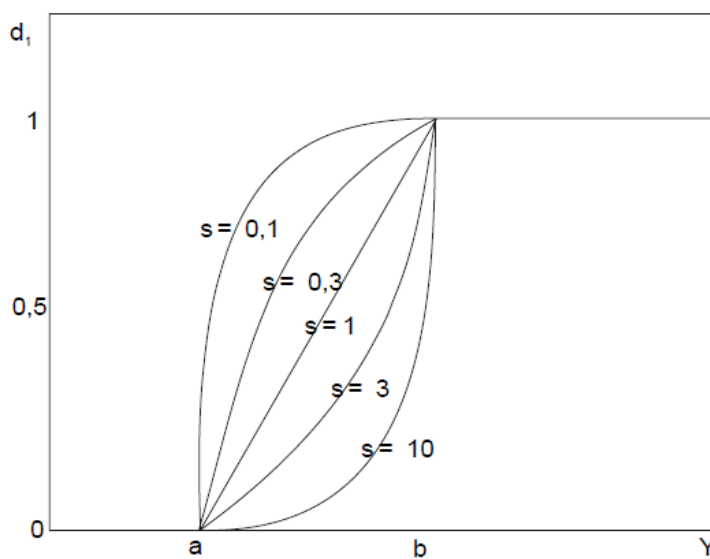


Figura 4 – Gráfico das respostas transformadas para diversos valores de s .

O valor de d_i aumenta proporcionalmente com o aumento da importância relativa de Y_i . As contribuições individuais são combinadas através de uma média geométrica, onde k representa o número de respostas otimizadas:

$$D_M = (d_1 \times d_2 \times \dots \times d_k)^{1/k} \quad (10)$$

O valor único de D_M representa a contribuição global de todas as respostas combinadas. Caso um valor de d_i seja zero toda a função de contribuição global será zerada. Este valor nulo de D_M indica uma combinação insatisfatória das respostas. Por outro lado, se o valor de conveniência for máximo ($d_i = 1$), a função de contribuição global não se altera. Esta condição permite que se continue a otimização das outras variáveis dependentes (DERRINGER; SUICH, 1980).

APÊNDICE C

REFERÊNCIAS

BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. **Planejamento e Otimização de Experimentos**, Campinas: Editora Unicamp, 1995. p. 217-237.

DERRINGER, G.; SUICH, R. Simultaneous optimization of several response variables. **Journal of Quality Technology**, v. 12, n. 4, p. 214-219, 1980.

HIMMELBLAU, D. M. **Applied nonlinear programming**. McGraw-Hill, 1972. p. 148-177.

NAKAI, S.; KOIDE, R.; EUGESTER, K. A. A new mapping super-simplex optimization for food products and process development. **J. Food Sci.** Chicago, v. 49, n. 4, p. 1143-1148, 1170, 1984.

NELDER, J. A.; MEAD, R. A simplex method for function minimization. **Computer J.** v. 7, p. 308-312, 1965.

PIRES, M. V. P.; BORSATO, D. E SILVA, R. S. F. Desenvolvimento de aplicativo para microinformática visando a otimização de sistemas alimentares. In: CONGRESSO BRASILEIRO DE CIÊNCIA E TECNOLOGIA DE ALIMENTOS, 16., 1998, Rio de Janeiro. **Anais**. SBCTA, v. 3, p. 1565-1568, 1998.

SPENDLEY, W.; HIMSWORTH, F. R.; HEXT, G. R. Sequential application of simplex designs in optimization and evolutionary operation. **Technometrics**. v. 4, p. 441-461, 1962.

ANEXOS

ANEXO A

NORMAS PARA PUBLICAÇÃO NA REVISTA CIÊNCIA E TECNOLOGIAS DE ALIMENTOS

Objetivo e política editorial

Ciência e Tecnologia de Alimentos publica artigos e comunicações científicas na área. Os trabalhos deverão ser apresentados em português, inglês ou espanhol, devendo observar as disposições normativas relacionadas abaixo.

Os trabalhos serão submetidos à revisão pela **Comissão Editorial**, sendo que cada trabalho será analisado por dois relatores. Em caso de discordância entre seus pareceres, um terceiro relator será ouvido, e os três pareceres serão analisados pela Diretoria de Publicações que tomará a decisão final.

Os pareceres dos relatores serão encaminhados aos autores para que verifiquem as sugestões e procedam às modificações que se fizerem necessárias. Os trabalhos devem ser submetidos *on-line* através do site www.sbcta.org.br, em *Revista sbCTA, submissão de Artigo*. O fornecimento de separatas deverá ser encomendado previamente à SBCTA.

Informações sobre os trabalhos deverão ser obtidas no programa de submissão *on-line* ou através do e-mail publicacoes@sbcta.org.br.

Normas para a apresentação de trabalhos

INSTRUÇÕES PARA PREPARAÇÃO E ENCAMINHAMENTO DE TRABALHOS¹

(O título abreviado do trabalho, com no máximo 40 caracteres deverá ser incluído) João A NONIMATTO², Mário E. SENOM², Clara O LAST^{2,*}

Endereços para correspondência e dos autores devem constar no rodapé da primeira página.

RESUMO/ SUMMARY Manuscritos sobre pesquisas originais, que mostrem contribuição técnico-científica na forma de artigos ou comunicações, escritos em português, inglês ou espanhol, serão considerados para publicação na revista *Ciência e Tecnologia de Alimentos*, após submissão *on-line* através do site www.sbcta.org.br, se preparados de acordo com as seguintes especificações: 01 arquivo completo do texto, com título, figuras e tabelas, 01 arquivo em pdf, sem os nomes e endereços dos autores, com tabelas e figuras em arquivos separados. Se o artigo for considerado próprio para publicação na Revista C&TA, a versão revisada deverá ser re-submetida ao editor de acordo com as mesmas recomendações acima descritas, também através do programa de submissão. Quando o artigo aprovado estiver na fase de revisão gramatical e diagramação, os autores receberão o artigo para efetuarem as correções finais. A avaliação do artigo será feita após os autores assinarem termo de concordância, conforme [modelo em anexo](#). Nos trabalhos apresentados em português, deve constar também um resumo em inglês, iniciando-se pelo título do trabalho também em inglês. Os trabalhos escritos em inglês ou espanhol deverão trazer um resumo em português, sendo que o resumo em português deverá ser iniciado com o título do trabalho. Resumo e Summary devem ser apresentados separadamente, constando cada um de um único parágrafo e com no máximo 200 palavras.

Palavras-chave/ Keywords. Três a seis palavras-chave deverão ser apresentadas após o sumário e no idioma deste, evitando-se a utilização de termos já usados no título e sumário.

1 - INTRODUÇÃO Os manuscritos submetidos para publicação na revista Ciência e Tecnologia de Alimentos deverão conter resultados de pesquisa original, relacionada à caracterização de novas matérias-primas e ingredientes, identificação de novos componentes ou contaminantes, avaliação de produtos típicos, desenvolvimento, melhoria ou avaliação de processos e equipamentos para obtenção de alimentos tradicionais ou novos produtos. Os trabalhos podem ser apresentados em qualquer uma das três línguas, com texto claro, conciso e de acordo com as seguintes orientações:

1.1 – Apresentação dos arquivos

A apresentação dos arquivos em formato eletrônico sem a devida qualidade prejudicará e prolongará o processo de revisão.

1.2 - Divisão do trabalho O trabalho científico deve conter as seguintes partes:

- Título do trabalho e nome(s) do(s) autor(es)
- Resumo (incluindo as Palavras-chave)
- **Summary** (incluindo o título do trabalho e **Keywords**)
- Introdução
- Material e Métodos
- Resultados e Discussão (podendo ser separados, se necessário)
- Conclusões
- Referências Bibliográficas
- Agradecimentos

A apresentação do texto deverá ser em uma coluna somente. As páginas devem ser numeradas seqüencialmente, bem como as linhas, iniciando-se a numeração destas em cada página.

2 - MATERIAL E MÉTODOS As informações desta seção devem ser consistentes e objetivas, permitindo a outros pesquisadores a identificação ou obtenção da correta matériaprima, o estabelecimento dos mesmos procedimentos e experimentos e reprodução dos resultados obtidos. Os equipamentos especializados e softwares utilizados deverão ser descritos quanto a sua origem (marca, modelo, cidade, país). O cumprimento cuidadoso destas orientações agiliza a avaliação do manuscrito. Títulos e subtítulos em caracteres bem definidos são recomendados, sempre que necessários, mas devem ser usados com critério, sem prejuízo da clareza do texto. As equações devem ser geradas por softwares apropriados e identificadas no texto com algarismos arábicos entre parêntesis (1) de acordo com a ordem que aparecem. As unidades usadas devem estar de acordo com o Sistema Internacional de Unidades.

3 - RESULTADOS E DISCUSSÃO Figuras, Tabelas e Quadros devem ser usados criteriosamente e numerados em seqüência lógica, com numerais arábicos (*Figura 1, Tabela 1, Quadro 1*), preferencialmente nos arquivos (Excel, SAS, TIF, BMP). Arquivos contendo estes itens devem ser enviados em separado, sendo adequadamente identificados com o número da figura (por exemplo, Figura 1.doc). O lugar exato onde Figuras, Tabelas e Quadros deverão ser inseridos deve ser indicado no texto. O autor deverá ser criterioso quanto à definição do número e tamanho destes, uma vez que estes itens são importantes para a compreensão do artigo e influem nos custos de impressão. As Tabelas e Quadros devem ser apresentados no mais simples formato usado, evitando sombreamento, cores ou linhas horizontais extras para itens do mesmo tipo. Nunca use linhas verticais ou diagonais nas Tabelas. Sempre considere a conveniência de consolidar duas ou mais Tabelas em uma, de forma a reduzir o número destas.

A utilização do número de algarismo significativo em Tabelas deve ser criteriosa. A legenda deve ser escrita acima da correspondente Tabela. Combine texto, tabelas e figuras

adequadamente de forma a produzir um texto consistente, de leitura fácil e contínua. Não apresente os mesmos dados na forma de gráfico e tabela.

4 - CONCLUSÕES As conclusões devem ser apresentadas de forma objetiva e clara, permitindo ao leitor a identificação da contribuição científica do trabalho, os pontos mais importantes encontrados, a contribuição para o avanço do conhecimento, ou de aspectos científicos a serem objeto de futuras pesquisas.

Os manuscritos serão avaliados por revisores nacionais e/ou estrangeiros, especialistas na área de cada trabalho. No caso de não ocorrer concordância entre os primeiros revisores, uma terceira opinião será solicitada. Com base nas avaliações dos revisores o Editor tomará a decisão final. Os comentários e questões dos revisores serão encaminhados à consideração do autor, que deverá responder por escrito às questões e comentários dos revisores.

É fundamental a observância minuciosa das normas para diminuir os prazos de tramitação, uma medida que interessa tanto ao autor quanto à Revista.

5 - REFERÊNCIAS BIBLIOGRÁFICAS

5.1 - Citações no texto As citações bibliográficas inseridas no texto devem ser indicadas pelo(s) último(s) nome(s) do(s) autor(es) em letra maiúscula, seguido(s) pelo ano da publicação (ex.: SILVA et al, 2005), sendo que a seguinte regra deverá ser obedecida:

- até 3 (três) autores: citam-se os sobrenomes dos autores;
- mais que três autores, cita-se o sobrenome do primeiro autor, seguido da expressão latina "et al";
- o nome do autor não é conhecido, a entrada é feita pela primeira palavra do título.

5.2 - Citações na lista de referências A literatura citada no texto deverá ser listada em ordem alfabética.

A lista de referências deve seguir o formato estabelecido pela Associação Brasileira de Normas Técnicas (ABNT) em "Regras Gerais de Apresentação" - NBR-6023, de agosto, 2002, resumido a seguir:

5.2.1 - Livros

autor(es), título, edição, local, editora e data de publicação.

BACCAN, N.; ALEIXO, L. M.; STEIN, E.; GODINHO, O. E. S. **Introdução à semimicroanálise qualitativa**, 6ª. edição. Campinas: EDUCAMP, 1995.

5.2.2 - Capítulos de livro

autor(es), título da parte seguido da expressão "in" e da referência completa do livro, ano de publicação, capítulo, paginação.

SGARBIERI, V. C. Composição e valor nutritivo do feijão *Phaseolus vulgaris* L. In: BULISANI, E. A (Ed.) **Feijão: fatores de produção e qualidade**. Campinas: Fundação Cargill, 1987. Cap. 5, p. 257-326.

5.2.3 - Artigos em periódicos e anais

autor(es), título da parte, título da publicação, local da publicação, volume, fascículo, paginação, data de publicação.

KINTER, P. K.; van BUREN, J. P. Carbohydrate interference and its correction in pectin analysis using the m-hydroxydiphenyl method. **J. Food Sci.**, v. 47, n. 3, p. 756-764, 1982.

5.2.4 - Artigos apresentados em encontros científicos

autor(es), título do trabalho apresentado, seguido da expressão "in": nome do evento, numeração do evento, se houver, ano e local (cidade) de realização, título do documento, local, editora, data de publicação e paginação.

JENSEN, G. K.; STAPELFELDT, H. Incorporation of whey proteins in cheese. Including the use of ultrafiltration. In: INTERNATIONAL DAIRY FEDERATION. **Factors Affecting the Yield of Cheese**. 1993, Brussels: International Dairy Federation Special Issue, n. 9301, chap. 9, p. 88-105.

5.2.5 - Dissertações, teses e relatórios

autor, título em negrito, local, ano da defesa, número de páginas, tese (grau e área), departamento, instituição.

CAMPOS, A C. **Efeito do uso combinado de ácido láctico com diferentes proporções de fermento láctico mesófilo no rendimento, proteólise, qualidade microbiológica e propriedades mecânicas do queijo minas frescal**. Campinas, 2000, 80p. Dissertação (Mestre em Tecnologia de Alimentos), Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas (UNICAMP).

5.2.6 - Trabalhos em meio-eletrônico

as referências devem obedecer aos padrões indicados, acrescidas das informações relativas à descrição física do meio eletrônico (disquetes, cd-room, on-line, etc.).

SÃO PAULO (Estado). Secretaria do Meio Ambiente. Tratados e organizações ambientais em matéria de meio ambiente. In: _____. **Entendendo o meio ambiente**. São Paulo, 1999. v. 1. Disponível em: <<http://www.bdt.org.br/sma/entendendo/atual.htm>>. Acesso em: 8 mar. 1999.

5.2.7 – Legislação

jurisdição e órgão judiciário competente, título, número, local, data e dados da publicação.
BRASIL. Portaria n. 451, de 19 de setembro de 1997. Regulamento técnico princípios gerais para o estabelecimento de critérios e padrões microbiológicos para alimentos. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 22 set. 1997, Seção 1, n. 182, p. 21005-21011.

6 - AGRADECIMENTOS Agradecimentos e outras formas de reconhecimento podem ser mencionados após a lista de referências.

Manuscritos devem ser submetidos *on-line* através do site www.sbcta.org.br, em *Revista sbCTA, submissão de Artigo*.

Qualquer dúvida entrar em contato com publicações@sbcta.org.br.

TERMO DE CONCORDÂNCIA E CESSÃO DE DIREITOS DE REPRODUÇÃO GRÁFICA

Os abaixo assinados, (nomes completos dos autores do manuscrito), intitulado "título", declaram ter lido e aprovado o manuscrito na sua totalidade e concordam em submetê-lo à revista *Ciência e Tecnologia de Alimentos* para avaliação e possível publicação como resultados originais. Esta declaração implica que o manuscrito, independente do idioma, não foi submetido a outros periódicos ou revistas com a mesma finalidade.

"Declaro(amos) que aceito(amos) ceder o direito de reprodução gráfica para a Sociedade Brasileira de Ciência e Tecnologia de Alimentos (sbCTA) no caso do artigo com o título descrito acima (ou título que posteriormente chegar a ser adotado, para atender às sugestões de editores e revisores) venha ser publicado por periódico ou revista da sbCTA.

Em adição (necessário se existir mais que um autor), concordamos em nomear _____ como sendo o autor a quem toda correspondência e separatas deverão ser enviadas."

Cidade Endereço Data Nomes e assinaturas

ANEXO B

NORMAS PARA PUBLICAÇÃO NA REVISTA SENSORS AND ACTUATORS B – CHEMICAL

Sensors & Actuators, B: Chemical is an interdisciplinary journal dedicated to covering research and development in the field of chemical sensors, actuators, micro- and nanosystems. The scope of the journal encompasses, but is not restricted to, the following areas:

- . Sensing principles and mechanisms
- . New materials development (transducers and sensitive/recognition components)
- . Fabrication technology including nanotechnology
- . Actuators
- . Optical devices
- . Electrochemical devices
- . Mass-sensitive devices
- . Gas sensors
- . Biosensors
- . Bio-MEMS
- . Analytical microsystems
- . Environmental
- . Process control
- . Biomedical applications
- . Signal processing
- . Sensor and sensor-array chemometrics

µTAS - Micro Total Analysis Systems *Microsystems for the generation, handling and analysis of (bio)chemical information*. The special section of *Sensors & Actuators, B: Chemical* on µTAS is dedicated to contributions concerning miniaturised systems for (bio)chemical synthesis and analysis, also comprising work on Bio-MEMS, Lab-on-a-chip, biochips and *microfluidics*. Topics covered by the µTAS section include:

- . Lab-on-a-chip
- . Physics and chemistry of microfluidics
- . Microfabrication technology for µTAS
- . Analytical chemical aspects
- . Detectors, sensors, arrays for µTAS
- . µTAS applications
- . DNA analysis
- . Microinstrumentation
- . Microsystems for combinatorial chemistry

Types of contribution

The journal publishes research papers, letters to the Editors and occasionally review articles. Short reports on current research can be submitted as a letter to the Editors. These should not exceed 2000 words or 4 printed pages. All papers will be reviewed by at least two independent referees. For all contributions the acceptance criteria are quality, originality, and scientific and technological relevance to the field. An adequate referencing to the state-of-the-art is essential. All contributions must be written in English.

Submission of Contributions - Online Submission of Papers

Authors are encouraged to submit their manuscript online to one of the editors by using the online submission tool for *Sensors and Actuators B: Chemical* at <http://www.elsevier.com/locate/snb>. To submit online, authors are required to go to this

website and upload their article (compuscript with figures in line) and separately its associated artwork, an electronic (PDF) proof is generated and the reviewing process is carried out using that PDF. Authors and editors send and receive all correspondence by email via the website and no paper correspondence is performed. Full instructions on how to use the online submission tool and how to prepare your manuscript for online submission are available at: <http://www.elsevier.com/locate/snb>.

Submission of Papers By Mail

Authors should submit three copies of their manuscripts, one complete set of original illustrations and two copies to the Editors. For the final version, in addition to the original and two copies, authors should submit an electronic version of their manuscript on **disk**. Papers should be sent to the Editor-in-Chief or the appropriate Regional Editor.

Editor-in-Chief:

Professor Milena Koudelka-Hep

Institute of Microtechnology

University of Neuchatel

Rue Jaquet-Droz 1

CH-2007 Neuchatel

Switzerland

Tel: +41 32 7205 305

Fax: +41 32 7205 711

E-mail: Milena.Koudelka@unine.ch

Regional Editor for North America:

Professor Marc Madou

Mechanical and Aerospace Engineering

University of California

Hery Samuelli School of Engineering

Irvine

CA 92697-3975

USA

Tel: +1-949-824-6585

Fax: +1-949-824-8585

E-mail: madsense@aol.com

Regional Editor for Asia:

Professor M. Egashira

Department of Materials Science and Engineering

Faculty of Engineering

Nagasaki University

1-14 Bunkyo-machi, Nagasaki 852-8521, Japan

Tel: +81-95-819-2642

Fax: +81-95-819-2643

E-mail: egashira@net.nagasaki-u.ac.jp

Papers for the iTAS Section

Send to Associate Editor uTAS Section:

Professor Shuichi Shoji

Department of Electrical Engineering and Bioscience

Major in Nano-science & Nano-engineering

Building 61 Room 411

Waseda University

3-4-1, Okubo

Shinjuku-ku

169-8555 Tokyo

Japan

Contributions are accepted on the understanding that authors have obtained the necessary authority for publication. Submission of an article is understood to imply that the article is

original and unpublished and is not being considered for publication elsewhere. Upon acceptance of an article by the journal, the author(s) will be asked to transfer the copyright of the article to the publisher. This transfer will ensure the widest possible dissemination of information. Authors are reminded that delays in publication may occur if the instructions for submission and disk and manuscript preparation are not strictly followed. Authors are strongly recommended to submit disks to aid rapid processing. To facilitate communication, authors are requested to provide their current **e-mail address**, **telephone** and **fax number**.

Manuscript Preparation

General

All papers should be concisely written. Papers should be written in English (American or British spelling but with use of only one form in the same paper). The author should remember that the journal is international and read widely by those whose first language may be other than that in which the paper is written. Clarity and precision are best achieved by the use of short words and simple sentences. Papers which do not satisfy the English language requirements will not be accepted. Authors in Japan kindly note that, upon request, Elsevier Japan will provide a list of people who can check and improve the English of an article before submission. For further information please contact our Tokyo office: Elsevier Japan K.K., 1-9-15 Higashi Azabu, Minato-ku, Tokyo 106-0044, Japan; tel.: +81-3-5561-5032; fax: +81-3-5561-5045; e-mail: jp.info@elsevier.com

When submitting their paper authors are requested to provide names and addresses of three competent **referees**, although the choice of referees used will be made by the editor.

Formats

The typescript should be in *double-spaced typing* on pages of uniform size with a wide margin on the left. This applies also to tables, legends for illustrations, references and footnotes. The margin and double spacing greatly facilitate editorial processing. Each table should be typed on a separate page, and the legends to illustrations should be typed in sequence on a separate page, widely spaced. Typescripts should be preceded by a page bearing the name and address of the person to whom proofs are to be sent and indicating the number of pages in the typescript. Some flexibility of presentation will be allowed, but authors are urged to arrange the subject matter clearly and logically under such headings as **Introduction**, **Experimental**, **Results**, **Discussion**, etc.

Title

Papers should be headed by a concise but informative title. This should be followed by the names of the authors and by the name and address of the laboratory in which the work was performed. If the address of the author at the time when the paper will appear will be other than that where the work was carried out, this may be stated in a footnote. Acknowledgements for financial support should not be made by a footnote to the title or name of the author but should be included in **Acknowledgements** at the end of the paper.

Abstract

All papers should have an **Abstract** on a separate sheet. The abstract (preferably 50-200 words) should comprise a brief and factual account of the contents and conclusions of the paper as well as an indication of any new information presented and its relevance. Complete sentences should be used, without unfamiliar abbreviations or jargon. The use of the present tense is customary.

Keywords

Authors are requested to provide 4 to 6 keywords. These should follow the Abstract.

Introduction

All papers should have a short **Introduction**. This should state the reasons for the work, with brief reference to previous work on the subject.

References

The references should be numbered consecutively throughout the text and should be collected together in a reference list (headed **References**) at the end of the paper. The list of references should be given on a separate sheet of the manuscript. Footnotes and legends should not include bibliographic material, and reference lists should not include material that could more appropriately appear as a footnote. When appropriate, authors may refer to

material available on the World Wide Web by citing the corresponding URL. Authors should ensure that every reference appearing in the text is in the list of references and vice versa. Numerals for references are enclosed in square brackets in the text, e.g., [1]; numerals referring to equations are enclosed in parentheses. The abbreviated titles of periodicals should conform to standard abbreviations such as those given in the INSPEC Science Abstracts Lists of Journals, regularly appearing in *Electrical and Electronics Abstracts*. In the reference list, periodicals [1], books [2], multi-author books [3] and conference proceedings [4] should be cited in accordance with the following examples.

[1] C. di Natale, F.A.M. Davide, A. D'Amico, W. Göpel, U. Weimar, Sensor array calibration with enhanced neural networks, *Sens. Actuators, B, Chem* 18-19 (1994) 654-657.

[2] A. Nadai, *Theory of Flow and Fracture of Solids*, vol. 1, 2nd ed., McGraw-Hill, New York, 1950, p. 350.

[3] B. Danielsson and K. Mosbach, in: K. Mosbach (Ed.), *Methods in Enzymology*, vol. 137, Academic Press, New York, 1988, pp. 181-197 (Chapter 16).

[4] K.E. Petersen, Silicon sensor technologies, *Tech. Digest, IEEE Int. Electron Devices Meet.*, Washington, DC, USA, Dec. 2-7, 1985.

A reference to "to be published in [title of periodical]" or "in press" implies that the paper has already been accepted for publication. A name appearing in the text which refers to a person as originator of an unpublished idea is listed in the References as a "personal communication". In the text an author's name is given without initials except where it is wished to avoid confusion with namesakes. When reference is made to a publication written by more than two authors it is preferable to give only the first author's name in the text followed by et al or the name of one of the authors followed by 'and coworkers'. In the list of references the names and initials of all authors must be given. This journal should be cited as ***Sensors and Actuators B, Chemical***.

Tables

Careful thought should be given to the layout of tables (and figures) so that the significance of the results may be quickly grasped by the busy reader. It should also be remembered that the length of a printed page is always greater than its width. Tables with only two or three headings are best printed horizontally.

Tables are to be numbered serially with arabic numerals, and should have headings which make their general meaning understandable without reference to the text. The units in which results are expressed should be given in parentheses at the top of each column and should not be repeated on each line of the table. Footnotes should be indicated by the use of lower case letters (a, b, c, etc.) as superscripts without parentheses.

Illustrations

Original line drawings and cyclic or aromatic chemical formulae should be in a form suitable for direct reproduction, prepared with a good contrast (i.e. black on a white background). Photographs and electronmicrographs should be black and white glossy prints and as rich in contrast as possible. Where magnifications are concerned, it is preferable to indicate the scale by means of a ruled line on the photograph. There is no need to specify reductions for figures; however, figures will generally be reduced in size before printing and authors are urged to ensure that any lettering is sufficiently large (minimum height 3-5 mm) to remain legible. Care should be taken when submitting computer graphics to ensure that labelling is of sufficient size and quality. All illustrations should preferably require the same degree of reduction and be submitted on paper of the same size, or smaller than the main text to prevent damage in transit. Legends to illustrations should be typed in sequence on a separate page or pages and be understandable without reference to the text. All illustrations should be clearly referred to in the text using Arabic numerals.

Colour Illustrations

Colour in print - please submit colour illustrations as original photographs, high-quality computer prints, transparencies or high resolution electronic files close to the size expected in publication, or as 35mm slides. Polaroid colour prints are not suitable. For colour reproduction in print, you will receive information regarding the costs from Elsevier after receipt of your accepted article. For further information on the preparation of electronic

artwork, please see <http://www.elsevier.com/artworkinstructions> Free colour on the web - if, together with your accepted article, you submit usable colour figures then Elsevier will ensure, at no additional charge, that these figures will appear in colour on the web (e.g., ScienceDirect and other sites) regardless of whether or not these illustrations are reproduced in colour in the printed version. Please note that if you do not opt for colour in print, you should submit relevant figures in both colour (for the web) and black and white (for print).

Supplementary data

Sensors and Actuators B: Chemical now accepts electronic supplementary material to support and enhance your scientific research. Supplementary files offer the author additional possibilities to publish supporting applications, movies, animation sequences, high-resolution images, background datasets, sound clips and more. Supplementary files supplied will, subject to peer review, be published online alongside the electronic version of your article in Elsevier web products, including ScienceDirect: www.sciencedirect.com. The presence of these files will be signified by a footnote to the article title, and by a description included in a 'Supplementary Data' section at the end of the paper. In order to ensure that your submitted material is directly usable, please ensure that data is provided in one of our recommended file formats and supply a concise and descriptive caption for each file. Please also clearly indicate whether data files are either i) for publication online or ii) only to be used as an aid for the refereeing of the paper. For more detailed instructions please visit our Author Gateway at <http://authors.elsevier.com>.

Biography

A short biography of the author(s) should follow the References. It should cover such information as the author(s) degree(s), where and in what year they were obtained, present employment, and current fields of interest. Photographs are optional.

Submission of electronic text

Preparation of manuscripts on disk

Articles prepared using any of the more popular word-processing packages are acceptable but please note the following points:

- . Submissions should be made on a double-density or high-density 3.5" disk.
- . The disk format, word-processor format, file name(s) and the title and authors of the article should be indicated on the disk.
- . The disk should always be accompanied by a hard-copy version of the article, and the content of the two should be identical.
- . The disk text **must** be the same as that of the final refereed, revised manuscript.
- . Disks formatted for either IBM PC compatibles or Apple Macintosh are preferred. If you can provide either of these, our preference is for the former.
- . The article should be saved in the native format of the word processor used, e.g. WordPerfect, Microsoft Word, etc.
- . Although most popular word processor file formats are acceptable, we cannot guarantee the usability of all formats. If the disk you send us proves to be unusable, we will publish your article from the hard copy.
- . Please do not send ASCII files as relevant data may be lost.
- . There is no need to spend time formatting your article so that the printout is visually attractive (e.g. by making headings bold), as most formatting instructions will be removed upon processing.
- . Leave a blank line between each paragraph and between each entry in the list of bibliographic references.
- . Tables should preferably be placed in the same electronic file as the text.
- . Graphics. We are processing graphic files in a growing number of cases. Both scanned and computer-generated illustrations, either in colour or black and white are acceptable. Graphics in electronic format must be TIFF (Tagged Image File Format) or JPEG; charts and graphs (line art) should be 1000 dpi, photographs 300 dpi, minimum. For more information please see <http://authors.elsevier.com>.

Symbols, formulae and equations

Symbols, formulae and equations should be written with great care, capitals and lower case letters being distinguished where necessary. Also a clear distinction in typewritten text should be made between the figure 1 (one) and the lower case l (ell), the letters "o" and zero, "k" and kappa, "u" and mu, "v" and nu, and "n" and eta. Particular care should be taken in writing mathematical expressions containing superscripts and subscripts. Greek letters and unusual symbols employed for the first time should be defined by name in the left-hand margin. The solidus / may be used in equations to economize vertical space but its use should be consistent.

For example:

$$A/b = x^2 / (u + v)^{1/2}$$

It is recommended that natural logarithms should be denoted by ln while decade logarithms should be denoted by lg.

Exponentials are better written as exp(a) than ea. The multiplication sign should be used in floating point numbers to avoid confusion, i.e., 4.25 x 10⁵, not 4.25.10⁵. The decimal point should always be denoted by a full stop.

Abstracting Services

This journal is cited by the following Abstracting Services: Analytical Abstracts, Cambridge Scientific Abstracts, Chemical Abstracts, Compendex, Computer and Control Abstracts, Current Contents, EIC/Intelligence, Electrical and Electronic Abstracts, Engineered Materials Abstracts, FIZ Karlsruhe, Metals Abstracts, PASCAL/CNRS, Physics Abstracts, Science Citation Index, The Engineering Index Annual, The Engineering Index Monthly.

Spellings used for some common words

Aging

Antireflection

Artifact

bandbending

bandgap

bandwidth

co-evaporate

cross section

cross-sectional

crosstalk

feedback (adj.)

flat-band (adj.)

Gaussian

Kirchhoff

Lifetime

Linewidth

Microelectronics

micromechanics

midpoint

multilayer

multi-target

non-crystalline

n-type (adj.)

open-circuit (adj.)

photoemission

photogenerate

photoresist

p-type (adj.)

printout

readout

reverse-bias (adj.)

rod-like (adj.)

semicontinuous

short-circuit (adj.)
single-crystal (adj.)
stepwise
submicron
thermoelectric
ultrahigh
waveband
waveform
wavelength
wavenumber

Proofs and Articles in Press

Proofs will be despatched via e-mail to the corresponding author, by the Publisher and should be returned with corrections as quickly as possible, normally within 48 hours of receipt. Proofreading is solely the author's responsibility. Authors should ensure that corrections are returned in one communication and are complete, as subsequent corrections will not be possible. Any amendments will be incorporated and the final article will then be published online as an Article in Press on ScienceDirect www.sciencedirect.com. For more information on proofreading please visit our proofreading page on <http://authors.elsevier.com/>

Articles in Press take full advantage of the enhanced ScienceDirect functionality, including the ability to be cited. This is possible due the innovative use of the DOI article identifier, which enables the citation of a paper before volume, issue and page numbers are allocated. The Article in Press will be removed once the paper has been assigned to an issue and the issue has been compiled.

Electronic Offprints (E-offprints)

The corresponding author, at no cost, will be provided with a PDF file of the article via e-mail or, alternatively, 25 free paper offprints. The PDF file is a watermarked version of the published article and includes a cover sheet with the journal cover image and a disclaimer outlining the terms and conditions of use.

Copyright

Upon acceptance of an article, Authors will be asked to transfer copyright (for more information on copyright see <http://authors.elsevier.com>). This transfer will ensure the widest possible dissemination of information. A letter will be sent to the corresponding Author confirming receipt of the manuscript. A form facilitating transfer of copyright will be provided. If excerpts from other copyrighted works are included, the Author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by Authors in these cases: contact Elsevier's Rights Department, Oxford, UK: phone (+44) 1865 843830, fax (+44) 1865 853333, e-mail permissions@elsevier.com

. Requests may also be completed on-line via the Elsevier homepage (<http://www.elsevier.com/locate/permissions>).

There are no page charges

Author enquiries

For enquiries relating to the submission of articles (including electronic submission) please visit Elsevier's Author Gateway at <http://authors.elsevier.com>. The Author Gateway also provides the facility to track accepted articles and set up e-mail alerts to inform you of when an article's status has changed, as well as detailed artwork guidelines, copyright information, frequently asked questions and more.

Contact details for questions arising after acceptance of an article, especially those relating to proofs, are provided after registration of an article for publication. Additional information on the journal, including a detailed Guide for Authors may be found on the Elsevier web site <http://www.elsevier.com/locate/snb>. For specific enquiries on the preparation of electronic artwork, Author Frequently Asked questions and any other enquiries relating to Elsevier, please consult the Author Gateway at <http://authors.elsevier.com>

ANEXO C

NORMAS PARA PUBLICAÇÃO NA REVISTA JOURNAL OF FOOD ENGINEERING

The Editor-in-Chief of *Journal of Food Engineering*, R. Paul Singh, distinguished professor of food engineering at the University of California, Davis, whose research has applications in areas ranging from food processing to space exploration, has been elected to the National Academy of Engineering. This is one of the highest professional distinctions for engineers in the United States. Elsevier would like to congratulate him on this prestigious honor.

Guide for Authors

Submission of Papers Submission of all types of manuscripts to *Journal of Food Engineering* proceeds totally online. Via the Elsevier Editorial System (EES) website for this journal (<http://ees.elsevier.com/jfoodeng>) you will be guided step-by-step through the creation and uploading of the various files. The system automatically converts source files to a single Adobe Acrobat PDF version of the article, which is used in the peer review process. Please note that even though manuscript source files are converted to PDF at submission for the review process, these source files are needed for further processing after acceptance. All correspondence, including notification of the Editor's decision and requests for revision, takes place by e-mail generated by EES and via the author's homepage, removing the need for a hard copy paper trail. Authors must submit revisions via EES. All manuscripts must be addressed to the Editor-in-Chief. Any manuscript sent as an email attachment to the Editor-in-Chief will not be processed. The Editor-in-Chief reviews all manuscripts and assigns them to handling Editors to initiate peer review. The authors may be contacted by the Editor-in-Chief or the handling Editor for any required changes before a manuscript is sent to reviewers. Authors may send queries concerning the submission process, manuscript status or journal procedures to authorsupport@elsevier.com.

Questions regarding content of a proposed submission can be directed to the Editor-in-Chief: R. Paul Singh Professor of Food Engineering Department of Biological and Agricultural Engineering University of California 2042 Bainer Hall, One Shields Avenue Davis, CA 95616, USA E-mail: rspingh@ucdavis.edu

Authors may wish to submit, with their manuscripts, the names and full contact details (including e-mail address) of 3 potential referees who may or may not be contacted to review the manuscript. It is expected that the authors who publish in the *Journal of Food Engineering* will be asked to review future manuscripts submitted to the Journal. It is the author's responsibility to ensure that papers are written in clear and comprehensible English. Authors whose native language is not English are strongly advised to have their manuscripts checked by an English-speaking colleague prior to submission. *Language Polishing:* For authors who require information about language editing and copyediting services pre-and post-submission (for which there will be a charge) please visit <http://www.elsevier.com/locate/languagepolishing> or contact authorsupport@elsevier.com for more information. Please note Elsevier neither endorses nor takes responsibility for any products, goods or services offered by outside vendors through our services or in any advertising. For more information please refer to our Terms and Conditions <http://www.elsevier.com/termsandconditions>.

Submission of a manuscript implies that it has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other

language, without the written consent of the Publisher. If a manuscript being submitted was previously rejected by another journal, it must be substantially revised for its technical content.

Review Process A peer review system involving two or three reviewers is used to ensure high quality of manuscripts accepted for publication. The Editor-in-Chief and Editors have the right to decline formal review of the manuscript when it is deemed that the manuscript is 1) on a topic outside the scope of the Journal, 2) lacking technical merit, 3) focused on foods or processes that are of narrow regional scope and significance, 4) fragmentary and provides marginally incremental results, or 5) is poorly written. *Scope Journal of Food Engineering* is a premier journal in the area of food engineering. Hence any manuscript submitted to the journal should have a clear engineering orientation. It could include engineering/design/modeling/evaluation aspects of all aspects of food from preparation to production to processing to consumption. It can also emphasize on all aspects of preservation from traditional canning, freezing, dehydration, and irradiation to more current and emerging technologies such as high pressure processing, and the use of pulsed electric fields. In the context of preservation, processing and handling of foods, all aspects related to chemistry, biochemistry, microbiology, and nutrition are welcome and acceptable. However, traditional scientific manuscripts such as those directly discussing chemistry, biochemistry, microbiology, and nutrition are inappropriate in this journal and hence should be directed to other journals which directly deal with such topics. The submitted manuscripts should have clear science content in methodology, results and discussion as opposed to being a technical report which may be more appropriate in trade journals. Appropriate scientific and statistically sound experimental designs must be included in methodology and statistics must be employed in analyzing data to discuss the impact of test variables. Moreover there should be clear evidence provided on how the given results advance the area of food engineering science and technology. Mere confirmation of existing published data is not acceptable. Manuscripts should not present progress reports; they should present results of completed works.

Types of Contributions Original papers; review articles; case studies; short communications; reports and announcements of conferences and meetings; book reviews and letters to the Editor.

Manuscript Preparation **General:** Most wordprocessing formats are accepted, but Microsoft Word, WordPerfect or LaTeX are preferred. Ensure that any security code/password protection that may have been incorporated into the document is removed. Always keep a backup copy of the electronic file for reference and safety. Save your files using the default extension of the program used. Manuscripts must be typewritten, double-spaced with wide margins. A font size of 12 or 10 pt is required. The corresponding author should be identified (an E-mail address is mandatory - if there is a change to this e-mail contact, the author must notify the publisher as soon as possible). Full postal addresses must be given for all coauthors. Authors should consult a recent issue of the journal for style - sample copies of the journal can be obtained from the journal website <http://www.elsevier.com/locate/jfoodeng>. The Editors reserve the right to adjust style to certain standards of uniformity. Authors should retain a copy of their manuscript since we cannot accept responsibility for damage or loss of papers.

Paper Length: It is recommended that the size of the paper should not exceed 6000 words or about 12 printed pages, including figures and tables. On occasion, manuscripts providing a comprehensive review of a topic important to the field of food engineering will be considered that exceed the suggested word limit. Authors planning to submit a review should first consult with the Editor-in-Chief by providing an abstract of the intended review.

Abstracts: Each paper should be provided with an Abstract of about 100-150 words, stating concisely the purpose and the content of the paper. The abstract should be written as one paragraph. It should include the objectives of the study, experimental approach used, significant results obtained and major conclusions drawn from the work. It must not include a reference to literature citations or acronyms without definition.

Text: Follow this order when typing manuscripts: Title, Authors, Affiliations, Abstract, Keywords, Main text, Acknowledgements, Appendix, References, Figure Captions, Figures and then Tables. Do not import the Figures or Tables into your text. Other than the cover page, every page of the manuscript, including the title page, references, tables etc. should be numbered; however, in the text no reference should be made to page numbers. **Lines must be numbered consecutively throughout the manuscript.** The corresponding author should be identified with an asterisk and footnote. All other footnotes (except for table footnotes) should be identified with superscript Arabic numbers. **Title:** The title of the paper should unambiguously reflect its contents. Where the title exceeds 70 characters a suggestion for an abbreviated running title should be given.

Introduction: Introduction should be brief and specific to the topic of your research. It should include citations of published papers that are most relevant to the work. It must not include description of processes, procedures, or mathematical expressions that are readily available in textbooks or review papers. Introduction section should emphasize on the originality of your work and how it contributes to new information. Introduction section should not exceed 3-4 paragraphs (one to two pages maximum), it should conclude with a list of key objectives of the reported work. The introduction should not describe or outline other sections of the manuscript. **Materials and Methods:** If any devices or instruments are used to obtain experimental data then provide the name of the manufacturer, model, city, state or country in parenthesis. Provide sufficient experimental detail that will allow your experiment to be repeated at another location. If a widely available published procedure is used in its entirety then provide the citation without repeating all the detailed steps. If small changes are made to a published methodology, provide the original reference and elaborate mainly on modifications made by providing appropriate justification. The statistical design employed to obtain experimental data must be clearly indicated with reference to the number of replicates used. If statistical software is used then provide the name and version of the software. Experimental design should include a quantitative discussion of the experimental uncertainty.

Equations and variables: All equations should be centered and sequentially numbered. All variables used in the manuscript should be fully defined in a nomenclature section at the start of the paper only.

Units: The SI system should be used for all scientific and laboratory data; if, in certain instances, it is necessary to quote other units, these should be added in parentheses. Temperatures should be given in degrees Celsius. The unit 'billion' (10^9 in America, 10^{12} in Europe) is ambiguous and should not be used.

References: All publications cited in the text should be presented in a list of references following the text of the manuscript. The manuscript should be carefully checked to ensure that the spelling of authors' names and dates are exactly the same in the text as in the reference list. All citations in text should refer to:

- 1 Single author: the author's name (without initials, unless there is ambiguity) and the year of publication;
- 2 Two authors: both authors' names and the year of publication;
- 3 **Three or more authors: first author's name followed by et al. and the year of publication.**

Citations may be made directly (or parenthetically). Groups of references should be listed first alphabetically, then chronologically. Examples: "as demonstrated (Allan, 1996a, 1996b, 1999); Allan and Jones, 1995). Kramer et al. (2000) have recently shown...." The list of references should be arranged alphabetically by authors' names, then further sorted chronologically if necessary. More than once reference from the same author(s) in the same year must be identified by the letters "a", "b", "c" etc., placed after the year of publication. References should be given in the following form:

Kumbhar, B.K., Agarwal, R.S., & Das, K. (1981). Thermal properties of fresh and frozen fish. *International Journal of Refrigeration*, 4(3), 143-146.

Machado, M.F., Oliveira, F.A.R., & Gekas, V. (1997). Modelling water uptake and soluble solids losses by puffed breakfast cereal immersed in water or milk. In *Proceedings of the Seventh International Congress on Engineering and Food*, Brighton, UK.

Neter, J., Kutner, M.H., Nachtsheim, C.J., & Wasserman, W. (1966). *Applied linear statistical models* (4th ed., pp. 1289-1293). Irwin, Chicago.

Thomson, F. M. (1984). Storage of particulate solids. In M. E. Fayed, L. Otten (Eds.), *Handbook of powder science and technology* (pp. 365-463). Van Nostrand Reinhold, New York.

Citing and listing of web references. As a minimum, the full URL should be given. Any further information, if known (author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

Illustrations

Photographs, charts and diagrams are all to be referred to as "Figure(s)" and should be numbered consecutively in the order to which they are referred. They should accompany the manuscript, but should not be included within the text. All illustrations should be clearly marked with the figure number and the author's name. All figures are to have a caption. Captions should be supplied on a separate sheet.

If, together with your accepted article, you submit usable colour figures then Elsevier will ensure, at no additional charge, that these figures will appear in colour on the web (e.g., ScienceDirect and other sites) regardless of whether or not these illustrations are reproduced in colour in the printed version. For colour reproduction in print, you will receive information regarding the costs from Elsevier after receipt of your accepted article. Please note: Because of technical complications which can arise by converting colour figures to 'grey scale' (for the printed version should you not opt for colour in print) please submit in addition usable black and white prints corresponding to all the colour illustrations.

Preparation of electronic illustrations

General points

- _ Make sure you use uniform lettering and sizing of your original artwork.
- _ Save text in illustrations as "graphics" or enclose the font.
- _ Only use the following fonts in your illustrations: Arial, Courier, Helvetica, Times, Symbol.
- _ Number the illustrations according to their sequence in the text.
- _ Use a logical naming convention for your artwork files.
- _ Provide all illustrations as separate files.
- _ Provide captions to illustrations separately.
- _ Produce images near to the desired size of the printed version. A detailed guide on electronic artwork is available on our website:

<http://www.elsevier.com/artworkinstructions> You are urged to visit this site.

Tables: Tables should be numbered consecutively and given a suitable caption and each table typed on a separate sheet. Footnotes to tables should be typed below the table and should be referred to by superscript lowercase letters. No vertical rules should be used. Tables should not duplicate results presented elsewhere in the manuscript (e.g. in graphs).

Preparation of Supplementary Data

Elsevier now accepts electronic supplementary material (ecomponents) to support and enhance your scientific research. Supplementary files offer the Author additional possibilities to publish supporting applications, movies, animation sequences, highresolution images, background datasets, sound clips and more. Supplementary files supplied will be published online alongside the electronic version of your article in Elsevier Web products, including ScienceDirect:

<http://www.sciencedirect.com>. In order to ensure that your submitted material is directly usable, please ensure that data is provided in one of our recommended file formats. Authors should submit the material in electronic format together with the final version of the article and supply a concise and descriptive caption for each file. For more detailed instructions please visit our artwork instruction pages at <http://www.elsevier.com/artworkinstructions>.
Proofs When your manuscript is received at the Publisher it is considered to be in its final form. Proofs are not to be regarded as 'drafts'. One set of page proofs in PDF format will be sent by e-mail to the corresponding author, to be checked for typesetting/editing. No changes in, or additions to, the accepted (and subsequently edited) manuscript will be allowed at this

stage. Proofreading is solely your responsibility. A form with queries from the copy editor may accompany your proofs. Please answer all queries and make any corrections or additions required. The Publisher reserves the right to proceed with publication if corrections are not communicated. Return corrections within two working days of receipt of the proofs.

Should there be no corrections, please confirm this. Elsevier will do everything possible to get your article corrected and published as quickly and accurately as possible. In order to do this we need your help. When you receive the (PDF) proof of your article for correction, it is important to ensure that all of your corrections are sent back to us in one communication. Subsequent corrections will not be possible, so please ensure your first sending is complete. Note that this does not mean you have any less time to make your corrections, just that only one set of corrections will be accepted. Proofs are to be e-mailed to the Log-in Department at proofcorrections@elsevier.com.

Offprints The corresponding author, at no cost, will be provided with a PDF file of the article via e-mail or, alternatively, 25 free paper offprints. The PDF file is a watermarked version of the published article and includes a cover sheet with the journal cover image and a disclaimer outlining the terms and conditions of use. Additional offprints and copies of the issue can be ordered at a specially reduced rate using the order form sent to the corresponding author after the manuscript has been accepted. Orders for reprints will incur a 50% surcharge.

Copyright Upon acceptance of an article, authors will be asked to transfer copyright (for more information on copyright see <http://authors.elsevier.com>). This transfer will ensure the widest possible dissemination of information. A letter will be sent to the corresponding author confirming receipt of the manuscript. A form facilitating transfer of copyright will be provided. If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source (s) in the article. Elsevier has preprinted forms for use by authors in these cases: contact Elsevier Ltd., Global Rights Department, The Boulevard, Langford Lane, Oxford, OX5 1GB, UK; phone: (+44) 1865 843830, fax: (+44) 1865 853333, e-mail:

permissions@elsevier.com

Author Enquiries

Authors can keep a track on the progress of their accepted article, and set up e-mail alerts informing them of changes to their manuscript's status, by using the "Track a Paper" feature <http://www.elsevier.com/trackarticle>. Other questions or queries will also be dealt with via the website

<http://authors.elsevier.com>. Contact details for questions arising after acceptance of an article, especially those relating to proofs, are provided when an article is accepted for publication.

Author's Checklist:

- All lines in the text are consecutively numbered.
- Page length does not exceed 6000 words
- Statistical design for the experimental work is clearly described
- Where applicable, figures contain error bars for experimental results
- Format of the text meets the Journal requirements
- Citations in the text and the references list have been crosschecked. All names given in the references have been carefully checked for spellings
- The manuscript has been proof-read by a person other than the authors with technical knowledge to check for writing and English quality.

***Journal of Food Engineering* carries no page charges**

© Copyright 2007 Elsevier | <http://www.elsevier.com>