

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA
DEPARTAMENTO DE QUÍMICA ANALÍTICA



Tese de Doutorado

APLICAÇÃO DE MÉTODOS QUIMIOMÉTRICOS EM ANÁLISE
DE ALIMENTOS

Autor: Marcelo Antônio Morgano

Orientadora: Profa. Dra. Márcia Miguel de Castro Ferreira

Campinas, SP
Dezembro de 2005

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO INSTITUTO DE
QUÍMICA DA UNICAMP**

M821a Morgano, Marcelo Antonio.
Aplicação de métodos quimiométricos em análise de alimentos / Marcelo Antonio Morgano. -- Campinas, SP: [s.n], 2005.

Orientadora: Márcia Miguel de Castro Ferreira.

Tese – Universidade Estadual de Campinas, Instituto de Química.

1. Análise de alimentos. 2. PLS. 3. Café cru.
4. Açúcar. I. Ferreira, Márcia Miguel de Castro.
II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Título em inglês: Application of chemometrics methods in food analysis

Palavras-chaves em inglês: Food analysis, PLS, Raw coffee, Sugar

Área de concentração: Química Analítica

Titulação: Doutor em Ciências

Banca examinadora: Sônia Cláudia do Nascimento de Queiroz; Marcos Flores Ferrão; Adriana Vitorino Rossi; Solange Cadore

Data de defesa: 05/12/2005

**Dedico este trabalho
aos meus pais Neusa e Carmine (*in memoriam*) e
aos meus irmãos Regina e Carlos.**

AGRADECIMENTOS

A Deus pela minha família, pelo meu trabalho, pelos meus amigos e pela presença em minha vida.

Meu agradecimento especial é para minha orientadora Profa. Dra. Márcia Miguel de Castro Ferreira, pela orientação, capacidade, conhecimento científico, amizade, apoio e incentivo em todos os momentos deste trabalho.

Ao Cristiano G. de Faria pela colaboração na execução do trabalho e amizade.

Ao Prof. Dr. Marco F. Ferrão pela colaboração na execução do trabalho, pelas sugestões e pela amizade.

Aos colegas do laboratório Dilza, Marta, Adriana, Tiago, Anderson, Sílvia, Paulo, Ana, Eduardo, Eni, Vanessa, Alessandra, Raquel, Liliana pela amizade e colaboração durante a realização deste trabalho.

A todos os funcionários do Instituto de Química da UNICAMP pela disponibilidade e atendimento.

Ao Instituto de Química da UNICAMP e ao Instituto de Tecnologia de Alimentos pelas facilidades concedidas durante a realização deste trabalho.

A todos que direta ou indiretamente colaboraram na realização deste trabalho.

SÚMULA CURRICULAR

Marcelo Antônio Morgano

Formação Acadêmica

Mestrado em Química Analítica – Instituto de Química da UNICAMP.

“Determinação quantitativa de compostos perhalogenados do tipo $C_xBr_yCl_z$ (onde $x = 1$ ou 2 , $y + z = 4$ ou 6) por cromatografia líquida de alta eficiência”, 1996.

Orientadora: Profa. Dra. Carol. H. Collins.

Bacharelado em Química – UNICAMP, 1990.

Atividade Profissional

Pesquisador Científico

Instituto de Tecnologia de Alimentos – ITAL, Campinas, SP, a partir de 2004.

Assistente de Pesquisa

Instituto de Tecnologia de Alimentos – ITAL, Campinas, SP, a partir de 1992.

Principais Publicações

MORGANO, M.A.; FARIA, C.G.; FERRÃO, M.F.; BRAGAGNOLO, N.; FERREIRA, M.M.C. Determinação de proteína em café cru por espectroscopia NIR e regressão PLS. **Ciência e Tecnologia de Alimentos**, v. 25, n. 1, p. 25-31, 2005.

MORGANO, M.A.; GOMES, P.C.; MANTOVANI, D.M.B, PERRONE, A.A.M; SANTOS, T.F. Níveis de mercúrio total em peixes de água doce de pisciculturas paulistas. **Ciência e Tecnologia de Alimentos**, v. 25, n. 2, p. 250-253, 2005.

QUINTAES, K.D.; AMAYA-FARFAN, J.; TOMAZINI, F.M.; MORGANO, M.A.; MANTOVANI, D.M.B. Migração de minerais de panelas brasileiras de aço inoxidável, ferro fundido e pedra sabão (esteatito) para simulantes de alimentos. **Ciência e Tecnologia de Alimentos**, v. 24, n. 3, p. 397-402, 2004.

OLIVEIRA, E.; ROSA, C.A.; MORGANO, M.A.; SERRA, G.E. Fermentation characteristics as criteria for selection of cachaça yeast. **World Journal of Microbiology and Biotechnology**, v. 20, p. 19-24, 2004.

FERRÃO, M.F.; FURTADO, J.C.; NEUMANN, L.G.; KONZEN, P.H.A.; MORGANO, M.A.; BRAGAGNOLO, N.; FERREIRA, M.M.C. Técnica não destrutiva de análise de tanino em café empregando espectroscopia no infravermelho e algoritmo genético. **Tecno Lógica**, v. 7, n. 1, p. 9-26, 2003

CAIRES, E.F.; FERRARI, R.A.; MORGANO, M.A. Produtividade e qualidade de grãos de soja em função da calagem na superfície em sistema plantio direto. **Bragantia**, v. 62, n. 2, p. 283-290, 2003.

MORI, E.E.M.; BRAGAGNOLO, N.; MORGANO, M.A.; ANJOS, V.D.A.; YOTSUYANAGY, K; FARIA, E.E.V. Brazil coffee growin regions and quality of natural, pulped natural and washed coffees. **Foods & Food Ingredients Journal of Japan**, v. 208, n. 6, p. 416-424, 2003.

MORGANO, M.A.; MORIYA, C.; FERREIRA, M.M.C. Determinação quantitativa do teor de sacarose em açúcar cristal por espectroscopia FT-IR/ATR e Regressão Multivariada, **Brazilian Journal of Food Technology**, v. 6, p.77-83, 2003.

MORGANO, M.A.; PAULUCI, L.F.; MANTOVANI, D.M.B; MORI, E.E.M. Determinação de minerais em café cru. **Ciência e Tecnologia de Alimentos**, v.22, n. 1, p. 19-23, 2002.

QUINTAES, K.D.; AMAYA-FARFAN, J.; MORGANO, M.A.; MANTOVANI, D.M.B. Soapstone (steatite) cookware as a source of minerals. **Food Additives & Contaminants**, v.19, n. 2, p. 134-143, 2002

FARIA, J.M.L.; SILVEIRA, L.A.; MORGANO, M.A., PAVIN, E.J.; FARIA, J.B.L. Erythrocyte sodium-litium countertransport and proliferative diabetic retinopathy. **Investigative Ophthalmology & Visual Science**, v. 41; n. 6, p. 1482-1485, 2000.

FERREIRA, M.M.C.; MORGANO, M.A.; QUEIROZ, S.C.N.; MANTOVANI, D.M.B. Relationship of the Mineral and Fatty Acidy Contents in Processed Turkey Meat. **Food Chemistry**, v. 69, p. 259-265, 2000.

MORGANO, M.A.; QUEIROZ, S.C.N.; FERREIRA, M.M.C. Determinação dos teores de minerais em sucos de frutas por espectrometria de emissão óptica em plasma indutivamente acoplado (ICP OES). **Ciência e Tecnologia de Alimentos**, v.19, n. 3, p. 344-348, 1999.

COLLINS, C.H.; MORGANO, M.A. Quantitative determination of several simple perhalogenated compounds by high-performance liquid chromatography. **Journal Chromatograph A**, v. 846, p. 395-399, 1999.

MORGANO, M.A.; QUEIROZ, S.C.N.; FERREIRA M.M.C. Aplicação da análise exploratória na diferenciação de vegetais. **Brazilian Journal of Food Technology**, v. 2, n. 1/2, p. 73-79, 1999.

Resumos expandidos apresentados em congressos: 21

Resumos apresentados em congressos internacionais: 09

Resumos apresentados em congressos nacionais: 83

RESUMO

APLICAÇÃO DE MÉTODOS QUIMIOMÉTRICOS EM ANÁLISE DE ALIMENTOS

Autor: Marcelo Antônio Morgano

Orientadora: Profa. Dra. Márcia Miguel de Castro Ferreira

Este trabalho teve como objetivo aplicar métodos quimiométricos em diferentes tipos de análise de alimentos. Foi desenvolvido um método para a determinação quantitativa do teor de sacarose em amostras de açúcar cristal, usando a espectroscopia no infravermelho médio com o acessório de reflectância total atenuada (ATR), associada à calibração multivariada baseada no método dos mínimos quadrados parciais (PLS). Outra aplicação foi a determinação quantitativa de algumas das propriedades químicas de amostras de café cru, como teor de umidade, de proteínas e de açúcares totais, usando a espectroscopia no infravermelho próximo (NIR) e a calibração multivariada com PLS. Para a construção dos modelos de calibração foram realizadas as análises químicas para a determinação dos teores de sacarose, de umidade, de proteína e de açúcares totais e em seguida obtidos os espectros no infravermelho. Diferentes pré-tratamentos matemáticos foram usados para a construção dos modelos. Os modelos de regressão desenvolvidos demonstram que é possível determinar com confiabilidade a sacarose em amostras de açúcar e o teor de umidade, de açúcares totais e de proteínas em amostras de café cru com erros menores que 10 %. Os métodos propostos apresentam como principais vantagens serem mais rápidos, mais econômicos, seguros e diminuem os resíduos químicos gerados. Uma terceira aplicação foi investigar os métodos de preparação de amostra, usando extração com ácido clorídrico a frio com agitação e digestão em sistema fechado assistido por microondas, sob pressão, para a determinação quantitativa dos minerais Ca, P, Na, K, Mg, Zn, Fe, Mn e Cu em amostras de sucos de frutas. As concentrações dos minerais foram obtidas empregando-se a espectrometria de emissão óptica com fonte de plasma indutivamente acoplado. A precisão e exatidão dos métodos foram avaliadas. A metodologia de extração com ácido clorídrico foi empregada para a determinação dos teores de minerais em sucos processados de abacaxi, acerola, caju, goiaba, manga, maracujá e uva. Para facilitar a interpretação dos teores dos minerais encontrados, foram utilizados os métodos de análise por componentes principais (PCA) e análise hierárquica de agrupamento (HCA).

ABSTRACT

APPLICATION OF CHEMOMETRICS METHODS IN FOOD ANALYSIS

Author: Marcelo Antônio Morgano

Adviser: Profa. Dra. Márcia Miguel de Castro Ferreira

The aim of the present study was the development of a methodology for quantitative determination of sucrose levels in crystal sugar samples using the Fourier-transform infrared spectroscopy with the accessory of attenuated total reflectance (ATR) associated with multivariate calibration based on the method of the partial least squares (PLS). A second objective was to quantitatively determinate some chemical properties of raw coffee samples such as moisture, protein and total sugar content using the near infrared spectroscopy (NIRS) and the multivariate calibration with PLS. For the calibration models construction, chemical analyses and determination of the sucrose, moisture, protein and total sugars content were performed and the NIR spectrums were obtained. Different mathematical pre-treatments were used for models construction. The regression models obtained demonstrate that it is possible to determine, within 10% error, the sucrose levels in sugar samples and the moisture, total sugar and protein contents in raw coffee samples. The proposed methods present several important advantages, such as being fast, economic, safe and generate small volumes of chemical residues. The third objective was to investigate sample preparation for the quantitative determination of the Ca, P, Na, K, Mg, Zn, Fe, Mn and Cu in fruit juice samples. For this purpose, methods using HCl extraction and digestion in closed microwaves system were tested. Minerals concentrations were obtained by inductively coupled plasma optical emission spectrometry. The precision and accuracy of the methods were evaluated. The methodology involving extraction with HCl was used for the mineral determination in processed juices of pineapple, west Indian cherry, cashew, guava, mango, passion fruit and grape. In order to interpret the levels of minerals determined, the methods of principal components analysis (PCA) and hierarchical cluster analysis (HCA) were used.

SUMÁRIO

LISTA DE ABREVIATURAS	xviii
LISTA DE FIGURAS	xx
LISTA DE TABELAS	xxii
1. JUSTIFICATIVA DO TRABALHO	1
2. OBJETIVOS	5
3. INTRODUÇÃO	6
3.1. AÇÚCAR.....	6
3.1.1. PRODUÇÃO E TIPOS DE AÇÚCAR	6
3.1.2. MERCADO DO AÇÚCAR.....	9
3.2. CAFÉ.....	10
3.2.1. COMPOSIÇÃO DO CAFÉ	11
3.3. SUCO DE FRUTA.....	12
4. ESPECTROSCOPIA NO INFRAVERMELHO	15
4.1. REFLECTÂNCIA TOTAL ATENUADA (ATR).....	16
4.1.1. PRINCÍPIO DA TÉCNICA DE ATR.....	17
4.2. REFLECTÂNCIA DIFUSA.....	19
4.2.1. PRINCÍPIO DA TÉCNICA DE REFLECTÂNCIA DIFUSA.....	21
5. ANÁLISE MULTIVARIADA	25
5.1. ANÁLISE DE COMPONENTES PRINCIPAIS.....	25
5.2. ANÁLISE DE AGRUPAMENTO POR MÉTODO HIERÁRQUICO (HCA).....	26
5.3. MODELOS DE CALIBRAÇÃO.....	28
5.3.1. REGRESSÃO POR MÍNIMOS QUADRADOS PARCIAIS (PLS).	29

5.4. VALIDAÇÃO DO MODELO	31
5.4.1. ESCOLHA DO NÚMERO DE VARIÁVEIS LATENTES.....	31
5.4.2. PREVISÃO DO MODELO DE CALIBRAÇÃO.....	32
6. PROCESSAMENTO DOS DADOS ESPECTRAIS.....	35
6.1. NORMALIZAÇÃO	35
6.2. DADOS CENTRADOS NA MÉDIA.....	35
6.3. DADOS ESCALADOS PELA VARIÂNCIA.....	36
6.4. DADOS AUTOESCALADOS.....	37
6.5. CORREÇÃO MULTIPLICATIVA DE SINAL (MSC).....	37
6.6. ALISAMENTO	38
6.6.1. MÉDIA MÓVEL.....	38
6.6.2. SPLINE.....	40
6.7. KUBELKA – MUNK.....	42
6.8. APLICAÇÃO DE DERIVADA.....	42
7. DETERMINAÇÃO DO TEOR DE SACAROSE EM AÇÚCAR CRISTAL USANDO ESPECTROSCOPIA MID E REGRESSÃO PLS.....	43
7.1. PARTE EXPERIMENTAL.....	43
7.1.1. PADRÃO SACAROSE	43
7.1.2. AMOSTRAS DE AÇÚCAR CRISTAL.....	43
7.1.3. INSTRUMENTAÇÃO (FT-IR) E CONDIÇÕES DE ANÁLISE.....	43
7.1.4. PREPARAÇÃO DA AMOSTRA PELO MÉTODO DE REFERÊNCIA (POLARIMÉTRICO).....	44
7.1.5. ANÁLISE DOS DADOS.....	44
7.1.6. TRATAMENTO DOS DADOS.....	45
7.2. RESULTADOS E DISCUSSÃO.....	46
7.3. CONCLUSÃO.....	52
8. DETERMINAÇÃO DO TEOR DE AÇÚCARES TOTAIS, DE PROTEÍNAS E DE UMIDADE EM AMOSTRAS DE CAFÉ CRU USANDO ESPECTROSCOPIA NIR E O MÉTODO DE REGRESSÃO PLS.....	53

8.1. PARTE EXPERIMENTAL.....	53
8.1.1. AMOSTRAS DE CAFÉ.....	53
8.1.2. ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO E ANÁLISE DOS DADOS.....	53
8.1.3. ESTATÍSTICA.....	54
8.1.4. MÉTODOS DE REFERÊNCIA.....	55
8.1.4.1. MÉTODO PARA A DETERMINAÇÃO DE AÇÚCARES TOTAIS.....	55
8.1.4.2. MÉTODO PARA A DETERMINAÇÃO DE UMIDADE.....	56
8.1.4.3. MÉTODO PARA A DETERMINAÇÃO DE PROTEÍNA....	57
8.2. RESULTADOS OBTIDOS USANDO OS MÉTODOS DE REFERÊNCIA	57
8.3. RESULTADOS E DISCUSSÃO.....	59
8.3.1. AÇÚCARES TOTAIS.....	59
8.3.2. PROTEÍNAS.....	66
8.3.3. UMIDADE.....	73
8.4. CONCLUSÃO.....	78
9. MINERAIS EM SUCOS DE FRUTAS	79
9.1. ESPECTROMETRIA ÓPTICA EM PLASMA INDUTIVAMENTE ACOPLADO.....	79
9.2. PARTE EXPERIMENTAL.....	82
9.2.1. AMOSTRAS	82
9.2.2. MÉTODO PARA PREPARAÇÃO DAS AMOSTRAS.....	82
9.2.3. INSTRUMENTAÇÃO (ICP OES).....	83
9.3. RESULTADOS E DISCUSSÃO.....	85
9.4. CONCLUSÃO.....	92
10. CONCLUSÕES GERAIS.....	94
REFERÊNCIAS BIBLIOGRÁFICAS.....	96

LISTA DE ABREVIATURAS

AR: Açúcares redutores

ART: Açúcares redutores totais

AOAC: Association of official analytical chemists

ATR: do inglês Attenuated total reflection (Reflectância total atenuada)

CV: Coeficiente de variação

DRIFTS: do inglês Diffuse Reflectance Infrared Fourier Transform Spectrometry (Espectroscopia por reflexão difusa no infravermelho com transformada de Fourier)

ER (%): Erro relativo percentual

FT-IR: do inglês Fourier transform infrared (Infravermelho com transformada de Fourier)

HCA: do inglês Hierarchical Cluster Analysis (Análise de agrupamento por método hierárquico)

ICP: do inglês Inductively coupled plasma (Plasma indutivamente acoplado)

ICP OES: do inglês Inductively coupled plasma optical emission spectrometry (Espectrometria de emissão óptica em plasma indutivamente acoplado)

IDR: Ingestão diária recomendada

KM: Kubelka e Munk

log1/R: Logaritmo do inverso da reflectância

L: Janela

LOD: Limite de detecção

LOQ: Limite de quantificação

MSC: do inglês Multiplicative signal correction (Correção multiplicativa de sinal)

MSE: do inglês Mean square error (Erro quadrático médio)

NIR: do inglês Near infrared (Infravermelho próximo)

NIRS: do inglês Near infrared spectroscopy (Espectroscopia no infravermelho próximo)

nº. VL.: Número de variáveis latentes

°C: Graus centígrados

PC: do inglês Principal component (Componente principal)

PCA: do inglês Principal component analysis (Análise de componentes principais)

PCR: do inglês Principal component regression (Regressão por Componentes Principais)

PLS: do inglês Partial least squares (Mínimos quadrados parciais)

PRESS: do inglês Predicted residual error sum of squares (Soma dos quadrados dos erros de previsão)

r: Coeficiente de correlação

R: Reflectância

REV: Variância relativa

RF: Gerador de rádiofreqüência

s: Estimativa de desvio padrão

s²: Estimativa da variância

RMSE: Raiz quadrada do erro quadrático médio

RMSEP: Raiz quadrada do erro médio quadrático de previsão

SEC: Erro padrão de previsão para calibração

SEP: Erro padrão de previsão

SEV: Erro padrão de previsão para validação

VL: Variáveis latentes

LISTA DE FIGURAS

Figura 3.1. Fluxograma simplificado da produção do açúcar.....	8
Figura 4.1. Representação da propagação da radiação infravermelha através do elemento de ATR.....	18
Figura 4.2. Representação da reflexão especular e difusa de uma onda eletromagnética em uma amostra particulada.....	22
Figura 6.1. Esquema do processo de filtragem utilizando média móvel com uma janela de tamanho 3.....	39
Figura 6.2. Ilustração de um ajuste utilizando SPLINE exemplificado para um segmento de um espectro ruidoso.....	41
Figura 6.3. Exemplificação do ajuste por SPLINE de uma gaussiana utilizando 4 funções cúbicas (f_1 , f_2 , f_3 e f_4).....	41
Figura 7.1. Espectros na região do infravermelho médio de vinte soluções de sacarose em água obtidos na região do infravermelho médio com concentrações variando de 3,3420 a 3,4108 mg/100mL.....	46
Figura 7.2. Gráfico de PRESS <i>versus</i> número de componentes principais....	47
Figura 7.3. Curva de regressão da etapa de calibração para o modelo usando derivada segunda.....	49
Figura 7.4. Gráfico de resíduo <i>versus</i> valores previstos para amostras do conjunto de calibração e amostras do conjunto de validação do modelo construído usando derivada segunda.....	49
Figura 8.1. A – Espectros NIR originais; B – Espectros NIR após aplicação da MSC e transformação Kubelka-Munk.....	59
Figura 8.2. A- Gráfico do PRESS de validação interna e de validação externa em função do valor de corte do correlograma (selecionando apenas as variáveis positivamente correlacionadas) dos modelos PLS utilizando 5, 6, 7 e 8 VLs. B- Gráfico do correlograma apresentando o valor de corte utilizado (0,1) e as variáveis selecionadas.....	61
Figura 8.3. Gráfico do PRESS dos conjuntos de validação interna e externa em função do número de variáveis latentes utilizadas nos modelos.....	62

Figura 8.4. Valor experimental (Y_{exp}) <i>versus</i> valor predito (Y_{pred}) (em g de açúcares totais/100g) para o conjunto de calibração do modelo de regressão usando 6 variáveis latentes.....	64
Figura 8.5. Histograma dos resíduos da calibração (%) do modelo PLS com 6 variáveis latentes.....	64
Figura 8.6. Representação gráfica da freqüência do erro percentual para o conjunto de validação externa.	65
Figura 8.7. Representação gráfica do PRESS de validação interna (- -) e de validação externa (—) em função do tamanho da janela (L), para modelos com 3, 4 e 5 variáveis latentes (VL).....	68
Figura 8.8. Representação gráfica do PRESS do conjunto de validação externa em função do número de VL consideradas nos modelos.....	68
Figura 8.9. A - Valores experimentais (Y_{exp}) <i>versus</i> valores preditos (Y_{pred}) do conjunto de calibração para o modelo usando 4 variáveis latentes. B - Histograma dos resíduos de calibração.....	70
Figura 8.10. A - Curva de calibração dos valores experimentais (Y_{exp}) <i>versus</i> valores preditos (Y_{pred}) para o conjunto de validação externa usando 4 variáveis latentes. B - Gráfico dos valores experimentais (Y_{exp}) <i>versus</i> resíduos para o conjunto de validação externa.....	71
Figura 8.11. Comportamento do PRESS x número de variáveis latentes (N_{VL}) x tamanho da janela da média móvel.....	73
Figura 8.12. Correlograma dos dados mostrando o valor de corte utilizado (<i>cut-off</i>) e as variáveis selecionadas (em preto) (número de onda) na determinação da umidade.....	74
Figura 8.13. Valores experimentais (Y_{exp}) <i>versus</i> valores preditos (Y_{pred}) da determinação de umidade para o conjunto de calibração externa usando 3 variáveis latentes.....	76
Figura 8.14. Valor experimental <i>versus</i> valor predito para as amostras do conjunto de validação externa.....	76
Figura 9.1. A - Gráfico de escores. B - Gráfico de pesos.....	90
Figura 9.2. Dendograma obtido pela HCA para as amostras de sucos de frutas.....	91

LISTA DE TABELAS

Tabela 3.1. Composição química aproximada do grão cru de café (valores expressos em g/100g em base seca).....	11
Tabela 7.1. Resultados obtidos para soma dos quadrados dos erros de previsão (PRESS), erro quadrático médio (MSE), coeficiente de correlação (r) e o número de variáveis latentes (VL) para os modelos de calibração.....	48
Tabela 7.2. Resultado da validação dos modelos para a determinação do teor de sacarose, em g 100g ⁻¹	50
Tabela 7.3. Comparação dos resultados obtidos pelo modelo de regressão construídos a partir dos dados espectroscópicos e de polarimetria para as amostras de açúcar cristal, provenientes de usinas de açúcar do Estado de São Paulo.....	51
Tabela 7.4. Comparação dos desvios-padrão e coeficientes de variação em uma amostra de açúcar cristal de usina, obtidos em três dias diferentes de análise, com a realização de dez repetições analíticas, utilizando as técnicas espectroscópica e polarimétrica (Precisão Intermediária).....	52
Tabela 8.1. Identificação das amostras de café cru e valores obtidos pelos métodos de referência para os teores de açúcares totais, de proteínas e de umidade.....	58
Tabela 8.2. Porcentagem de variância capturada para cada variável latente do modelo PLS.....	63
Tabela 8.3. Resultados de predição do teor de açúcares totais para as amostras do conjunto de validação.....	66
Tabela 8.4. Parâmetros de ajuste e erro dos modelos obtidos utilizando 4 variáveis latentes (g/100g) na determinação do teor de proteína.....	69
Tabela 8.5. Porcentagem de variância capturada para cada variável latente do modelo PLS.....	70
Tabela 8.6. Comparação dos resultados de proteína total (g/100g) para as amostras do conjunto de validação externa obtidos pelos métodos de Kjeldhal e NIRS / PLS.....	72

Tabela 8.7. Parâmetros de ajuste e erro dos modelos PLS (g/100g) sem seleção (modelo I) e com seleção de variáveis (modelo II) usados na construção dos modelos para a determinação de umidade.....	74
Tabela 8.8. Porcentagem de variância capturada por cada variável latente do modelo PLS.....	75
Tabela 8.9. Comparação dos resultados do teor de umidade (g/100g) para as amostras do conjunto de validação externa obtidos pelos métodos de referência e NIRS – PLS, aplicando o modelo de regressão de melhor ajuste obtido.....	77
Tabela 9.1. Condições de operações do espectrômetro ICP OES.....	84
Tabela 9.2. Elementos minerais determinados, comprimento de onda, limite de detecção e parâmetros da curva analítica.....	85
Tabela 9.3. Comparação dos teores dos minerais em suco de uva obtidos pelo método de extração com HCl e agitação e pelo método de digestão usando energia de microondas.....	86
Tabela 9.4. Avaliação da exatidão dos métodos usando cálculo de recuperação para uma amostra de suco de uva.....	86
Tabela 9.5. Teores de minerais (mg 100mL ⁻¹) para os dois lotes dos diferentes sucos de frutas estudados (Lotes A e B). Para cada lote os valores correspondem a média de três repetições analíticas (desvio padrão).....	88
Tabela 9.6. Contribuição nutricional média estimada para adultos de um copo de 300 mL de suco após a diluição recomendada pelo fabricante.....	92

1. JUSTIFICATIVA DO TRABALHO

Os principais métodos analíticos para a avaliação da qualidade dos alimentos ou determinação de suas propriedades físico-químicas utilizam as técnicas de cromatografia gasosa, a cromatografia líquida de alta eficiência, a espectrometria de massas, a espectrofotometria na região do ultravioleta e do visível, a espectrometria de emissão óptica com fonte de plasma indutivamente acoplada, a absorção atômica e ainda os métodos clássicos como os volumétricos ou os gravimétricos que ainda continuam sendo muito utilizados nos dias de hoje para as determinações das propriedades dos alimentos.

As análises convencionais para a avaliação da qualidade de alimentos empregam muitos testes físicos e químicos que dependem muito tempo, consomem muitos reagentes químicos, necessitam de equipamentos e mão-de-obra especializada. Para contornar estes problemas, os métodos multivariados de análise associados aos métodos usando a espectroscopia na região do infravermelho têm tido grandes avanços com diversas aplicações nas análises de alimentos e são considerados como alternativos às análises físico-químicas convencionais de controle de qualidade de alimentos e produtos da agricultura. A aplicação analítica da espectroscopia no infravermelho teve um crescimento muito rápido com o aumento da eficácia da instrumentação em combinação com o desenvolvimento de métodos quimiométricos. Os métodos quimiométricos [Morgano *et al.*, 1999; Ferreira *et al.*, 2000] são considerados como uma poderosa ferramenta para processar a vasta quantidade de informações geradas pelas técnicas instrumentais modernas e aliando as técnicas de espectroscopia na região do infravermelho a métodos multivariados de análise, tempo, reagentes e mão-de-obra podem ser poupados.

O desenvolvimento de métodos multivariados tem possibilitado análises quantitativas no infravermelho de diferentes propriedades dos alimentos como, por exemplo, o teor de umidade, de proteínas [Morgano *et al.*, 2005; Konzen *et al.* 2002], de carboidratos, de cinzas, de açúcares [Morgano *et al.*, 2003], de lipídios, de cafeína [Morgano *et al.*, 2001], de taninos [Ferrão *et al.*, 2003], entre outras, em diferentes matrizes de alimentos. Os modelos de regressão por componentes

principais (PCR) e os por mínimos quadrados parciais (PLS) estão entre os mais utilizados em calibração multivariada.

Neste trabalho foram feitas análises quantitativas em diferentes matrizes de alimentos como açúcar cristal, café cru e sucos. Foram desenvolvidos modelos de regressão com aplicação analítica da espectroscopia no infravermelho e do método quimiométrico PLS para a determinação quantitativa de sacarose em amostras de açúcar cristal e determinação quantitativa do teor de açúcares totais, de umidade e de proteínas em amostras de café cru. Também, foi estudada a composição mineral de algumas variedades de sucos de frutas e aplicados os métodos de análise de componentes principais (PCA) e análise hierárquica de agrupamentos (HCA) aos resultados obtidos dos teores dos minerais.

A indústria açucareira tem procurado métodos alternativos para a avaliação da qualidade de açúcar e de sua composição. Os principais parâmetros para avaliar a qualidade do açúcar são a densidade, o tamanho do cristal, a cor e o teor de sacarose. O setor agroindustrial da cana-de-açúcar tem um papel extremamente importante no mercado nacional e internacional; internamente abastecendo o mercado consumidor e externamente como gerador de divisas. Tanto a produção direcionada ao mercado interno como aquela destinada à exportação é avaliada em centros de pagamento de cana-de-açúcar. A implementação de novos métodos alternativos de análise, principalmente para a determinação da concentração de sacarose são primordiais pelas seguintes razões:

- a análise da cana dos fornecedores é geralmente própria, afetando o valor da tonelada de cana nos centros de pagamento pelo teor de sacarose. As medidas têm assim um impacto direto sobre as transações comerciais entre os produtores de cana e as usinas.
- as medidas afetam a qualidade no controle do processo de fabricação (rendimento, confiabilidade....) dos vários tipos de açúcares, desde a sua extração até o produto final.
- a verificação final da qualidade do produto, antes de submetê-lo ao mercado interno ou à exportação, é de grande importância no fechamento dos contratos [Luna *et al.*, 2003].

A qualidade do açúcar cristal está relacionada com o seu grau de umidade, de sacarose, cor da solução do açúcar, entre outros. No caso da sacarose, quanto maior a sua concentração no açúcar, melhor a qualidade do produto. Assim, propôs-se desenvolver uma metodologia de análise para a determinação do teor de sacarose em açúcar cristal, usando a técnica de espectroscopia no infravermelho com transformada de Fourier (FT-IR) e o acessório de ATR associada ao método PLS de regressão com a finalidade de se ter um método alternativo para a avaliação da qualidade do açúcar cristal.

Outra matriz de alimento estudada foi o café cru. A qualidade do café é medida principalmente em função de duas classificações: uma que utiliza o aspecto e a pureza do café e outra, o sabor e o aroma, referindo-se esta última às qualidades sensoriais da bebida. Sabe-se ainda que a qualidade depende na fase pré e pós-colheita da interação entre fatores que garantem a expressão final das características de sabor e aroma como os processos de armazenamento, beneficiamento, torra e moagem.

O sabor característico do café é devido à presença de vários constituintes químicos voláteis e não-voláteis, proteínas, aminoácidos, ácidos graxos, compostos fenólicos e, também, da ação de enzimas sobre alguns desses constituintes, o que gera, como produto de reações, compostos que interferem no sabor e odor da bebida.

A composição química do grão cru de café depende de fatores como a espécie e variedade em questão, das práticas agrícolas, grau de maturação e condições de estocagem dos grãos. Durante o processo de torração do café alguns componentes sofrem degradação alterando assim sua composição final [Carvalho *et al.*, 1997].

Neste trabalho, foram determinados os teores dos seguintes componentes: umidade, proteína e açúcares totais.

O potencial das informações analíticas provenientes de determinações de elementos minerais em alimentos é grande, uma vez que seus níveis de concentração podem ser importantes para sua caracterização e principalmente pelo fato dos minerais serem essenciais na dieta humana. Os constituintes considerados

essenciais à dieta humana podem estar presentes em alimentos, como sucos de frutas em concentrações tais que podem ser classificados como uma fonte rica ou boa de determinados minerais [Miller e Miller, 1996].

Quanto ao desenvolvimento de métodos para a quantificação de elementos minerais em sucos de frutas, observa-se um grande interesse na determinação de diversos elementos minerais. A determinação da composição dos macro e micronutrientes nesta matriz por espectrometria de emissão óptica com plasma indutivamente acoplado (ICP OES) é muito adequada, pois a técnica de ICP OES proporciona altas temperaturas, o que promove a destruição da matéria orgânica, dissociação de moléculas, atomização e ionização para um grande número de espécies diferentes, com poucas interferências químicas, proporcionando a análise multielementar com rapidez, precisão e exatidão [Boss e Fredeen, 1997].

Neste trabalho foi feito um estudo de dois métodos de preparo de amostra, extração com ácido clorídrico a frio e mineralização assistida por microondas, para a quantificação simultânea dos macro e micronutrientes cálcio, cobre, sódio, manganês, magnésio, ferro, potássio, fósforo e zinco em amostras de sucos de frutas de abacaxi, acerola, caju, goiaba, manga, maracujá e uva, por ICP OES. A exatidão dos métodos foi avaliada por ensaio de recuperação dos analitos em uma amostra de suco de uva. Foram usados os métodos quimiométricos, análise por componentes principais (PCA) e análise de agrupamento por método hierárquico (HCA) para a interpretação dos resultados encontrados referentes aos elementos minerais presentes nas diferentes amostras de sucos de frutas.

2. OBJETIVOS

Os principais objetivos deste trabalho foram:

- Desenvolver um método para a determinação quantitativa de sacarose em açúcar cristal usando a espectroscopia no infravermelho médio com transformada de Fourier e como acessório uma cela de reflectância total atenuada (ATR) combinada com o método PLS de regressão.
- Desenvolver métodos de análise para a determinação de algumas das propriedades físico-químicas da qualidade do café, como: o teor de umidade, de açúcares totais e de proteínas, usando a técnica de espectroscopia no infravermelho próximo, associada ao método de regressão multivariada PLS.
- Otimizar os modelos de regressão multivariada, fazendo o uso de pré-tratamento de dados espectroscópicos para se obter os menores valores possíveis de erros na construção dos modelos de regressão para as amostras de café.
- Avaliar dois métodos de preparo de amostra, extração com ácido clorídrico a frio e digestão assistida por microondas, para a quantificação simultânea dos macro e micronutrientes cálcio, cobre, sódio, manganês, magnésio, ferro, potássio, fósforo e zinco em amostras de sucos de frutas de abacaxi, acerola, caju, goiaba, manga, maracujá e uva. Utilizar a técnica de PCA e HCA na identificação das propriedades dos sucos de frutas estudados.

3. INTRODUÇÃO

3.1. AÇÚCAR

Os açúcares são substâncias com gosto adocicado amplamente utilizados em alimentos. Eles fazem parte de um grupo denominado de carboidratos. A sacarose ou açúcar-de-cana é formada por 2 monossacarídeos, D-glicose e D-frutose, originando um dissacarídeo denominado α -D-glicopiranosil- β -D-frutofuranosídeo de fórmula molecular $C_{12}H_{22}O_{11}$.

Os açúcares em geral são utilizados para o consumo humano, na indústria alimentícia, como fonte de energia. Devido ao sabor adocicado é usado como componente de muitos alimentos e bebidas e tem grande aplicação por suas propriedades físicas. Ele ajuda a manter unidos os ingredientes de vários produtos alimentares e farmacêuticos, por causa de sua propriedade de adesividade e viscosidade em solução. É utilizado como conservador, pois em altas concentrações, inibe o crescimento de microorganismos responsável por alterações. Na indústria alimentícia é utilizado principalmente como adoçante, conservante, modificador de estrutura e substituto na fermentação [Vettorazi e MacDonald, 1989].

3.1.1. PRODUÇÃO E TIPOS DE AÇÚCAR

Um fluxograma da obtenção do açúcar-de-cana é mostrado na Figura 3.1.

Após a colheita, a cana é submetida a uma fase de preparação, onde é cortada, desfibrada e esmagada, fazendo-se uso de diversos equipamentos, antes de passar pelas moendas. Nas moendas ocorre a extração da maior parte do caldo bruto e a separação deste do bagaço. O caldo bruto posteriormente sofre um processo de clarificação, visando separar a maior quantidade possível de impurezas dissolvidas e em suspensão, sem afetar a concentração da sacarose.

Após todo o tratamento, o suco é aquecido e posteriormente decantado e filtrado para separar o caldo claro da borra. O caldo claro passa por um processo de

evaporação visando retirar a maior parte de sua água e um cozimento, que transforma o xarope em massa cozida. A massa cozida é resfriada em constante movimento e parte se transforma em cristais. O material obtido sofre um processo de centrifugação, onde ocorrerá a separação do açúcar cristal bruto do melaço. No melaço são realizados outros tratamentos para a sua utilização, como a fabricação de álcoois, adubos, combustível, fermentos e outros. O açúcar cristal bruto é lavado e secado e posteriormente levado ao consumo direto ou antes refinado.

Os tipos de açúcar produzidos podem variar dependendo do tamanho dos cristais de sacarose, do grau de clarificação do açúcar, cor, teor de sólidos, nível de polarização do produto. A polarização (Pol) é a porcentagem em massa de sacarose aparente contida em uma solução açucarada, determinada pelo desvio provocado pela solução no plano de luz polarizada. Diversos tipos de açúcar são produzidos pelas usinas e refinarias:

- Açúcar mascavo - é obtido diretamente da concentração do caldo-de-cana recém-extraído, eliminando-se a etapa de clarificação. Sua cor pode variar de dourado ao marrom escuro em função da variedade e da estação do ano em que a cana é colhida.

- Açúcar demerara - é produzido como matéria-prima para exportação. Pelo processo verifica-se que o açúcar demerara não sofre a sulfitação, processo de purificação e branqueamento. Não é lavado com água dentro da centrífuga, ou seja não são refinados e portanto, o produto final é mais escuro que o açúcar cristal e levemente úmido, dificultando o fácil escoamento dos cristais. Deve conter no mínimo 96,0% de sacarose.

- Açúcar cristal - contém no mínimo 99,3 % de sacarose. A produção deste tipo de açúcar caracteriza-se pelo processo de sulfitação do caldo de cana seguido de calagem e decantação. Na fase final da separação do mel dos cristais na centrífuga, realiza-se uma lavagem com água potável quente que remove praticamente todo o mel residual em volta dos cristais. A granulometria dos cristais pode variar de 0,6 a 0,7 mm. Pode ser classificado em três tipos: açúcar cristal especial, açúcar cristal de alta cor e açúcar cristal de granulometria controlada. Essa

classificação é função principalmente do teor de sacarose, cor e umidade. O teor de açúcares redutores (glicose e frutose) não deve ser maior que 0,2 % para os açúcares cristal especial e cristal de granulometria controlada e deve ser isento no açúcar cristal de alta cor.

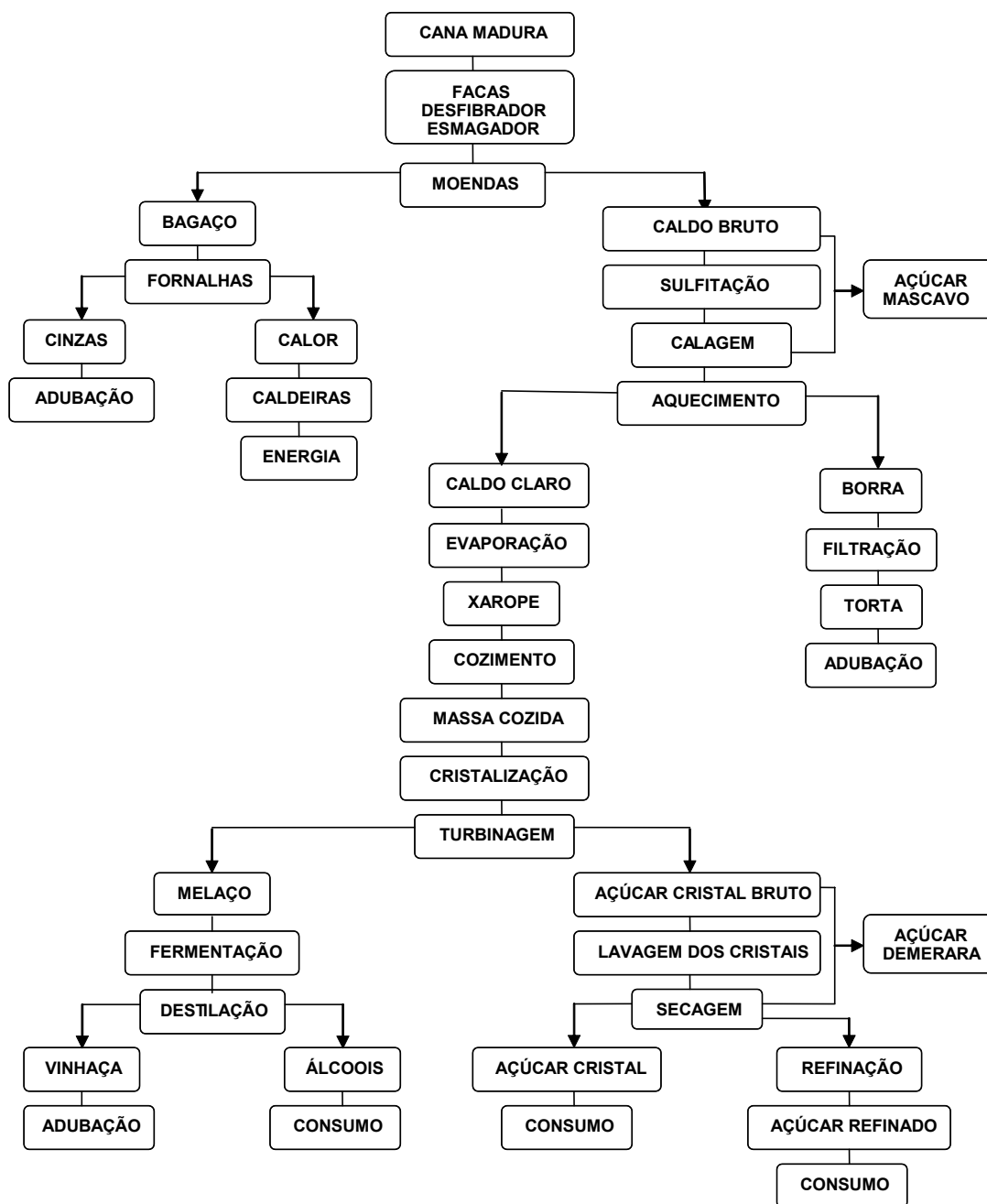


Figura 3.1. Fluxograma simplificado da produção do açúcar.

- Açúcar refinado - contém em sua composição no mínimo 98,5 % de sacarose. É obtido quase totalmente a partir do açúcar cristal. O processo de refinação consiste na dissolução do açúcar cristal e remoção do material insolúvel e dos corantes naturais solúveis por processos físicos. Possui granulometria menor que a do açúcar cristal (0,1 a 0,2 mm). Disponível comercialmente é o mais utilizado no consumo doméstico [Quast, 1986].

- Açúcar orgânico - seu processamento segue os princípios internacionais da agricultura orgânica, que exclui o uso de fertilizantes sintéticos e agrotóxicos, além de utilizar como matéria-prima o açúcar verde, ou seja, foi obtido da cana-de-açúcar que não foi queimada durante a colheita.

Atualmente a qualidade da cana-de-açúcar e seu preço são determinados pela concentração em açúcares redutores totais (ART) em quilograma por tonelada, o qual é função do Pol, do Brix e dos açúcares redutores (AR):

$$ART = (10 \times 0,88 \times 1,0526 \times Pol) + (10 \times 0,88 \times AR) \quad (3.1)$$

onde: $Pol \% \text{ caldo} = \text{leitura sacarimétrica} \times (0,2605 - 0,0009882 \times Brix \% \text{ caldo})$; AR corresponde a porcentagem de açúcares redutores [Consecana-PR, 2005].

Para o pagamento dos fornecedores de cana-de-açúcar, os açúcares redutores são apenas estimados considerando o teor de Brix e o Pol (equação 3.2), devido a amostragem que seria necessário para cada fornecedor de acordo com a área colhida [Procana, 2005].

$$AR_{\text{estimado}} = 9,9408 - (0,1049 \times (Pol / Brix) \times 100) \quad (3.2)$$

3.1.2. MERCADO DO AÇÚCAR

O agronegócio sucroalcooleiro brasileiro movimentava cerca de R\$ 40 bilhões por ano (aproximadamente 2,35 % do PIB nacional). O setor gera cerca de 3,6 milhões de empregos direto e indiretos, envolvendo agricultores [Procana, 2005].

O Brasil é o maior produtor mundial de cana e de açúcar, seguido pela Índia e Austrália. O Estado de São Paulo é responsável por mais de 60 % da produção

nacional de açúcar e álcool e por 70 % das exportações. Em média, 55 % da cana brasileira são convertidas em álcool e 45% em açúcar. Planta-se cana no Centro-Sul (São Paulo, Mato Grosso do Sul, Mato Grosso e Paraná) e no Norte-Nordeste, o que permite dois períodos de safra: de maio a novembro no Centro-Sul e de setembro a março no Norte-Nordeste [Única, 2005].

3.2. CAFÉ

O café provém de uma árvore do gênero *Coffea* e pertence à família das rubiáceas. Esta família engloba mais de quinhentos gêneros e seis mil espécies, a maioria das quais são cafeeiros tropicais e arbustos. Dentro do gênero *Coffea* existem duas espécies principais, assim como duas espécies inferiores. As espécies principais são o *Coffea arabica* variedade *Arábica* e o *Coffea canephora* variedade *Robusta*. As espécies inferiores são o *Coffea liberica* e o *Coffea excelsa* [Rubayiza e Meurens, 2005].

O *Coffea arabica* produz os grãos arábicos, sendo o café de maior qualidade do mundo. As duas variedades mais conhecidas do *Coffea arábica* são a *Típica* e a *Bourbon*, apesar de terem se desenvolvido várias linhagens como a *Caturra* (cultivada no Brasil e na Colômbia), a *Tico* (presente na América Central), a *San Ramon* (uma estirpe anã), a *Mundo Novo* (proveniente do Brasil) e a mais célebre, a *Montanha Azul* Jamaicana [Illy e Viani, 1998].

A variedade robusta é cultivada em grande escala. Provindo de um arbusto forte ou de uma árvore que chega a ultrapassar os dez metros de altura, embora com um sistema radicular pouco profundo, os seus frutos são redondos e necessitam de onze meses para amadurecerem. Esta variedade de cafeeiro é cultivada na África Central e Ocidental e também no Brasil, onde é conhecido pelo termo Conilon. O café Arábica é responsável por cerca de 64% da produção mundial, enquanto o café Robusta aproxima-se dos 35% [Café Damasco, 2005].

Tanto os cafezais de Arábica como os de Robusta permitem a safra após três ou quatro anos de plantio, sendo rentáveis por um período de 20 a 30 anos. Após

este período de vida útil, as plantas terão de ser substituídas. Ambas as espécies necessitam de muito sol e chuva (uma precipitação média anual de um metro e meio é a mais indicada) [Café Damasco, 2005].

3.2.1. COMPOSIÇÃO DO CAFÉ

O sabor característico do café deve-se à presença e aos teores de muitos constituintes químicos, destacando-se entre eles os ácidos, aldeídos, açúcares, proteínas, aminoácidos, ácidos graxos, compostos fenólicos. É muito complexa a composição química do café. Na Tabela 3.1 é apresentada a composição química aproximada do grão cru [Monteiro e Trugo, 2005].

A torração do café provoca mudanças tanto físicas no grão, como modificações na forma, cor e tamanho, quanto químicas devidas principalmente às reações pirolíticas que provocam alterações nos compostos orgânicos do grão cru gerando produtos tais como: caramelos, ácidos voláteis, carbonilas voláteis e sulfetos. O sabor e aroma do café são formados durante a torração dos grãos.

Tabela 3.1. Composição química aproximada do grão cru de café (valores expressos em g/100g em base seca).

Componentes	Café Arábica	Café Robusta
Cafeína	1,2	2,2
Trigonelina	1,0	0,7
Cinzas (41% corresponde a K)	4,2	4,4
Ácidos:		
Clorogênico total	6,5	10,0
Alifáticos	1,0	1,0
Quínico	0,4	0,4
Açúcares:		
Sacarose	8,0	4,0
Redutores	0,1	0,4
Polissacarídeos	44,0	48,0
Lignina	3,0	3,0
Pectina	2,0	2,0
Proteína	11,0	11,0
Aminoácidos livres	0,5	0,8
Lipídeos	16,0	10,0

Durante a torração, os açúcares presentes no café são transformados em produtos caramelizados, responsáveis pela cor do café torrado. Com a elevação da temperatura os açúcares são desidratados, polimerizados e parcialmente degradados a compostos orgânicos voláteis, água e gás carbônico. O amido e as dextrinas sofrem processos semelhantes.

Os ácidos predominantes no café torrado são o clorogênico, o acético e o cítrico e os compostos voláteis são predominantemente formados por aldeídos e cetonas, provenientes da decomposição de carboidratos. Pequenos teores de sulfetos provenientes de proteínas são de grande importância para o aroma do café.

Durante a torração do café ocorre hidrólise nas ligações peptídicas das moléculas protéicas com liberação de carbonilas e aminas. O sulfeto de dimetila é importante para o bom aroma dos cafés. Partículas insolúveis de proteínas podem se ligar a substâncias graxas formando no café coado partículas coloidais responsáveis pela turbidez do café. Em teores aproximados de 13% no grão cru, as proteínas dão contribuição marcante ao sabor do café através dos produtos de sua decomposição.

A fração lipídica do café na presença dos ácidos são hidrolisados para glicerina e ácidos graxos durante a torração. Os ácidos graxos assim formados são importantes por reduzir a tensão superficial dos cafés coados, evitando assim a formação de espumas [Amorin, 1975; Amorin, 1976].

3.3. SUCO DE FRUTA

Suco de fruta é todo o produto obtido por extração de frutas maduras por processos tecnológicos adequados. Eles podem ser classificados quanto à concentração [Portal de Embalagens, 2005]:

- Sucos simples, prontos para beber, com concentração de sólidos solúveis na faixa de 8 a 13 graus Brix (concentração de sólidos solúveis);
- Sucos integrais, com concentração de sólidos variável, em função do tipo de fruta;
- Sucos concentrados, com teor de sólidos solúveis de 55 a 66 graus Brix.

Quanto à forma de preservação :

- Sucos congelados, preservados pela temperatura baixa de estocagem e comercialização (cerca de $-18\text{ }^{\circ}\text{C}$);
- Sucos estáveis à temperatura ambiente, conservados principalmente pela ação de conservantes químicos;
- Sucos refrigerados, com vida útil curta, comercializados na cadeia de refrigeração (4 a $8\text{ }^{\circ}\text{C}$).

Usando métodos de reconhecimento de padrões, Bayer *et al.* [1980] conseguiram determinar a origem geográfica de amostras de sucos de laranja concentrados congelados do Brasil e dos Estados Unidos (Flórida) pela discriminação por cinco elementos (Ba, B, Ga, Mn e Rb).

Informações sobre a composição mineral de sucos de frutas nacionais são em geral escassas e podem dar uma orientação para um bom balanceamento das dietas com relação à ingestão de bebidas. Os elementos minerais reconhecidos como essenciais são comumente divididos entre macroelementos (cálcio, fósforo, potássio, sódio, cloro, magnésio, enxofre) e microelementos (ferro, cobre, cobalto, manganês, zinco, iodo, flúor, molibdênio, selênio, cromo, silício), de acordo com as quantidades maiores ou menores em que são encontrados no organismo humano. A importância de sua inclusão na dieta tem sido amplamente discutida em textos sobre nutrição [Sgarbieri, 1987].

Cinqüenta e três amostras de sucos de uva produzidos no Estado do Rio Grande do Sul foram analisados e os teores dos elementos minerais mais importantes (K, Na, Ca, Mg, Fe, Zn, Cu, Mn, Li e P) foram determinados [Rizzon e Miele, 1995]. No exterior poucos estudos da composição de sucos de frutas tropicais comerciais têm sido efetuados [Câmara *et al.*, 1995].

Simpkins *et al.* [2000] determinaram as diferenças de sucos de laranja concentrados do Brasil e da Austrália, utilizando duas técnicas para quantificação dos elementos minerais, a espectrometria de emissão atômica com plasma induzido

e ICP usando detecção por massas. Eles também verificaram as diferenças regionais de sete variedades de laranja na Austrália.

Oito sucos de laranja integrais comerciais brasileiros foram analisados por Filli *et al.* [2003] para a determinação de Ca, Cu, Fe, K, Mg, Mn, Na, P e Zn por espectrometria de emissão óptica com plasma indutivamente acoplado.

No trabalho de Turra [2005] foram estudados 8 elementos químicos (Br, Ca, Cs, K, La, Na, Rb e Zn) em amostras de sucos de laranja comerciais e de frutas *in natura* pela análise de ativação neutrônica instrumental com o objetivo de caracterizar quimicamente as diferentes marcas de sucos de laranja e, também, conhecer a composição dos diferentes tipos de sucos (integral, concentrado congelado, reconstituído e natural) na tentativa de estabelecer *fingerprints* para a discriminação destes tipos de sucos. O autor concluiu que as marcas comerciais de sucos de laranja apresentam grande variação de composição química que podem ser decorrentes dos diversos processamentos tecnológicos, dificultando o estabelecimento de padrões de discriminação. Relataram, também, que parâmetros como a variedade da planta, local de origem e sistema de cultivo devem interferir na composição dos sucos.

Sucos concentrados brasileiros foram estudados por Soares *et al.* [2004] com respeito a 8 minerais (K, Na, Ca, Mg, Fe, Zn, Cu, Mn) para avaliar a importância nutricional dos sucos de abacaxi, acerola, caju, goiaba, manga, maracujá e pitanga. A contribuição nutricional dos sucos para a dieta de crianças após a diluição recomendada pelo fabricante e considerando a ingestão diária de um copo de 300 mL é elevada quanto a potássio para todos os sucos examinados (170 – 930 % da ingestão diária recomendada, RDA). Os sucos de abacaxi e acerola podem fornecer 6 e 12 % da RDA de ferro, respectivamente. O manganês aparece nos sucos de abacaxi, manga, goiaba e acerola com 38, 14, 8 e 7 % do RDA, respectivamente. Magnésio varia entre um máximo de 9% do RDA em abacaxi e 2% em maracujá e caju. Zinco e cobre variaram entre menos de 1 – 2 % do RDA em sucos de caju e pitanga, a 2 – 6 % nos outros sucos. Para adultos as contribuições são proporcionalmente menores, porém em nada desprezíveis.

4. ESPECTROSCOPIA NO INFRAVERMELHO

Para que ocorra absorção da radiação infravermelha é necessário que ocorra variação no momento de dipolo elétrico da molécula como consequência de seu movimento vibracional ou rotacional. O momento de dipolo é determinado pela magnitude das diferenças de cargas e distância entre dois centros de cargas. O espectro de absorção no infravermelho tem origem quando a radiação eletromagnética incidente tem um componente com valor de número de onda correspondente a uma transição entre dois níveis vibracionais. As vibrações podem ser estiramento ou deformação. As vibrações de estiramento envolvem uma mudança contínua na distância interatômica ao longo do eixo da ligação entre dois átomos. Vibrações de deformação envolvem mudanças no ângulo entre duas ligações [Williams e Norris, 1998].

Na faixa de radiações do infravermelho distinguem-se três regiões: infravermelho próximo (12800 a 4000 cm^{-1}), médio (4000 a 200 cm^{-1}) e distante (200 a 10 cm^{-1}). As absorções moleculares ou sobretons ocorrem com freqüências de 12000 a 9000 cm^{-1} e as bandas de combinação ocorrem de 5500 a 3500 cm^{-1} . O espectro é formado de bandas de absorções para cada substância e pode ser determinado devido a que as ligações entre átomos vibram com freqüências que podem ser descritas por leis físicas, sejam moléculas diatômicas simples ou moléculas poliatômicas [Miller e Miller, 2000].

O avanço da técnica de espectroscopia no infravermelho médio e próximo, como técnica para análise quantitativa, se deve à combinação da operação matemática com transformada de Fourier e da geometria dos espectrômetros que utilizam o interferômetro de Michelson, gerando uma grande eficiência no transporte de radiação até o detector, aumentando a resolução e a reprodutibilidade, além de realizar medidas simultaneamente [Coates, 1998].

As medidas feitas usando espectroscopia no infravermelho próximo (NIR) em sólidos podem ser por reflexão especular, reflectância total atenuada ou reflectância difusa e possuem a vantagem de não serem destrutivas para a amostra. A técnica de reflectância difusa, conhecida como DRIFTS (do inglês, *Diffuse Reflectance*

Infrared Fourier Transform Spectrometry) não requer grande preparo de amostras, apenas moagem para transformá-la em pó [Williams e Norris, 1998].

Na década de 90, a utilização da espectroscopia no infravermelho para a análise qualitativa e quantitativa de alimentos obteve seu maior avanço. Wilson [1990] introduziu as potencialidades da região do infravermelho médio associada às técnicas de reflexão total atenuada (ATR), detecção fotoacústica e reflexão difusa (DRIFTS), na análise e controle de qualidade de alimentos.

Salgo *et al.* [1998] mostraram as primeiras aplicações da espectroscopia NIR para a indústria de açúcar. Tewari *et al.* [2003] determinaram o conteúdo dos açúcares sacarose, glicose e frutose em caldo de açúcar-de-cana usando a espectroscopia NIR e um acessório de fibra ótica associada à regressão por mínimos quadrados parciais (PLS). O método desenvolvido apresentou a vantagem de ser mais rápido que o polarimétrico.

4.1. REFLECTÂNCIA TOTAL ATENUADA (ATR)

A reflexão total atenuada é uma técnica de análise que normalmente requer uma mínima preparação da amostra para análise rápida e fácil para a obtenção de espectros no infravermelho médio. Existem muitos elementos de ATR e uma variedade de acessórios foi desenvolvida para se obter espectros de materiais líquidos, sólidos e viscosos. Van de Voort e Ismail [1991; 1992] demonstraram aplicações do uso da espectroscopia por ATR em análises qualitativa e quantitativa de diferentes produtos, como: leite, carne, óleos, gorduras, manteiga e frutas.

Uma outra aplicação da espectroscopia no infravermelho e de métodos quimiométricos é na quantificação da composição de alimentos com misturas de variedades de produtos diferentes. Briantet *et al.* [1996] descreveram um método para a quantificação do conteúdo da espécie Robusta em misturas das espécies Arábica e Robusta entre 0 e 60 % em café. Usando um modelo de regressão PLS e o acessório ATR obtiveram um erro padrão de previsão de 1,2 % (m/m). Em um trabalho similar [Brianted *et al.*, 1996] foi usado o mesmo princípio para a quantificação de adições de xilose, glicose e frutose em amostras de café.

Reeves III e Zapf [1998] empregaram a espectroscopia por reflexão difusa no infravermelho com transformada de Fourier (DRIFTS) para discriminar vários ingredientes alimentares. Wilson e Tapp [1999] fizeram uma revisão de métodos que empregam a espectroscopia no infravermelho médio na análise de alimentos.

A determinação do conteúdo de cafeína na bebida de café usando a espectroscopia no infravermelho com transformada de Fourier em combinação com a técnica de reflectância total atenuada foi realizada por Singh *et al.* [1998] usando a absorção da cafeína em apenas um comprimento de onda. A sensibilidade da técnica foi de 5 ppm. Em um trabalho similar, Bouhsain *et al.* [1999] desenvolveram um sistema totalmente automatizado usando a FTIR e injeção da amostra em fluxo para a determinação da cafeína com extração *on-line* usando clorofórmio como solvente. O limite de detecção encontrado foi de 9 mg L^{-1} , com desvio padrão relativo de 0,6 % e velocidade de análise de 6 amostras por hora. O método proporcionou uma considerável redução de volume de solvente (250 mL para 30 mL de solvente) e de tempo, quando comparado aos métodos cromatográficos e espectrofotométricos tradicionais.

Como método de análise quantitativa para os açúcares glicose, frutose e maltose a espectroscopia de reflectância total atenuada foi usada por Costa Filho e Poppi [2002]. Outra aplicação de estudos usando ATR/FT-IR foi para os compostos limoneno e terpineno presentes em óleo essencial de citrus [Schulz *et al.*, 2002].

Ferrão e Davanzo [2005] determinaram simultaneamente proteínas e cinzas em amostras de farinha de trigo usando a espectroscopia no infravermelho médio com o acessório de reflectância total atenuada e regressão PLS. Os melhores modelos foram obtidos quando os espectros sofreram pré-tratamentos usando a correção multiplicativa de sinal (MSC).

4.1.1. PRINCÍPIO DA TÉCNICA DE ATR

O fenômeno de reflexão interna é observado sob certas condições. Quando a radiação entra em um prisma construído com um índice de refração alto em relação ao meio externo (cristal de ATR) a radiação será refletida totalmente em seu interior.

Esta reflexão interna gera uma onda evanescente a qual se estende na amostra em contato com o cristal e na região do infravermelho esta onda é atenuada.

Para um índice de refração tal que $n_1 > n_2$, pela lei de Snell o ângulo de refração é imaginário para valores do ângulo de incidência tal que $\theta_i > \text{sen}^{-1}(n_2/n_1)$. O ângulo acima do qual esta onda refratada deixa de ser real é denominado de ângulo crítico.

$$\theta_c = \text{sen}^{-1}(n_2/n_1) \quad (4.1)$$

nesta condição o feixe é totalmente refletido na interface n_1/n_2 .

As propriedades do campo evanescente são ilustradas na Figura 4.1. A propagação da radiação no meio 1, com índice de refração n_1 , sofre reflexão interna na interface com o meio 2, com índice de refração n_2 , quando o ângulo de incidência excede o ângulo crítico. O ângulo de incidência θ_i é medido em relação à normal. O ângulo crítico é definido pela equação 4.1. A propagação do campo evanescente no meio 2 sob essas condições pode ser enumerada [Mirabella, 1985]:

1- A intensidade do campo no meio 2 não é zero, mas um componente normal instantâneo de fluxo de energia no meio 2 cujo termo médio é zero. Assim, não há diminuição da energia e a propagação da radiação no meio 1 é refletida totalmente.

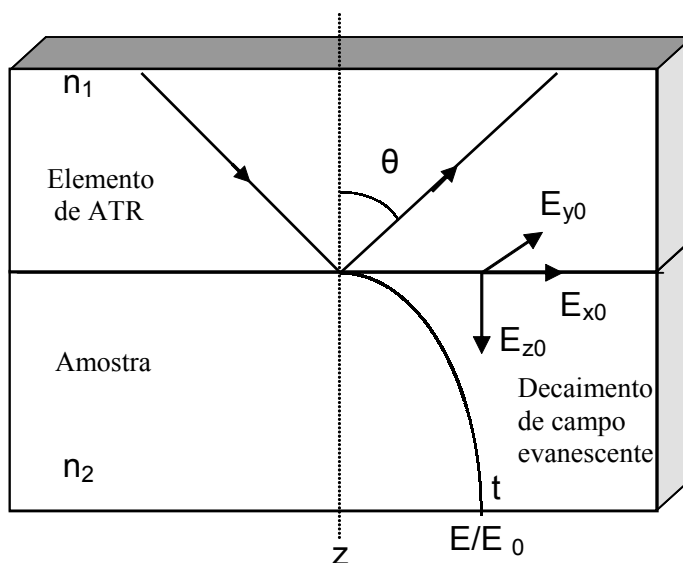


Figura 4.1. Representação da propagação da radiação infravermelha através do elemento de ATR.

2- O campo evanescente no meio 2 é uma onda não transversa e tem componentes em todas orientações espaciais.

3- O campo evanescente é confinado à vizinhança da superfície do meio 2 e diminui em intensidade com a distância no meio 2 normal à superfície.

4- Há um fluxo de energia diferente de zero paralelo à superfície (ao longo do eixo x, resultando num deslocamento das ondas incidente e refletida).

A profundidade de penetração da radiação no segundo meio pode ser quantitativamente definida em termos de distância requerida para a intensidade do campo elétrico decair por um fator de $1/e$ do valor inicial [Mackenzie, 1988] e pode ser calculada usando a seguinte equação:

$$dp = \frac{\lambda}{2\pi n_1 \left(\text{sen}^2 \theta_i - \left(\frac{n_2}{n_1} \right)^2 \right)^{1/2}} \quad (4.2)$$

θ_i é o ângulo de incidência, n_1 é o índice de refração do elemento de ATR e n_2 é o da amostra.

A equação 4.2 mostra que a profundidade de penetração depende do comprimento de onda da radiação no meio com índice de refração mais alto (elemento de ATR).

A radiação que incide na amostra penetra poucos microns na mesma. Assim, qualquer material que esteja em contato com o cristal de ATR pode absorver a radiação atenuando sua intensidade, dando origem ao espectro de infravermelho.

Um dos materiais mais usados no ATR é o ZnSe (insolúvel em água); o seu índice de refração é 2,4 e sua faixa de trabalho é de 20000 até 650 cm^{-1} . Silício é recomendado para amostras altamente espalhadoras, o seu índice de refração é 3,5, é usado na faixa de 9000 até 400 cm^{-1} e é relativamente inerte [Coleman, 1993].

4.2. REFLECTÂNCIA DIFUSA

Os primeiros trabalhos utilizando a espectroscopia no infravermelho próximo na análise de alimentos faziam uso da técnica de reflexão difusa [Tkachuk, 1987]. Esta técnica também foi utilizada na determinação de proteínas em substituição ao

método Kjeldahl [Schuster, 1978]. Davies [1984] apresentou os principais avanços desta técnica, discutindo novos métodos de calibração e regressão multivariada, de seleção de comprimentos de onda para a calibração e as novas formas de utilização da espectroscopia NIRS na indústria de alimentos.

A espectroscopia de reflexão no infravermelho próximo (NIRS) e a espectroscopia por reflexão difusa no infravermelho com transformada de Fourier (DRIFTS) são bem conhecidas na literatura [Birth e Hecht, 1987].

Ikegaya *et al.* [1987] utilizaram a espectroscopia no infravermelho próximo para determinar o conteúdo de cafeína, xantina, teofilina e teobromina em chá verde. Os espectros no infravermelho das amostras de chás foram obtidos na região de 1100 a 2500 nm, utilizaram métodos multivariados (regressão linear múltipla) para a quantificação da cafeína. Foram usadas 30 amostras na calibração e 28 amostras para a previsão dos resultados. O método usando NIR foi comparado ao método químico e não demonstrou diferença significativa ao nível de 99 % de confiança.

Williams e Stevenson [1990] demonstraram diversas aplicações da técnica NIRR na análise de alimentos.

Downey e Spengler [1996] utilizaram NIR e infravermelho médio para a identificação de duas variedades de café. Espectros no infravermelho ($12500 - 500 \text{ cm}^{-1}$) foram obtidos de 56 amostras de pó de café liofilizado usando acessório de reflectância difusa (DRIFT). Para o tratamento dos dados, foi usada a análise de componentes principais (PCA). Um modelo com três componentes foi capaz de descrever 100% de informações dos dados originais e as amostras foram classificadas corretamente entre Arábica e Robusta. Os autores concluíram que a composição em cafeína e ácido clorogênico foram os componentes químicos mais importantes para a diferenciação entre as duas variedades.

Determinação de umidade e gordura na manteiga com espectroscopia na região do infravermelho tem sido feita levando em consideração as vantagens apresentadas por esta técnica de análise como o uso de pequena quantidade de amostra, sem destruí-la e rapidez na análise com bons resultados e boa repetitividade (menor que 3%) [Hermida *et al.*, 2001].

A técnica NIR foi usada como uma ferramenta rápida para quantificar os estágios nutricionais de plantas e monitorar a eficiência de fertilizantes no solo. Determinações com boa precisão foram obtidas de teores de nitrogênio, fósforo, magnésio e ferro em folhas de milho através de medidas de reflectância em diferentes regiões do espectro. Esta metodologia contribuiu para melhor interpretar os estágios fisiológicos das folhas pela análise dos seus espectros e pode ser usada com outras espécies para discriminar a fertilização recomendada [Graeff e Claupein, 2003; Gislum *et al.*, 2004; Schwab *et al.*, 2003].

Para o controle de qualidade de farinhas, as indústrias de moagens têm usado a técnica de espectroscopia de infravermelho para a determinação de proteína bruta, além de outros parâmetros, como a determinação do valor energético e do conteúdo de matéria seca [Miralbés, 2004]. Ferrão *et al.* [2004] desenvolveram uma metodologia para a determinação simultânea dos teores de cinza e de proteína em amostras de farinha de trigo utilizando as técnicas de reflexão no infravermelho próximo e por reflexão difusa no infravermelho médio com transformada de Fourier empregando o método de regressão PLS para a determinação simultânea destas duas propriedades. Foram obtidos modelos de regressão bem ajustados com erros padrão de calibração e erro padrão de validação inferiores a 0,3 %.

4.2.1. PRINCÍPIO DA TÉCNICA DE REFLECTÂNCIA DIFUSA

A reflexão difusa ocorre em superfícies não totalmente planas, podendo o substrato ser contínuo ou fragmentado (na forma de pó). Neste processo de reflexão, o feixe incidente penetra a superfície da amostra interagindo com a matriz, retornando à superfície da mesma, após absorção parcial e múltiplos espalhamentos, conforme ilustrado pela Figura 4.2. Na reflexão difusa, a radiação incidente entra em contato diversas vezes com as partículas da amostra sendo conseqüentemente atenuada. Assim, a radiação que sofre a reflexão difusa fornece informações qualitativas e quantitativas, sobre a natureza química da amostra.

Uma característica importante desta técnica é a possibilidade de ocorrência de reflexão especular (reflexão que ocorre na superfície do substrato com igual

ângulo de incidência), que é função das características da matriz com a qual o feixe incidente interage, sendo este fenômeno mais intenso nas regiões do espectro onde a absorção da radiação é maior e, em alguns casos, pode ocasionar o deslocamento de bandas.

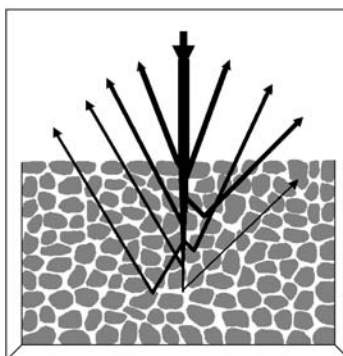


Figura 4.2. Representação da reflexão especular e difusa de uma onda eletromagnética em uma amostra particulada.

Um efeito que merece destaque sobre o espectro devido à reflexão especular é o tamanho das partículas. No trabalho descrito por Devaux *et al.* [1995], sugere-se fazer transformações no espectro como a primeira ou a segunda derivada ou correção do espalhamento multiplicativo (MSC) para reduzir esses efeitos.

Algumas propriedades da amostra, como a geometria ótica, a absorção pela matriz, o índice de refração, a absorvidade, o empacotamento, a morfologia (tamanho da partícula) e a concentração do analito, afetam o formato da banda, a posição e a intensidade nos espectros de reflexão difusa tanto no infravermelho próximo quanto no infravermelho médio. A principal causa de desvio da linearidade da intensidade da banda da reflexão difusa *versus* a concentração é a radiação refletida da superfície das partículas da amostra, chamada reflexão especular (ou reflexão de Fresnel) [Olinger e Griffiths, 1993].

Na ausência de absorção, com espessura infinita e múltiplos espalhamentos, um máximo de radiação retorna à superfície. Quando ocorre absorção, a radiação é transmitida e espalhada nas interfaces no interior da amostra, antes de retornar a superfície.

O tamanho, a forma das partículas, o vazio entre elas e a compactação afetam a quantidade (concentração) do material através do qual a radiação é transmitida. Alguns desses fatores também afetam o caminho entre as interfaces individuais. E a morfologia da superfície da amostra é uma consideração importante na magnitude do componente especular [Wetzel, 1983].

A atenuação da parte da reflexão difusa pela absorção no meio é dada pela lei Lambert-Bourguer, conforme a equação 4.3:

$$I = I_0 \cdot e^{-\varepsilon \cdot \delta} \quad (4.3)$$

onde I representa a intensidade da radiação, I_0 corresponde a intensidade da radiação incidente, ε é o coeficiente de extinção molar da amostra e δ é a espessura média da camada.

A profundidade de penetração da radiação, δ , em um meio absorvente depende de vários fatores. Para uma dada matriz e concentração do analito, o tamanho da partícula do analito e da matriz diluente governam a profundidade de penetração efetiva [Tevrucht e Griffiths, 1991].

Kubelka e Munk desenvolveram uma teoria considerando que uma camada é composta de partículas absorventes e espalhativas uniformemente e aleatoriamente distribuídas, cuja dimensão das partículas é muito menor que a espessura da camada [Birth e Hecht, 1987]. Eles consideram somente a direção da radiação incidente e refletida perpendicular à superfície da camada, assumindo que a radiação é monocromática e que a espessura é infinita (o aumento da espessura não resulta em diferença na reflexão). A equação 4.4 é conhecida como equação de Kubelka-Munk:

$$F(R_\infty) = \frac{(1 - R_\infty)^2}{2R_\infty} = \frac{k}{s} \quad (4.4)$$

onde R_∞ é a reflexão da amostra para profundidade infinita, k é o coeficiente de absorção da amostra (que é proporcional à concentração) e s é o coeficiente de espalhamento da amostra.

O coeficiente de espalhamento determina a extensão da interação da luz incidente com a amostra antes da radiação retornar à superfície. Para uma certa extensão, s controla a profundidade na qual a luz penetra na amostra. Quanto menor o valor de s , maior é o valor $F(R_{\infty})$.

Entretanto, existe um limite para este comportamento, sendo somente válido para as partículas que efetivamente espalham a radiação incidente [Korte e Otto, 1988].

Como consequência, o efeito do tamanho da partícula desloca a linha de base, comportamento que se torna muito pronunciado em comprimentos de onda de grande absorção pela amostra. Por exemplo, em duas amostras com mesma composição, porém com diferente granulometria, há maior reflexão das partículas menores (mais finas), e os efeitos da reflexão especular são minimizados em uma superfície composta por partículas de menor tamanho.

A função de Kubelka-Munk pode relacionar os espectros de reflexão difusa com a concentração da amostra, transformando o espectro de reflexão em formato semelhante a um espectro de absorvância. Por esta razão, a função de Kubelka-Munk é conhecida como a Lei de Beer da espectroscopia de reflexão.

A correta quantificação de analitos em amostras em pó é dependente do efetivo poder de penetração, sendo grande o suficiente para fornecer um espectro representativo da amostra por inteiro. A equação de Kubelka-Munk é mais simples e é geralmente a forma mais usada para definir uma afinidade linear entre intensidade da banda e a concentração da amostra na espectroscopia por reflexão difusa [Olinger, Griffiths, 1993].

O rápido aumento na utilização da espectroscopia no infravermelho por reflexão difusa surgiu devido à facilidade em se obter espectros de pós, sólidos e espécies adsorvidas em sólidos e, também, devido ao surgimento de vários tipos de acessórios comerciais para se obter os espectros DRIFTS.

5. ANÁLISE MULTIVARIADA

5.1. ANÁLISE DE COMPONENTES PRINCIPAIS

A base fundamental da maioria dos métodos modernos para tratamento de dados multivariados é a análise por componentes principais (PCA – *Principal Components Analysis*) [Beebe *et al.*, 1998; Sharaf *et al.*, 1986], que consiste numa transformação da matriz de dados com o objetivo de representar as variações presentes em muitas variáveis, através de um número menor de “fatores”. Constrói-se um novo sistema de eixos (denominados rotineiramente de fatores, componentes principais, variáveis latentes ou ainda autovetores) para representar as amostras, no qual a natureza multivariada dos dados pode ser visualizada em poucas dimensões. Estas novas variáveis são obtidas em ordem decrescente de quantidade de informação estatística que descrevem, ou seja, a primeira componente principal aponta a direção de maior variação dos dados, a segunda, que é ortogonal à primeira, aponta outra direção que descreve a maior variação restante dos dados e assim por diante.

Usando a notação matricial, as componentes principais são obtidas por meio de transformações lineares conforme a equação:

$$\mathbf{X} \mathbf{P} = \mathbf{T} \quad (5.1)$$

em que \mathbf{X} é a matriz original dos dados, \mathbf{T} é a matriz de escores que contém as coordenadas das amostras no novo sistema de eixos e \mathbf{P} é a matriz dos pesos (*loadings*) onde os elementos de cada coluna correspondem aos coeficientes das combinações lineares das variáveis originais.

O método PCA está baseado na correlação entre variáveis e, na realidade, agrupa aquelas que estão altamente correlacionadas. As colunas da matriz de pesos correspondem aos autovetores da matriz de correlação (ou variância e co-variância dependendo do pré-processamento utilizado), enquanto os autovalores representam uma fração da variância total dos dados. Assim, cada componente principal descreve uma porcentagem da variância total dos dados.

Uma das vantagens destas transformações é que ruídos experimentais podem ser eliminados, pois estes não estão correlacionados com as informações contidas na matriz de dados original. Outra vantagem é que podemos escolher as variáveis originais mais importantes do ponto de vista estatístico.

Os dados experimentais originais podem não ter uma distribuição adequada para a análise, dificultando a extração de informações úteis e interpretação dos mesmos. Nestes casos, um pré-processamento nos dados originais pode ser de grande valia (Capítulo 6). Medidas em diferentes unidades e variáveis com diferentes variâncias são algumas das razões que levam a estes problemas. Os métodos de pré-processamento mais utilizados consistem basicamente em centrar na média ou autoescalar os dados. No primeiro caso, calcula-se a média dos valores experimentais para cada variável e subtrai-se cada valor experimental do respectivo valor médio. Autoescalar significa centrar os dados na média e dividi-los pelo respectivo desvio-padrão, sendo um para cada variável.

5.2. ANÁLISE DE AGRUPAMENTO POR MÉTODO HIERÁRQUICO (HCA)

Esta é uma ferramenta excelente para análise preliminar dos dados [Sharaf *et al.*, 1986; Beebe *et al.*, 1998] sendo útil para determinar a semelhança entre objetos e identificar amostras anômalas. O método relaciona as amostras de forma que as mais semelhantes são agrupadas entre si com relação às variáveis usadas no processo de agrupamento. Cada amostra é representada como um ponto no espaço de m variáveis e, portanto, a questão fundamental reside em transformar a informação de um espaço m -dimensional para um espaço bidimensional. O parâmetro usado neste processo é a medida de distância entre as amostras. Estas distâncias, quando visualizadas através de diagramas conhecidos por dendrogramas, indicam as amostras mais semelhantes entre si, ou seja, quanto menor a distância entre os pontos, maior será a semelhança entre as mesmas. Para proceder à análise de agrupamentos, é necessário estabelecer o conceito de distância entre duas amostras a e b (d_{ab}) em um espaço de m variáveis. Há várias

medidas usuais e sendo a mais utilizada a distância euclideana (d_{ab}), que pode ser determinada pela expressão matemática:

$$d_{ab} = \left(\sum_{j=1}^m (x_{aj} - x_{bj})^2 \right)^{1/2} \quad (5.2)$$

em que x_{aj} é a j -ésima coordenada da amostra a .

No método de análise por agrupamento hierárquico as distâncias entre todas as amostras são calculadas e comparadas. Como as distâncias reais entre as amostras mudam com o tipo e tamanho do conjunto de dados, é necessário transformá-las numa escala de similaridade que é calculada segundo a expressão:

$$s_{ab} = 1 - \{ d_{ab}/d_{max} \} \quad (5.3)$$

onde: s_{ab} é a similaridade entre as amostras a e b , d_{ab} é a distância entre as mesmas e d_{max} é a maior distância entre todas as amostras do conjunto de dados. Nestes diagramas a escala varia de um (amostras idênticas) a zero (amostras sem similaridade).

Após o cálculo das similaridades, as duas amostras mais próximas são conectadas formando um agrupamento. Este processo é repetido até que todas as amostras sejam conectadas formando um único grupo. Uma vez que as amostras são conectadas pela proximidade entre elas, é necessário definir a distância entre uma amostra e um grupo ou entre grupos de amostras e há várias técnicas para tal. As mais usuais são: conexão pelo vizinho mais próximo (*single linkage* ou *nearest neighbour*); conexão pelo vizinho mais distante (*complete link* ou *farthest neighbour*); conexão pela distância média (*average link*). Quando a distância entre dois agrupamentos for obtida pela distância média, pode-se utilizar a média aritmética ou ponderada. No primeiro caso, cada agrupamento tem o mesmo peso e no último, o coeficiente de ponderação é correspondente ao número de elementos nos respectivos *clusters* (agrupamentos) e, neste caso, o método é denominado *incremental*.

5.3. MODELOS DE CALIBRAÇÃO

Construir um modelo de calibração significa desenvolver expressões matemáticas que, de alguma forma, possam descrever o comportamento de uma variável de interesse (variável dependente), designada por y . Pode-se citar como exemplo, a concentração de um analito de interesse.

Para a construção do modelo de regressão é necessário se ter alguma informação sobre a variável dependente. Estas informações são obtidas pelas variáveis independentes. Os sinais de reflectância ou absorvância em diferentes comprimentos de onda de um espectro no infravermelho de uma amostra são exemplos de variáveis independentes e podem estar relacionados com a concentração da amostra. Estes sinais podem ser organizados na forma de uma matriz X . Os modelos ainda geram constantes denominadas de coeficientes que controlam o comportamento do modelo e são estimados a partir dos dados (variáveis). Uma representação do modelo matemático:

$$y = X b \quad (5.4)$$

onde: y = vetor contendo os valores das variáveis dependentes; b = vetor contendo os coeficientes de regressão e X matriz de dados contendo as absorvâncias.

O objetivo da construção do modelo matemático é encontrar relações entre as variáveis dependentes e as independentes, por alguma função de transferência.

Devido ao grande avanço da instrumentação e da informática, tem sido favorecido o uso de modelos de calibração multivariada nas determinações de propriedades físico-químicas de alimentos.

Os métodos multivariados de análises quantitativa empregados com maior frequência são os de calibração [Martens e Naes, 1996; Thomas, 1994] sendo os mais conhecidos a regressão linear múltipla (MLR), regressão por componentes principais (PCR) e regressão por mínimos quadrados parciais (PLS) [Thomas, 1990; Geladi e Kowalski, 1986; Beebe e Kowalski, 1987]. Todos estes métodos são lineares, porque os modelos resultantes descrevem uma relação linear entre as respostas e as concentrações dos componentes das amostras. Outros métodos,

como PARAFAC e redes neurais, também têm sido aplicados nos últimos anos [Sena *et al.*, 2005; Vale e Zambiasi, 2000; Gutteres *et al.*, 2004].

PLS é um exemplo de método de calibração indireto, pois não exige que os espectros individuais dos constituintes de interesse, nem dos interferentes, sejam conhecidos previamente. Entretanto, requer análise de uma série de amostras de calibração (conjunto de calibração), cuja composição seja conhecida [Coates, 1998]. Essa calibração deve levar em conta todos os fenômenos físicos e químicos que possam afetar os espectros das amostras cujas concentrações deverão ser previstas (conjunto de validação).

Duas etapas são importantes na calibração multivariada: a construção do modelo (calibração) e a verificação de sua validade, ou seja, se o modelo construído é ou não realístico (previsão ou validação). Na etapa de calibração, são obtidas uma matriz com as respostas instrumentais (variáveis independentes – \mathbf{X}) e uma outra matriz com os resultados das concentrações (variáveis dependentes – \mathbf{y}), normalmente obtidos pelos métodos convencionais de análise. Os dados assim organizados constituem o conjunto de calibração. Diante disto, cria-se um modelo matemático adequado que melhor possa reproduzir \mathbf{y} a partir dos dados de \mathbf{X} . O modelo construído será, então, utilizado na fase de previsão para estimar as concentrações dos constituintes de novas amostras, a partir de suas respostas instrumentais (variáveis independentes). Os dados utilizados nesta fase formam o conjunto de validação.

As linhas da matriz \mathbf{X} correspondem às amostras e as colunas são formadas pelos valores das intensidades dos espectros em cada comprimento de onda. Na matriz \mathbf{y} , as linhas correspondem às mesmas amostras da matriz \mathbf{X} e as colunas são formadas pelos valores das concentrações destas amostras.

5.3.1. REGRESSÃO POR MÍNIMOS QUADRADOS PARCIAIS (PLS)

Esta técnica foi desenvolvida na década de 70 usando dados espectroscópicos na região do infravermelho próximo, onde é difícil designar bandas para componentes particulares. No modelo PLS, as informações espectrais (matriz

X) e as informações das concentrações (matriz **y**) são usadas ao mesmo tempo correlacionando-as a fim de se obter uma relação linear na fase de calibração.

A regressão dos mínimos quadrados parciais, para construção do modelo, retira informações do conjunto de dados da matriz espectral (matriz **X**) para correlacioná-las com as informações retiradas do conjunto de dados de referência (matriz **y**). Por meio de combinações lineares dos dados espectroscópicos (da matriz **X**) e dos dados de referências (matriz **y**), se obtém o número de variáveis latentes necessários para correlacionar os espectros e as concentrações. É usado para construção do modelo de calibração um número de variáveis latentes que proporcione o menor erro possível de previsão, ou seja, que as diferenças entre os valores de referência e os valores previstos sejam as menores. Pelo método dos mínimos quadrados parciais, as amostras anômalas são identificadas pelos valores de resíduos de Student, quando a amostra está fora do limite de confiança de 95%, e pelo valor limite de “*leverage*” (três vezes o número de variáveis latentes dividido pelo número de amostras).

Na modelagem por mínimos quadrados parciais, tanto a matriz das variáveis independentes (**X**), como o vetor das variáveis dependentes (**y**) são representados pelos escores e pesos conforme a equação 5.5:

$$\mathbf{X} = \mathbf{T}\mathbf{P} + \mathbf{E} \quad (5.5)$$

As informações são incorporadas de modo que cada componente principal (CP) do modelo seja modificada para que a covariância de **T** seja maximizada. Nesta etapa, a CP recebe o nome de variável latente (VL). Quando a matriz **X** é decomposta, **T** contém os escores, **P** contém os pesos, **E** é a matriz de erros da decomposição de **X**.

A decomposição pode ser feita utilizando o algoritmo NIPALS. Os valores de escores são relacionados com os *loadings* para cada componente de cada vez pela equação:

$$\mathbf{q} = \mathbf{y} \cdot (\mathbf{t}^T \cdot \mathbf{t}) \cdot \mathbf{y}^T \quad (5.6)$$

O vetor de regressão é calculado pela equação:

$$\mathbf{b} = \mathbf{W} (\mathbf{P} \mathbf{W}^T)^{-1} \cdot \mathbf{q}^T \quad (5.7)$$

onde **W** é a matriz de pesos do método PLS [Martens e Naes, 1996].

5.4. VALIDAÇÃO DO MODELO

5.4.1. ESCOLHA DO NÚMERO DE VARIÁVEIS LATENTES

A seleção do número de variáveis latentes a serem incluídas no modelo de regressão criado é um ponto muito importante. Deve-se sempre selecionar um número de variáveis latentes com informações suficientes que permita modelar o sistema de forma a não ser superajustado e também não deve modelar ruídos. Um modelo com maior número de variáveis latentes é mais ajustado, porém pode estar modelando ruídos. Já um número muito baixo de variáveis latentes pode deixar informações importantes de fora. O número ideal de variáveis latentes é aquele que permite a criação de um modelo com boa capacidade de previsão para amostras externas. Um dos métodos mais usados para a escolha do número de variáveis latentes é a validação cruzada [Martens e Naes, 1996]. Neste procedimento, deixa-se uma amostra de fora, de uma série de n amostras, $n-1$ amostras são usadas na calibração e é feita a previsão da concentração da amostra que não foi usada na calibração. O processo é repetido até que todas as amostras tenham sido excluídas uma vez. Compara-se a concentração de cada amostra prevista pelo modelo com o valor de referência e determina-se o erro. A soma dos quadrados dos erros de previsão (PRESS – *predicted residual error sum of squares*) é medida do mesmo modo para cada variável latente [Lindgren *et al.*, 1994] e é definida como:

$$\text{PRESS} = \sum (y_i - \hat{y}_i)^2 \quad (5.8)$$

onde y_i é o valor conhecido e \hat{y}_i é o valor estimado.

Uma forma alternativa e complementar para se determinar o número de variáveis latentes é pela inspeção dos vetores pesos (*loadings*) espectrais. Os vetores pesos maiores que zero estão positivamente correlacionados com a concentração e os menores que zero estão negativamente correlacionados. Variáveis latentes significativas têm picos em comum com os espectros e estes

podem ser positivamente ou negativamente correlacionados. Variáveis latentes com alta correlação em área onde os espectros não possuem picos podem estar modelando ruídos [Niemeyer *et al.*, 1992].

5.4.2. PREVISÃO DO MODELO DE CALIBRAÇÃO

Após a construção do modelo de regressão, pode-se utilizar uma série de amostras externas para verificar a capacidade de previsão do modelo de regressão. Uma forma de validar o modelo é usando uma série de amostras de referência (validação externa), de composição diferente das amostras que faziam parte com conjunto de calibração. Para verificar a capacidade de estimativa de previsão do modelo é recomendado utilizar de uma série de amostras representativa para a população [Martens e Naes, 1996].

Foram utilizados os seguintes erros para avaliar a capacidade de previsão dos modelos desenvolvidos:

- erro quadrático médio (MSE), mais conhecido como erro padrão de predição (SEP)

$$\text{MSE} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (5.9)$$

- raiz quadrada do MSE (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (5.10)$$

onde n representa o número de amostras do conjunto de calibração, y_i é o valor de referência e \hat{y}_i é o valor previsto pelo modelo para a i ésima amostra.

Como estes erros podem ser medidos tanto para o conjunto de calibração quanto para o de previsão é comum adicionar no final da sigla destes erros a letra C indicando ser estes relativos à calibração C (tipo MSEC ou RMSEC) ou P quando forem relativos à previsão (tipo MSEP ou RMSEP). Estes erros podem dar uma boa medida do desempenho do modelo.

Em contrapartida aos erros absolutos (PRESS, MSE, RMSE) têm-se as medidas de erros relativos que também são interessantes de ser estimados, pois fornecem um parâmetro de comparação entre os erros absolutos e a variância intrínseca dos dados. Um destes erros relativos utilizados foi o de variância relativa explicada (REV).

- variância relativa (REV)

$$\text{REV} = 1 - \text{MSE}/s^2 \quad (5.11)$$

A estimativa da variância total dos dados (s^2) para n objetos, utilizada na expressão acima, é calculada segundo a expressão:

$$s^2 = \sum_{i=1}^n (y_i - \bar{y})^2 / (n-1) \quad (5.12)$$

onde: \bar{y} é o valor médio.

Quanto mais próximo o valor do REV for da unidade, melhor e mais útil é o modelo de regressão criado, sendo que o oposto pode-se dizer para os valores de REV próximos de zero.

- Erro relativo percentual (ER (%)) entre o método de referência e o método desenvolvido (NIR-PLS) foi usado para verificar a capacidade de previsão do modelo. Ele expressa a exatidão que pode ser esperada para amostras de composição desconhecida [Haaland, 1990].

$$\text{ER} (\%) = \frac{\sum (y_i - \hat{y}_i)}{y_i} \times 100 \quad (5.13)$$

- Coeficiente de correlação linear (r) exibido pelo gráfico do valor experimental (obtido pelos métodos de referência) em função do valor predito (calculado pelo modelo de regressão) é outro importante parâmetro medido e de fácil interpretação. Quanto mais próximo de 1 melhor o ajuste (caso o gráfico seja feito com as amostras do conjunto de calibração) ou a habilidade de previsão do modelo de

regressão (caso o gráfico seja feito com as amostras do conjunto de validação externa).

$$r = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})(y_i - \hat{y})}{\left[\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 (y_i - \hat{y})^2 \right]^{1/2}} \quad (5.14)$$

Valores de repetitividade e de reprodutibilidade também são parâmetros importantes, principalmente quando diferentes métodos de calibração são comparados. A repetitividade é a tendência de um sinal dar a mesma resposta quando usado várias vezes no mesmo experimento e a reprodutibilidade é a tendência de um sinal dar a mesma resposta quando usado várias vezes nas mesmas condições experimentais, mas não no mesmo experimento [Bellonmaurel *et al.*, 1995; Currie e Svehla, 1994]. O termo precisão intermediária é mais adequado para este parâmetro.

A repetitividade do método foi calculada pela estimativa do desvio padrão (s) para cada amostra determinada em triplicata.

$$s = \sqrt{(\hat{y}_i - \bar{y})^2 / (n - 1)} \quad (5.15)$$

onde \bar{y} é o valor médio das concentrações previstas e n o número de medidas independentes para cada amostra.

A precisão intermediária foi estimada pelo coeficiente de variação de n repetições analíticas.

$$CV = \frac{s}{y} \times 100 \quad (5.16)$$

onde s é a estimativa de desvio-padrão e \bar{y} o valor médio de n repetições analíticas.

A observação visual do gráfico das concentrações conhecidas em função das concentrações calculadas é uma forma simples para a avaliação do desempenho do modelo de regressão. Discrepâncias entre as correlações obtidas para calibração e para a validação sugerem que o melhoramento na calibração pode estar sendo obtido com o uso de ruídos nos dados espectrais.

6. PROCESSAMENTO DOS DADOS ESPECTRAIS

O tratamento dos dados é aplicado antes da fase de calibração para melhorar a eficiência do método, produzindo um modelo de regressão com a melhor capacidade de previsão de novas amostras. Diversos pré-tratamentos da resposta (sinal medido) têm sido relatados na literatura, como por exemplo: normalização, correção da linha de base, cálculo da derivada, alisamento, correção multiplicativa de sinal [Martens e Naes, 1996]. Determinações quantitativas a partir de espectros obtidos no infravermelho podem apresentar problemas de deformação da linha de base, variações no caminho óptico, ruídos, entre outros.

6.1. NORMALIZAÇÃO

A normalização pela soma das intensidades é realizada dividindo-se cada intensidade espectral pela raiz quadrada da soma dos quadrados de todas as intensidades do espectro correspondente, calculada como:

$$X_{ij} = x_{ij} / (\sum x_{ij}^2)^{1/2} \quad (6.1)$$

Este processo é o mesmo que multiplicar cada elemento no vetor linha por uma constante b:

$$b = 1 / (\sum x_{ij}^2)^{1/2} \quad (6.2)$$

6.2. DADOS CENTRADOS NA MÉDIA

Consiste em fazer com que para cada variável seus valores tenham média zero. Isto é feito, obtendo-se para cada coluna o valor médio e, em seguida, subtrai-se este valor de cada variável dessa mesma coluna. Desta forma, ocorre a mudança do sistema de coordenadas para o centro dos dados. Esta operação geralmente reduz a dimensão do modelo [Thomas, 1994]. A equação 6.3 é utilizada para centrar os dados na média.

$$x_{(i,k)} = x_{(i,k)} - \bar{x}_i \quad (6.3)$$

onde \bar{x}_i é o valor médio correspondente a i -ésima coluna.

6.3. DADOS ESCALADOS PELA VARIÂNCIA

Nesse tipo de pré-processamento os dados são ponderados pelos recíprocos dos desvios-padrão de suas respectivas variáveis (equação 6.4). O desvio-padrão (s_j) é calculado extraindo-se a raiz quadrada da variância (s_j^2 - equação 6.5). As coordenadas do centróide dos dados escalados pela variância ($\bar{x}_{j(sv)}$) mantém uma relação direta (equação 6.6) com as coordenadas do centróide dos dados originais (\bar{x}_j). As variáveis que antes possuíam variância s_j^2 passam a ter, para os dados escalados, variância unitária (equação 6.7). Tais relações podem ser verificadas comparando-se a média e o desvio padrão entre dados escalados e os dados originais.

$$x_{ij(sv)} = \frac{x_{ij}}{s_j} \quad (6.4)$$

$$s_j^2 = \frac{1}{m-1} \sum_{i=1}^m (x_{ij} - \bar{x}_j)^2 \quad (6.5)$$

$$\bar{x}_{j(sv)} = \frac{1}{m} \sum_{i=1}^m x_{ij(sv)} = \frac{1}{m} \sum_{i=1}^m \frac{x_{ij}}{s_j} = \frac{1}{s_j} \cdot \frac{1}{m} \sum_{i=1}^m x_{ij} = \frac{\bar{x}_j}{s_j} \quad (6.6)$$

$$s_{j(sv)}^2 = \frac{1}{m-1} \sum_{i=1}^m (x_{ij(sv)} - \bar{x}_{j(sv)})^2 = \frac{1}{m-1} \sum_{i=1}^m \left(\frac{x_{ij}}{s_j} - \frac{\bar{x}_j}{s_j} \right)^2 = \frac{1}{s_j^2} \cdot \frac{1}{m-1} \sum_{i=1}^m (x_{ij} - \bar{x}_j)^2 = \frac{s_j^2}{s_j^2} = 1 \quad (6.7)$$

A variância e o desvio-padrão para os dados escalados pela variância é igual a um. Deve-se atentar para este fato ao utilizar esse tipo de pré-processamento, pois ao se tornar igual a variância de todas as variáveis pressupõe-se que a informação que se deseja extrair dos dados distribui-se de maneira igualitária entre todas as variáveis.

6.4. DADOS AUTOESCALADOS

Neste tipo de pré-processamento se escalam pela variância os dados centrados na média. A expressão utilizada para auto-escalar os dados é dada na equação 6.8, em que x_{ij} representa genericamente os elementos da matriz original \mathbf{X} e \bar{x}_j e s_j respectivamente o valor médio e o desvio padrão calculados para a variável j . Os dados auto-escalados $x_{ij(as)}$ apresentam média zero e variância unitária.

$$x_{ij(as)} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (6.8)$$

6.5. CORREÇÃO MULTIPLICATIVA DE SINAL (MSC)

Quando da aquisição dos espectros por técnicas de reflexão, alguns fenômenos não desejados podem ocorrer, como por exemplo o espalhamento da luz, provocado pela não homogeneidade da amostra, principalmente pelas diferenças de granulometria, geometria, empacotamento e orientação das partículas. Também costumam acompanhar os sinais analíticos, obtidos por técnicas instrumentais, ruídos das mais diversas origens. A presença destes ruídos pode dificultar a interpretação, bem como a modelagem das propriedades associadas aos sinais espectrais, sendo recomendada sua eliminação ou minimização. A técnica de correção multiplicativa de sinal (MSC) é um processo matemático que visa corrigir o efeito do espalhamento de luz presente nos espectros obtidos por técnicas de reflexão. Na MSC a interferência relativa ao espalhamento de luz é minimizada com base no espalhamento médio de todos os espectros formadores na matriz de dados \mathbf{X} . Para tanto, é inicialmente calculado o espectro médio x_i , fazendo-se a média das absorvâncias para todas as amostras da matriz de dados \mathbf{X} , para cada uma das colunas (referentes a cada um dos comprimentos de onda do espectro). Calcula-se a regressão linear para cada um dos espectros originais contra o espectro médio,

em cada um dos comprimentos de onda, o que resulta na equação da reta do tipo apresentado pela expressão abaixo.

$$x_i = u_i \mathbf{1} + v_i \cdot x_{\text{médio}} \quad (6.9)$$

Finalmente utilizam-se os coeficientes linear (u_i) e angular (v_i) desta reta para corrigir cada um dos espectros originais, conforme a relação (6.10):

$$x_{\text{corrigido}} = (x_i - u_i \mathbf{1}) / v_i \quad (6.10)$$

6.6. ALISAMENTO

O alisamento dos espectros é realizado com o objetivo de reduzir os efeitos causados pelos ruídos. No tratamento dos dados espectrais foram utilizados dois métodos de alisamento: filtro de média móvel e SPLINE.

6.6.1. MÉDIA MÓVEL

É o tipo mais simples de alisamento utilizado para eliminar ruído de alta frequência. Consiste na aplicação de um algoritmo utilizado para tratamento dos dados, em uma janela móvel que percorre seqüencialmente os dados, calculando a média dos dados contidos dentro da janela, de modo que no final, os dados são substituídos pela média calculada. Considerando uma janela de tamanho L , os L pontos de um vetor (que aqui no caso correspondem a uma linha da matriz de dados (espectro)) são substituídos pelo valor médio calculado utilizando a expressão:

$$x_{ik} = \sum_{l=1}^L x_{i,j+l} u_l \quad (6.11)$$

Sendo $x_{i,j+l}$ ($j=1,2,\dots,L$) os L pontos dos quais será calculada a média, u_l os pesos de convolução e x_{ik} o k -ésimo valor médio calculado que substituirá os L pontos

considerados. No tratamento dado aos espectros deu-se o mesmo peso a todos os pontos considerados no cálculo da média móvel, que equivale ao cálculo de uma média simples, de modo que os pesos de convolução são todos iguais a:

$$u_l = 1/L \quad (l = 1, 2, \dots, L) \quad (6.12)$$

Devido ao fato que os pontos utilizados no cálculo são substituídos pela média calculada sobre estes pontos, o uso da média móvel leva ao final do processo a uma redução do número de variáveis, a extensão em que se dá esta redução depende, portanto, do tamanho escolhido para a janela móvel.

O diagrama mostrado na Figura 6.1 exemplifica o processo de filtragem dos dados utilizando uma janela móvel de tamanho igual a 3 aplicado a um vetor \mathbf{X}_i que poderia representar dados espectrais ou dados de qualquer outra espécie. As médias tomadas de três em três atributos vêm a formar, então, o vetor \mathbf{X}_i (Filtrado).

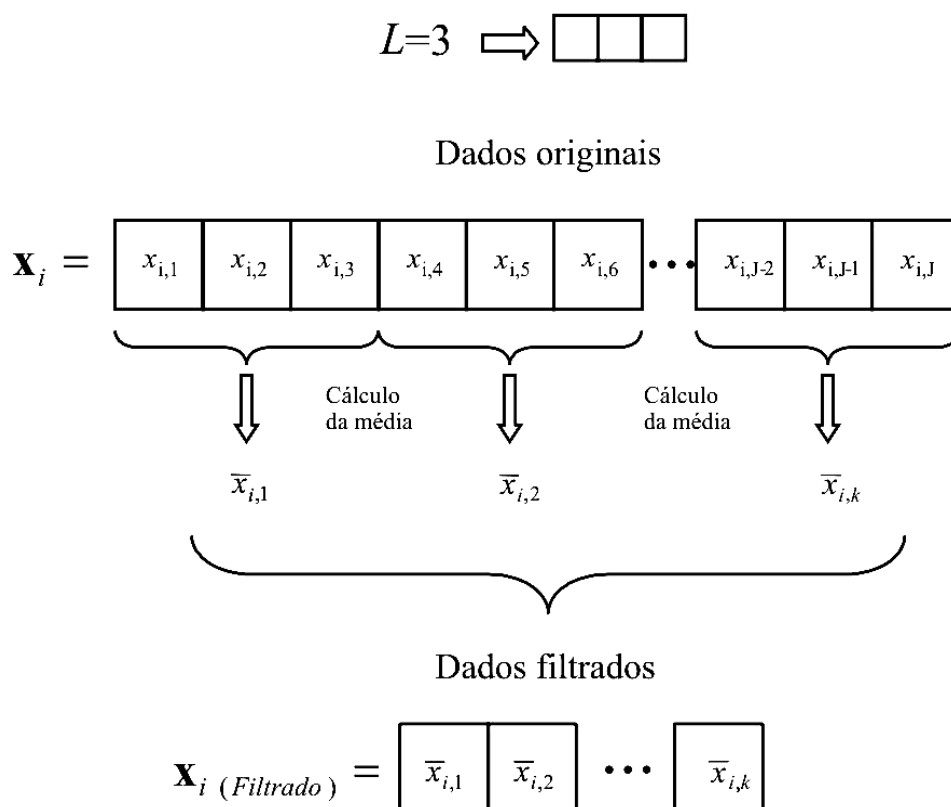


Figura 6.1. Esquema do processo de filtragem utilizando média móvel com uma janela de tamanho 3.

6.6.2. SPLINE

Neste tipo de alisamento, curvas de curvatura mínima são ajustadas aos dados contendo ruído (Figura 6.2). Ao invés de se tentar encontrar uma função genérica que se ajuste a um conjunto de pontos ou uma curva, o que seria praticamente impossível devido à infinita possibilidade de diferentes arranjos e curvas existentes, com o ajuste utilizando SPLINE o que se faz é segmentar o conjunto de pontos ou a curva que se deseja ajustar, em pequenos intervalos, estes sim, fáceis de serem ajustados por meio de uma classe específica de funções pré-estabelecidas (funções-base), geralmente polinômios de baixo grau.

A primeira etapa importante do ajuste por SPLINE é estabelecer o número de segmentos a serem utilizados. O início e o fim destes segmentos são determinados pelo posicionamento dos nós ou pontos de quebra. A Figura 6.3 ilustra este processo para o ajuste de uma gaussiana. Neste exemplo, a gaussiana foi dividida em 4 regiões distintas (região 1, 2, 3 e 4) delimitadas pelos nós (círculos vermelhos). Cada uma destas faixas estabelecida foi ajustada utilizando como funções base polinômios de terceiro grau (f_i).

$$f_i = a_{3i}x^3 + a_{2i}x^2 + a_{1i}x + a_{0i} \quad (i = 1,2,3,4) \quad (6.13)$$

O domínio ao qual se aplica cada uma destas funções é estabelecido pela posição dos nós. Para que estas funções sejam agrupadas de modo a gerar uma curva suave e contínua, os parâmetros das funções de base, que no caso do polinômio são os coeficientes ($a_{k,i}$ ($i=0,1,2,3$)), são escolhidos de modo que na junção das 2 curvas, ou seja, no nó, as derivadas primeira e segunda destas funções sejam iguais, o que é feito resolvendo um sistema de equações. Isto garante que nos nós a função seja bem comportada, ou seja, diferenciável e contínua.

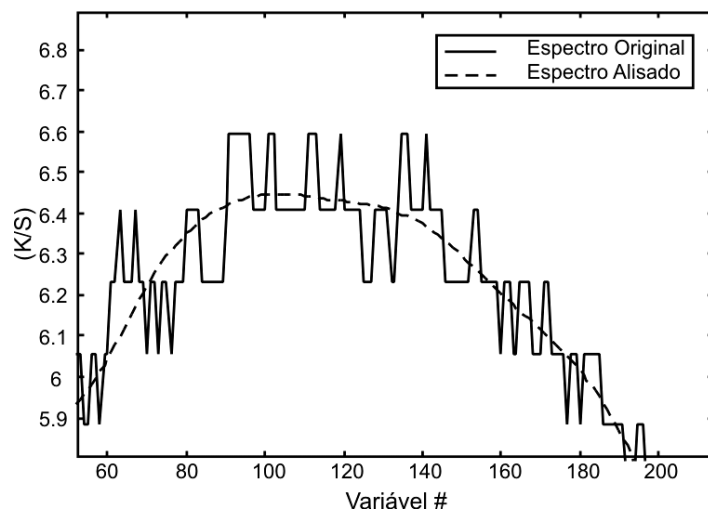


Figura 6.2. Ilustração de um ajuste utilizando SPLINE exemplificado para um segmento de um espectro ruidoso.

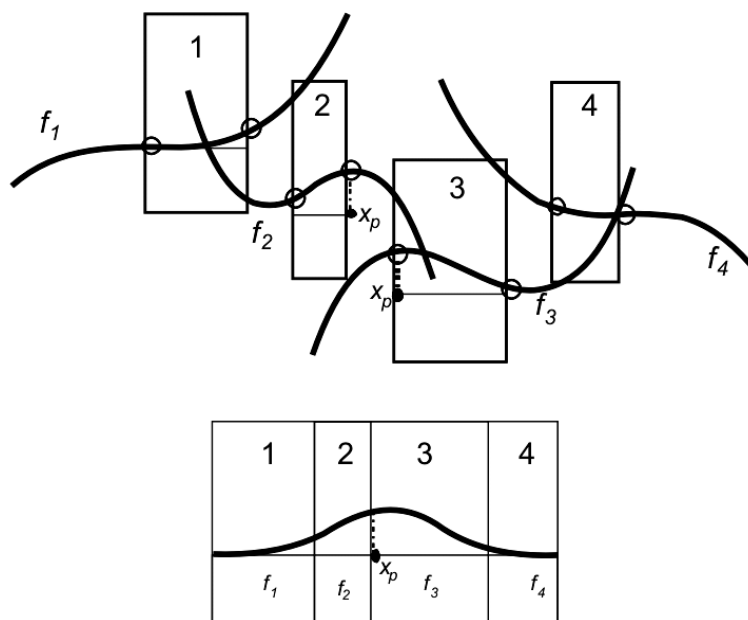


Figura 6.3. Exemplificação do ajuste por SPLINE de uma gaussiana utilizando 4 funções cúbicas (f_1 , f_2 , f_3 e f_4). Os círculos correspondem aos nós utilizados no ajuste.

6.7. KUBELKA – MUNK

A função Kubelka-Munk é normalmente usada para descrever o comportamento da reflectância difusa na região do infravermelho médio e próximo. A equação 4.4, definida anteriormente, descreve esta função.

6.8. APLICAÇÃO DE DERIVADA

A aplicação da primeira ou segunda derivada sobre os dados espectrais brutos é um procedimento que pode destacar ombros espectrais, bem como minimizar o efeito de inclinações provocadas na linha de base dos espectros, devido à morfologia das partículas. Ao aplicarmos as operações de derivação aos espectros, as informações contidas ao longo dos diferentes comprimentos de onda são geralmente acentuadas [Thomas, 1990]. Deve-se ter cuidado com a qualidade dos espectros com os quais deseja-se aplicar o cálculo das derivadas, uma vez que não só os sinais espectrais, como também os ruídos, tornam-se acentuados.

7. DETERMINAÇÃO DO TEOR DE SACAROSE EM AÇÚCAR CRISTAL USANDO ESPECTROSCOPIA MID E REGRESSÃO PLS

7.1. PARTE EXPERIMENTAL

7.1.1. PADRÃO SACAROSE

Para a construção do modelo de regressão foi usado um padrão de sacarose (99,73 %) grau para análise (Merck) na preparação da curva de regressão. Foram preparados, por pesagem em balança analítica, vinte soluções padrão em concentrações de 3,3420 a 3,4108 g 100mL⁻¹ com incrementos de aproximadamente 0,004 g 100mL⁻¹. O padrão de sacarose foi dissolvido em água bidestilada em balão volumétrico calibrado de 100 mL.

7.1.2. AMOSTRAS DE AÇÚCAR CRISTAL

As amostras de açúcar cristal foram coletadas em vinte diferentes usinas de açúcar do Estado de São Paulo e foram pesadas em balança analítica diretamente em balão volumétrico de 10 mL e dissolvidas com água bidestilada.

7.1.3. INSTRUMENTAÇÃO FT-IR E CONDIÇÕES DE ANÁLISE

Os espectros no infravermelho médio com transformada de Fourier foram registrados no espectrômetro Nicolet FT-IR (modelo 520), equipado com cristal de reflectância de seleneto de zinco, material transparente a região de interesse. Entre uma amostra e outra, o cristal foi limpo e seco com papel absorvente. Os espectros no FT-IR foram registrados entre 1400 a 810 cm⁻¹, com resolução de 4 cm⁻¹ para $\log(1/R)$, onde R é a razão da intensidade de reflectância do *background* (cela vazia) para a amostra. Os sinais inicialmente obtidos como $\log(1/R)$ foram expressos como absorvância para facilitar o entendimento e a visualização dos resultados, utilizando-se o *software* do equipamento. Foram realizadas cento e vinte e oito varreduras para cada espectro e a combinação de todas as varreduras resultou num espectro médio.

7.1.4. PREPARAÇÃO DA AMOSTRA PELO MÉTODO DE REFERÊNCIA (POLARIMÉTRICO)

A determinação de sacarose utilizando a técnica da análise polarimétrica realizada envolve três etapas: a preparação da solução de açúcar em água, a clarificação da solução por filtração e a determinação polarimétrica por medidas da rotação ótica da solução clarificada.

Para a preparação da solução da amostra foram pesados $26,0000 \pm 0,0002$ g de açúcar, o qual foi transferido quantitativamente, com água destilada, para um balão volumétrico de 100 mL até ser obtido um volume aproximado de 70 mL. O açúcar foi dissolvido e a solução permaneceu em repouso por, no mínimo, dez minutos (tempo necessário para atingir a temperatura ambiente da sala de análise, $+20^{\circ}\text{C}$). Posteriormente, foi completado o volume do balão volumétrico com água destilada.

A solução foi filtrada em papel de filtro Whatman 41 (15 cm) após a homogeneização.

Os primeiros 10 mL foram descartados e um volume de 50 a 60 mL foi necessário para a medida polarimétrica.

O polarímetro utilizado, Polartronic NIR W2 (SCHMIDT + HAENSCH), mediu valores polarimétricos na região do visível, 589,44 nm (comprimento de onda da luz amarela de sódio). Para minimizar erros na medida dos volumes, a temperatura da sala foi estabilizada em $20,0 \pm 0,5^{\circ}\text{C}$.

7.1.5. ANÁLISE DOS DADOS

O método de regressão utilizado no tratamento de dados deste estudo foi o dos quadrados mínimos parciais (PLS), onde as informações espectrais foram organizadas numa matriz de dados e correlacionada a um vetor que armazena os dados sobre a concentração.

7.1.6. TRATAMENTO DOS DADOS

Os espectros foram inicialmente normalizados (equação 6.1). Os tratamentos matemáticos e estatísticos foram feitos com e sem aplicação da derivada aos valores de absorvância. Assim, foram obtidos diversos modelos de calibração para a quantificação da sacarose presente nas amostras de açúcar cristal. Os programas MATLAB (1997) e PIROUETTE (2001) foram usados para a análise dos dados. Os valores de absorvância registrados entre 810 a 1400 cm^{-1} foram centrados na média.

Para encontrar o número de componentes principais (PC) do modelo PLS foram usados o método de validação cruzada e a soma dos quadrados dos erros de predição estimada (PRESS). O número adequado de componentes principais usado foi aquele que apresentou o menor valor de PRESS.

Para a construção do modelo de regressão foram utilizadas vinte amostras na calibração e seis na validação. O desempenho do modelo foi representado pelo erro quadrático médio (MSE) ou a raiz quadrada do erro (RMSE) mais conhecido como erro padrão de predição (SEP).

A repetitividade que é a tendência de um sinal dar a mesma resposta várias vezes no mesmo experimento, mesmo dia e mesmo analito foi determinada e expressa em termos de estimativa de desvio padrão (s) de acordo com o trabalho de Bellon-Maurel *et al.* [1995]. Para os espectros no infravermelho, a repetitividade do sinal em um determinado comprimento de onda foi calculado pela estimativa de desvio padrão de dez leituras de amostras de mesma concentração (0,3389 g de sacarose por 10 mL de água bidestilada), durante o mesmo experimento e nas mesmas condições. No polarímetro, a repetitividade foi determinada pela medida polarimétrica de dez soluções aquosas de sacarose preparadas pesando-se 26,0000 g da sacarose e dissolvendo-se em 100 mL de água bidestilada.

A precisão intermediária é a tendência de um sinal dar a mesma resposta nas mesmas condições experimentais, mas não no mesmo experimento [Bellon-Maurel *et al.*, 1995; Wojciechowski *et al.*, 1998]. Essas condições expressam a estabilidade do modelo em tempo real. A precisão intermediária do método foi determinada pelos espectros das soluções preparadas para a medida da repetitividade, mas com

medidas no infravermelho e de polarimetria em diferentes dias (3 dias) e foi expressa pelo coeficiente de variação (CV).

7.2. RESULTADOS E DISCUSSÃO

A determinação e quantificação da sacarose pela técnica de reflexão atenuada (ATR), para a aquisição dos espectros no infravermelho médio, foi empregada devido à facilidade de obtenção dos espectros em solução aquosa e pela sua simplicidade. Além disso, nesta região não existe interferentes que possam comprometer a realização dos ensaios.

A Figura 7.1 mostra os espectros obtidos na região do infravermelho médio (870 a 1400 cm^{-1}) para o conjunto de dados de padrões de sacarose usados na construção dos modelos de regressão.

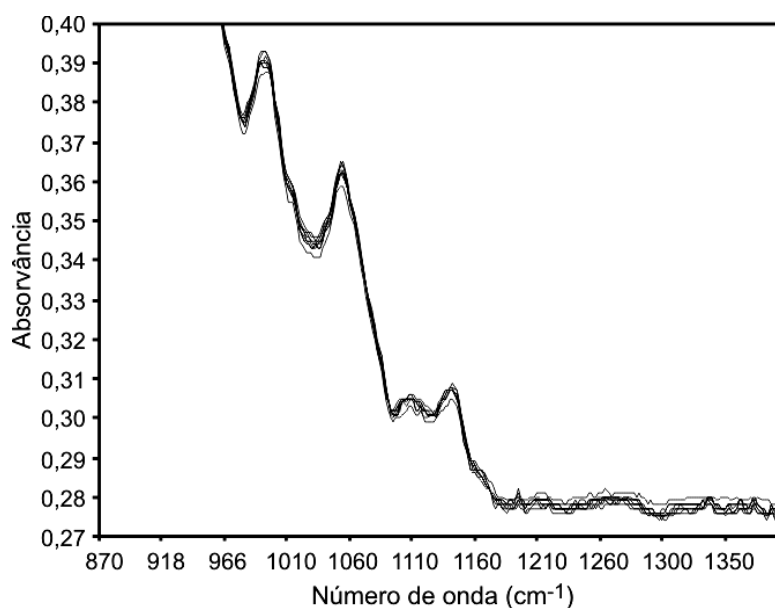


Figura 7.1. Espectros na região do infravermelho médio de vinte soluções de sacarose em água obtidos na região do infravermelho médio com concentrações variando de 3,3420 a 3,4108 $\text{mg } 100\text{g}^{-1}$.

Segundo a literatura [Bellon-Maurel *et al.*, 1995; Costa Filho e Poppi, 2002], os estiramentos encontrados na região entre 1153 a 904 cm^{-1} são atribuídos aos modos normais vibracionais C-O e C-C da sacarose. Na região entre 1199 a 1400 cm^{-1} ocorrem as deformações angulares das ligações O-C-H, C-C-H e C-O-H. As regiões de maior interesse para este trabalho foram: 972 a 984 cm^{-1} relacionadas a deformação angular das ligações C-C; 1056 a 1080 cm^{-1} ligações C-O-C; 1092 a 1148 cm^{-1} ligações C-O e 1192 cm^{-1} ligações O-C-H, C-C-H, C-O-H.

Na Figura 7.2, é apresentado o gráfico que determina o número de componentes principais necessários para estabelecer o modelo com aplicação da derivada segunda. A partir deste gráfico foi verificado que o valor mínimo da soma dos quadrados dos erros de predição foi obtido com três componentes principais.

Para o sistema usando três componentes principais o valor de PRESS é mínimo, da ordem de 0,0084. O uso de um número maior de componentes principais não é conveniente, pois a capacidade de predição do modelo cai, uma vez que os valores de PRESS aumentam como pode ser visto na Figura 7.2.

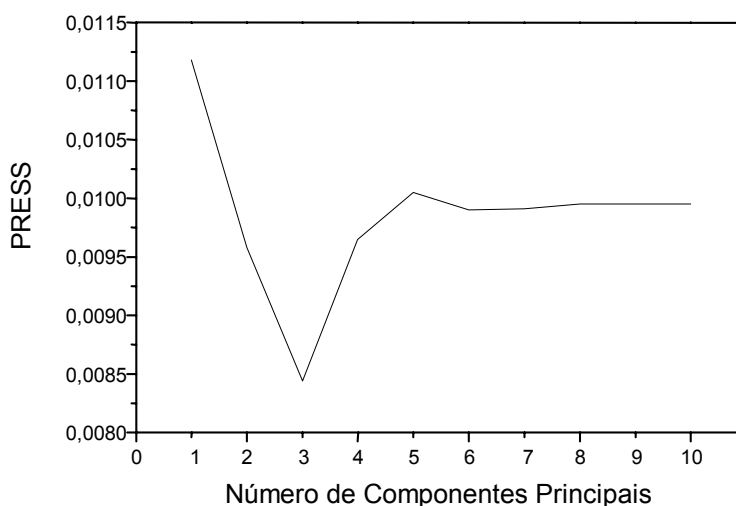


Figura 7.2. Gráfico de PRESS *versus* número de componentes principais.

A Tabela 7.1 apresenta os valores de PRESS, de MSE, o coeficiente de correlação (r) para cada modelo (sem aplicação de derivada; com aplicação de derivada primeira; com aplicação de derivada segunda) e o número de variáveis latentes mais conveniente.

Tabela 7.1. Resultados obtidos para soma dos quadrados dos erros de previsão (PRESS), erro quadrático médio (MSE), coeficiente de correlação (r) e o número de variáveis latentes (VL) para os modelos de calibração.

Método	PRESS (g 100g ⁻¹)	MSE 10 ⁻³ (g 100g ⁻¹)	r	VL
Sem derivada	0,0005	0,02	0,9928	3
Derivada primeira	0,003	0,2	0,9896	4
Derivada segunda	0,008	0,4	0,9755	3

O erro padrão de predição dos métodos, sem derivada, usando derivada primeira e usando derivada segunda foram: 0,002; 0,01 e 0,02 g 100g⁻¹, respectivamente.

Schindler *et al.* [1998], ao analisar açúcares (glicose, frutose e sacarose) por duas técnicas de injeção em fluxo, usando a espectroscopia no infravermelho com transformada de Fourier, encontraram valores de 0,22 e 0,14 mg 100g⁻¹ para o erro padrão de predição na análise de sacarose. No trabalho de Valderrama [2005] foram obtidos erros de 0,3 % para o teor de sacarose em caldo de cana usando NIRS/PLS.

A Figura 7.3 ilustra a curva de regressão obtida a partir dos dados espectrais das soluções-padrões de sacarose na etapa de calibração com a aplicação da derivada segunda. Foi verificado que os pontos apresentaram-se bem próximos à reta do gráfico com um coeficiente de correlação de 0,9755.

Na Figura 7.4, podem ser verificados os resíduos da calibração e de validação para o mesmo modelo com a aplicação da derivada segunda. O modelo de calibração apresentou baixos valores de resíduos, sendo a média igual a 0,004 g 100g⁻¹, comprovando a validade do modelo desenvolvido.

A Tabela 7.2 apresenta os resultados da validação dos modelos construídos, testados para seis amostras externas. As diferenças entre os valores reais e os preditos foram inferiores a 1,3 % para o modelo usando derivada segunda. Para os outros dois modelos os erros foram maiores.

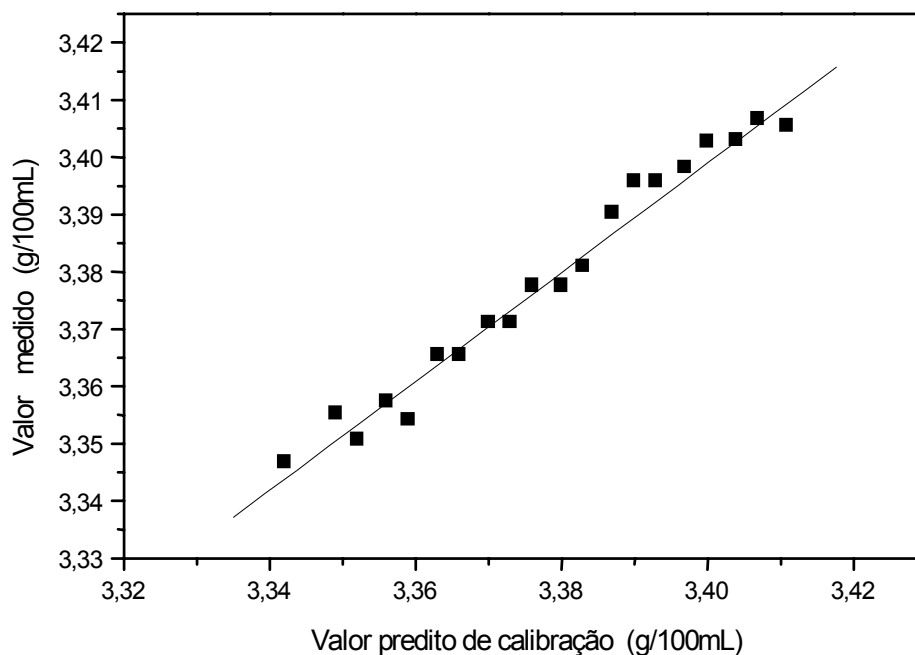


Figura 7.3. Curva de regressão obtida na etapa de calibração para o modelo usando derivada segunda.

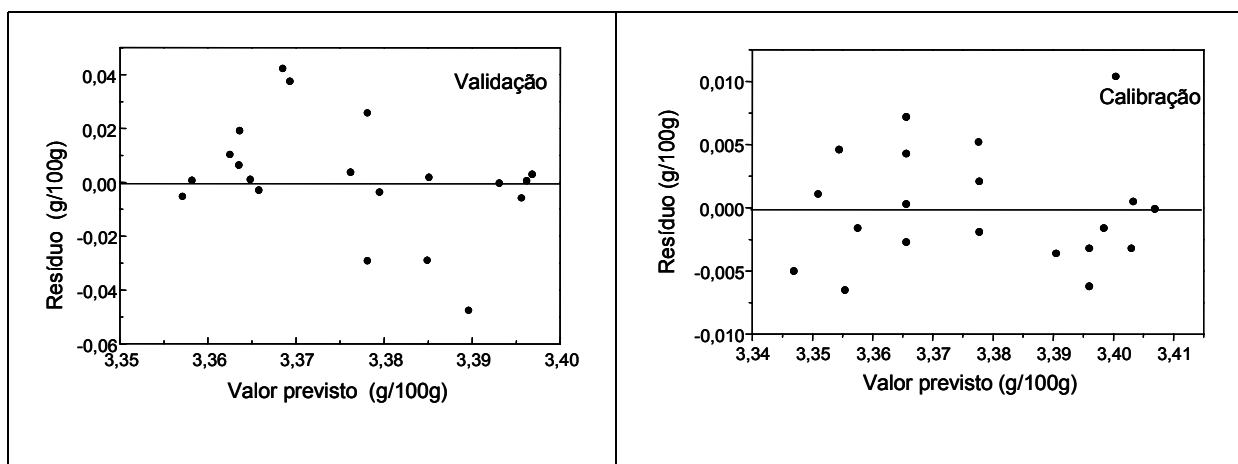


Figura 7.4. Gráfico de resíduos *versus* valores previstos para amostras do conjunto de calibração e amostras do conjunto de validação do modelo construído usando derivada segunda.

Tabela 7.2. Resultado da validação dos modelos para a determinação do teor de sacarose, em g 100g⁻¹.

Valor real	Modelo sem aplicação derivada		Modelo usando derivada primeira		Modelo usando derivada segunda	
	Valor previsto de validação	Erro relativo (%)	Valor previsto de validação	Erro relativo (%)	Valor previsto de validação	Erro relativo (%)
3,4108	3,3974	0,4	3,3998	0,3	3,3915	0,6
3,3968	3,4126	-0,5	3,4186	-0,6	3,3935	0,1
3,3828	3,4295	-1,4	3,4212	-1,1	3,3739	0,3
3,3699	3,4098	-1,2	3,4126	-1,3	3,3931	-0,7
3,3559	3,4115	-1,7	3,4109	-1,6	3,3997	-1,3
3,3420	3,3350	0,2	3,3837	-1,2	3,3703	-0,8
		$\bar{E} = 0,9$		$\bar{E} = 1,0$		$\bar{E}^* = 0,6$

* \bar{E} = Módulo do erro relativo médio (%).

Após a análise dos resultados dos modelos, sem derivada e com aplicação da derivada primeira ou derivada segunda, foi concluído que o modelo, usando a derivada segunda, apresentou melhores resultados para a determinação de sacarose em amostras de açúcar cristal, pois na etapa de validação ficou demonstrado que é o modelo que fornece os melhores resultados para a previsão de amostras externas.

A Tabela 7.3 apresenta os resultados obtidos na determinação do teor de sacarose para o melhor modelo de regressão construído, usando os dados dos espectros no infravermelho e a comparação com os resultados obtidos pelo método empregando a polarimetria. As diferenças observadas foram mínimas, sendo que a maior diferença foi de 0,5 %, a menor de 0,03 % e a média de 0,34 %.

A exatidão do método encontrada empregando o modelo com derivada segunda e expressa pela raiz quadrada do erro médio quadrático de previsão (RMSEP) foi de 0,02 g 100⁻¹.

Usando métodos multivariados de análise na determinação do teor de sacarose em açúcar-de-cana por FT-IR/ATR e polarimetria, Cadet *et al.* [1991] obtiveram um erro relativo médio de 11,5 %. Já no trabalho de Valderrama [2005] foram obtidos erros de 0,3 a 0,4 % para amostras de caldo-de-cana usando a espectroscopia NIR e regressão PLS.

A repetitividade dos métodos avaliada para dez preparações de uma mesma amostra é apresentada na Tabela 7.4. Os desvios-padrão para o método, usando medidas no infravermelho, foram melhores em relação ao método usando as medidas polarimétricas, sendo os desvios da ordem de $0,035 \text{ g } 100\text{g}^{-1}$. Este pequeno desvio-padrão demonstra a boa repetitividade das medidas.

Bellon-Maurel *et al.* [1995] determinaram os teores de glicose, maltodextrina, maltose e maltotriose por FT-IR/ATR, usando métodos multivariados e obtiveram desvios entre $0,39$ a $0,64 \text{ g kg}^{-1}$ em relação aos valores de referência. Cadet *et al.* [1991] determinaram o teor de sacarose em açúcar-de-cana por FT-IR/ATR e polarimetria, obtendo resultados no método polarimétrico com um desvio de $1,8 \%$ em relação aos valores de referência, enquanto o método espectroscópico apresentou um desvio de $0,3 \%$. Valderrama [2005] obteve $0,6 \%$ na medida da repetitividade para sacarose usando NIRS/PLS.

Tabela 7.3. Comparação dos resultados obtidos pelo modelo de regressão construído a partir dos dados espectroscópicos e de polarimetria para as amostras de açúcar cristal, provenientes de usinas de açúcar do Estado de São Paulo.

Usinas	Sacarose ($\text{g } 100\text{g}^{-1}$)		Erro Relativo (%)
	Polarimetria	FT-IR/ATR*	
1	99,71	99,67	0,04
2	99,67	100,06	-0,4
3	99,77	99,73	0,04
4	99,76	99,94	-0,2
5	99,70	100,18	-0,5
6	99,74	100,12	-0,4
7	99,81	99,61	0,2
8	99,80	99,76	0,04
9	99,81	100,14	-0,3
10	99,74	99,38	0,4
11	99,70	100,14	-0,4
12	99,80	99,51	0,3
13	99,80	99,91	-0,1
14	99,76	99,80	-0,04
15	99,85	99,88	-0,03
16	99,71	100,21	-0,5
17	99,78	99,81	-0,03
18	99,80	99,52	0,3
19	99,77	100,07	-0,3

* Modelo com aplicação de derivada segunda e 3 variáveis latentes.

A precisão intermediária dos métodos avaliada em preparações de soluções de sacarose com medidas em três dias diferentes foi expressa em termos do coeficiente de variação (Tabela 7.4). Neste caso, o método polarimétrico apresentou uma melhor precisão intermediária.

Tabela 7.4. Comparação dos desvios-padrão e coeficientes de variação em uma amostra de açúcar cristal de usina, obtidos em três dias diferentes de análise, com a realização de dez repetições analíticas, utilizando as técnicas espectroscópica e polarimétrica (Precisão Intermediária).

Dia	Infravermelho ¹		Polarimetria ²	
	Média ± DP	CV ⁴	Média ± DP ³	CV ⁴
Primeiro	1,63 ± 0,04	2	99,6 ± 0,1	0,1
Segundo	1,64 ± 0,04	2	99,9 ± 0,1	0,1
Terceiro	1,64 ± 0,04	2	99,5 ± 0,1	0,1

¹ Valores de reflectância obtidos no espectro de infravermelho utilizando seis comprimentos de onda diferentes. Estas medidas são adimensionais; ² Valores expressos em g 100g⁻¹; ³ DP = desvio-padrão (g 100g⁻¹); ⁴ CV = coeficiente de variação, em %.

7.3. CONCLUSÃO

- Os resultados obtidos na determinação do teor de sacarose em açúcar de usina por FT-IR/ATR foram muito satisfatórios e apresentaram um erro padrão de predição igual a 0,02 g 100g⁻¹ para o modelo de regressão usando derivada segunda.
- A repetitividade do método usando espectroscopia no infravermelho foi melhor em relação ao método polarimétrico, porém a precisão intermediária do método empregando polarimetria foi melhor. A precisão do método por FT-IR/ATR usando derivada segunda foi de 0,04 g 100g⁻¹.
- Os resultados obtidos demonstram que o método utilizando a técnica de espectroscopia FT-IR/ATR e regressão multivariada, empregando o método dos quadrados mínimos parciais (PLS), pode ser aplicado na determinação do teor de sacarose em amostras de açúcar cristal com boa precisão e exatidão.
- A espectroscopia usando FT-IR também pode ser adaptada para determinar outros componentes do açúcar, como o teor de umidade ou teor de sólidos.

8. DETERMINAÇÃO DO TEOR DE AÇÚCARES TOTAIS, DE PROTEÍNAS E DE UMIDADE EM AMOSTRAS DE CAFÉ CRU USANDO ESPECTROSCOPIA NIR E O MÉTODO DE REGRESSÃO PLS

Nesta etapa do trabalho foram desenvolvidos modelos de calibração para a determinação dos teores de açúcares totais, de proteínas e de umidade aplicados a amostras de café cru.

8.1. PARTE EXPERIMENTAL

8.1.1. AMOSTRAS DE CAFÉ

Para a construção dos modelos de regressão foram usadas entre 51 a 56 amostras de café cru, da variedade arábica, procedentes de diferentes regiões produtoras de café do Brasil, dos estados de São Paulo, Minas Gerais e Bahia. Todas as amostras foram inicialmente homogeneizadas em moinho de facas até a obtenção de partículas de tamanho reduzido que passaram por peneira de 0,84 mm de tamanho de abertura de poro. Algumas das amostras foram excluídas pois apresentavam comportamento de outliers.

8.1.2. ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO E ANÁLISE DOS DADOS

Os espectros no infravermelho próximo foram coletados em um espectrômetro BOMEM DA-08, sendo realizadas 3 réplicas para cada amostra, com resolução de 4 cm^{-1} , 16 varreduras para cada espectro e a região espectral utilizada foi de $4500\text{ a }10000\text{ cm}^{-1}$. O instrumento foi equipado com um acessório de reflectância difusa (JASCO), sendo os sinais registrados como $\log(1/R)$ (R = reflectância) e transformados para absorvância através do *software* do equipamento.

O método de regressão multivariada utilizado para o tratamento de dados deste estudo, foi o PLS e o *software* utilizado foi o Matlab 5.1, para o tratamento dos

dados e na construção dos modelos das rotinas presentes no toolbox QUIMIOPAC [Faria e Ferreira, 2002).

8.1.3. ESTATÍSTICA

O conjunto total de espectros foi dividido em dois subconjuntos: um de calibração e outro de validação externa. Para compor estes dois conjuntos foi aplicada uma análise de componentes principais aos resultados obtidos pelos métodos tradicionais de análise das propriedades das amostras de café (teor de umidade, proteínas, açúcares totais, cinzas, lipídios, carboidratos....) e agruparam-se as amostras de forma que as que fizeram parte do conjunto de calibração apresentavam a maior faixa de concentração, ou seja, as amostras do conjunto de validação estavam dentro da faixa de concentração das amostras do conjunto de calibração para as propriedades de interesse (teor de umidade, proteínas e açúcares totais). O Quadro 8.1 identifica as amostras utilizadas em cada um destes conjuntos e a Tabela 8.1 apresenta a identificação completa das amostras, como a origem e os valores dos teores de umidade, açúcares totais e proteínas.

Quadro 8.1. Conjunto das amostras de café cru usadas na etapa de calibração e das amostras usadas na etapa de validação.

AMOSTRAS DE CALIBRAÇÃO	AMOSTRAS DE VALIDAÇÃO
1, 3, 5, 6, 8, 9, 12, 14, 16, 17, 19, 20, 22, 23, 24, 27, 28, 29, 30, 31, 32, 36, 39, 44, 45, 47, 51, 52, 54, 56, 57	2, 4, 7, 10, 11, 13, 15, 21, 25, 26, 33, 34, 35, 37, 38, 40, 41, 42, 43, 46, 48, 49, 50, 53, 55

Para se determinar o número de variáveis latentes (LV) utilizadas no modelo, foi feita uma validação cruzada (validação interna) no conjunto de calibração: uma amostra do conjunto de calibração é excluída, o modelo é construído e, então, estimado o seu teor de açúcar. O processo é repetido até que todas as amostras sejam previstas para 1, 2, ... variáveis latentes. A habilidade do modelo de

calibração para estimar (ou prever) o teor de açúcar baseado nos dados dos espectros NIR gerados foi avaliada usando os erros de predição e os coeficientes de correlação entre os valores dos teores de açúcar estimados pelo modelo utilizando espectros NIR e os valores do método de referência das amostras do conjunto de calibração. Os parâmetros de erro avaliados foram o SEP (eq. 5.9), o PRESS (eq. 5.8), o coeficiente de correlação (eq. 5.14), a variância relativa (eq. 5.11) e o erro percentual relativo (eq. 5.13).

8.1.4. MÉTODOS DE REFERÊNCIA

8.1.4.1. MÉTODO PARA A DETERMINAÇÃO DE AÇÚCARES TOTAIS

O método utilizado para a determinação do teor de açúcares totais e redutores foi o de Munson-Walker [AOAC, 2000]. Foram pesados 10 g de amostra em balança analítica, transferidos para balão volumétrico de 250 mL com cerca de 100 mL de água destilada. Foi neutralizado com solução de hidróxido de sódio 35 % (m/v) até $\text{pH} \leq 7$. Adicionaram-se 6 mL de solução de ferrocianeto de potássio $0,25 \text{ mol L}^{-1}$ e 7 mL de solução de acetato de zinco $1,0 \text{ mol L}^{-1}$. Após agitação, completou-se o volume e filtrou-se para erlenmeyer de 500 mL.

A- Açúcares redutores: juntaram-se 25 mL das soluções A (34,639 g de sulfato de cobre pentahidratado e 0,5 mL de ácido sulfúrico concentrado em 500 mL de água destilada) e B (172,0 g de tartarato de sódio e potássio e 50,0 g de hidróxido de sódio e 500 mL de água destilada) de Fehling e 50 mL de água destilada. Aqueceu-se até ebulição. Adicionaram-se 10 mL do filtrado e deixou-se em ebulição por 2 minutos. Filtrou-se a vácuo, em cadinho de porcelana P-2 previamente tarado, lavando-se com água a 60°C e com 10 mL de etanol e 10 mL de éter etílico. Secou-se em estufa a 105°C por 1 hora, esfriou-se em dessecador e pesou-se o óxido cuproso (Cu_2O).

B- Açúcares invertidos: Pipetaram-se 10 mL do filtrado para balão volumétrico de 100 mL. Adicionaram-se 5 mL de ácido clorídrico concentrado, aqueceu-se durante 5 minutos à temperatura de 38-70°C e esfriou-se. Neutralizou-se a pH ≤ 7 com solução de hidróxido de sódio e completou-se o volume. Procedeu-se como descrito acima para os açúcares redutores.

O teor de açúcares redutores (em glicose) é determinado pela expressão de Hammond usando os valores da massa de óxido cuproso obtida. A porcentagem de açúcares redutores é calculada pela expressão:

$$\% \text{ açúcares redutores} = \frac{\text{mg de glicose} \times 25}{\text{amostra} \times 10} \quad (8.1)$$

O teor de açúcares não-redutores (Ac.n.red.) expressos em sacarose é determinado pela expressão de Hammond usando os valores da massa de óxido cuproso obtida expresso como açúcar invertido. A porcentagem de açúcares invertidos é calculada pela expressão:

$$\% \text{ açúcares invertidos} = \frac{\text{mg de açúcar invertido} \times 250}{\text{g de amostra} \times 50} \quad (8.2)$$

$$\% \text{ Ac.n.red.} = (\% \text{ açúcares invertidos} - \% \text{ açúcares redutores}) \times 0,95 \quad (8.3)$$

O teor de açúcares totais é determinado pela seguinte expressão:

$$\% \text{ açúcares totais} = \% \text{ açúcares redutores} + \% \text{ Ac.n.red.} \quad (8.4)$$

8.1.4.2. MÉTODO PARA A DETERMINAÇÃO DE UMIDADE

O método utilizado para a determinação de umidade foi o da AOAC [2000], nº 968.11. Foram pesados em balança analítica 5 g da amostra homogeneizada em

cápsula de alumínio tarada. Em seguida, as cápsulas foram colocadas em estufa a vácuo por 5 ½ horas, a 98-100°C e pressão < 25 mmHg. As cápsulas foram colocadas em dessecador, esfriadas e pesadas. O teor de umidade foi expresso como a razão entre o peso da amostra dessecada e o peso da amostra pesada vezes 100.

8.1.4.3. MÉTODO PARA A DETERMINAÇÃO DE PROTEÍNA

O método utilizado para a determinação de proteínas foi o de Kjeldhal [IAL, 1985]. Foram transferidos para balão de Kjeldhal 0,5000 ± 0,0001 g de amostra. Juntaram-se 15 g da mistura de sulfato de potássio anidro e de sulfato de cobre 1:1 (m/m), 0,02 g de selênio metálico e 20 mL de ácido sulfúrico. Digeriu-se a quente até destruição completa da matéria orgânica (mínimo de 1 hora). Retirou-se o balão, esfriou-se e juntaram-se 180 mL de água. Deixou-se escorrer pelas paredes do balão 90 mL de solução de hidróxido de sódio 50 % (m/v), conectou-se o balão ao condensador e destilou-se, recolhendo 120 mL do destilado em erlenmeyer contendo cerca de 30 mL de solução de ácido bórico e titulou-se com solução de ácido clorídrico 0,1 mol L⁻¹ usando solução de verde de bromocresol até mudança de coloração de verde para rosado. Foi calculado o teor de proteínas usando a equação 8.5 e o foi usado o fator 5,75 para converter o teor de nitrogênio em proteína.

$$\text{g/100g de proteína} = \frac{\text{vol. HCl (mL)} \times \text{mol L}^{-1} \text{ HCl} \times 14 \times 100 \times 5,75}{\text{massa amostra (g)} \times 1000} \quad (8.5)$$

8.2. RESULTADOS OBTIDOS USANDO OS MÉTODOS DE REFERÊNCIA

A Tabela 8.1 apresenta a identificação das amostras e os valores obtidos pelos métodos de referência para os teores de açúcares totais, de proteínas e de umidade.

Tabela 8.1. Identificação das amostras de café cru e valores obtidos pelos métodos de referência para os teores de açúcares totais, de proteínas e de umidade.

Amostra - Origem	Açúcares Totais (g/100g)	Proteínas (g/100g)	Umidade (g/100g)
1 – São Manuel – SP	9,1 ± 0,2	13,4 ± 0,1	11,0 ± 0,2
2 – Franca – SP	8,5 ± 0,1	13,8 ± 0,1	10,36 ± 0,03
3 – Pedregulho – SP	7,7 ± 0,5	13,1 ± 0,1	10,33 ± 0,09
4 – Jeriquara – SP	8,9 ± 0,2	14,3 ± 0,2	9,75 ± 0,01
5 – Espírito Santo do Pinhal - SP	6,5 ± 0,1	13,3 ± 0,1	10,12 ± 0,00
6 – Espírito Santo do Pinhal - SP	9,6 ± 0,3	13,70 ± 0,00	10,3 ± 0,2
7 – Espírito Santo do Pinhal - SP	7,9 ± 0,1	14,1 ± 0,1	10,05 ± 0,05
8 – Águas da Prata - SP	9,1 ± 0,2	12,9 ± 0,2	10,55 ± 0,01
9 – Santo Antônio do Jardim - SP	8,6 ± 0,4	13,6 ± 0,1	10,4 ± 0,1
10 – Espírito Santo do Pinhal - SP	9,5 ± 0,1	13,3 ± 0,1	9,98 ± 0,05
11 – Mogi Guaçu - SP	9,6 ± 0,4	14,1 ± 0,1	10,521 ± 0,005
12 – Espírito Santo do Pinhal - SP	10,9 ± 0,6	13,55 ± 0,00	9,507 ± 0,007
13 – Santo Antônio do Jardim - SP	9,0 ± 0,2	13,35 ± 0,07	10,01 ± 0,03
14 – Brejões - BA	10,3 ± 0,1	14,51 ± 0,09	9,665 ± 0,002
15 – Paracatu - MG	8,9 ± 0,2	12,61 ± 0,07	11,1 ± 0,2
16 – Patrocínio - MG	9,2 ± 0,3	14,51 ± 0,07	10,94 ± 0,08
17 – Dois Córregos - MG	9,0 ± 0,2	13,2 ± 0,2	10,31 ± 0,06
18 - *	-	-	-
19 – Poços de Caldas - MG	8,4 ± 0,2	13,87 ± 0,09	10,96 ± 0,08
20 – Patrocínio - MG	11,2 ± 0,7	13,57 ± 0,05	10,4 ± 0,1
21 – Poços de Caldas - MG	9 ± 1	12,8 ± 0,1	10,554 ± 0,002
22 – Poços de Caldas - MG	9,8 ± 0,8	12,86 ± 0,04	10,76 ± 0,07
23 – Patrocínio - MG	10,89 ± 0,01	13,28 ± 0,05	9,62 ± 0,04
24 – Patrocínio - MG	10,2 ± 0,1	13,7 ± 0,1	8,8 ± 0,2
25 – Alfenas - MG	9,2 ± 0,5	13,77 ± 0,05	9,410 ± 0,008
26 – Pedregulho – SP	9,6 ± 0,1	14,22 ± 0,09	8,8 ± 0,1
27 – Brejões - BA	7,6 ± 0,3	14,1 ± 0,2	11,906 ± 0,004
28 – Brejões – BA	9,1 ± 0,4	13,9 ± 0,1	12,71 ± 0,03
29 – Muzambinho – MG	8 ± 1	14,55 ± 0,03	9,8 ± 0,3
30 – Arceburgo – MG	7,5 ± 0,6	13,6 ± 0,4	9,6 ± 0,2
31 – Rio Paranaíba – MG	8,1 ± 0,9	14,0 ± 0,2	8,6 ± 0,1
32 – Carmo de Rio Claro – MG	8,1 ± 0,3	13,5 ± 0,1	9,86 ± 0,03
33 – Cabo Verde – MG	6,8 ± 0,6	12,4 ± 0,2	9,84 ± 0,04
34 – Borda da Mata – MG	9,2 ± 0,3	14,14 ± 0,09	10,4 ± 0,2
35 – Ouro Fino – MG	9,3 ± 0,3	12,9 ± 0,2	9,90 ± 0,06
36 – Coromandel – MG	9,3 ± 0,6	13,34 ± 0,04	9,6 ± 0,3
37 – Marília – SP	7,4 ± 0,1	12,9 ± 0,4	9,67 ± 0,02
38 – Varginha - MG	8 ± 1	13,614 ± 0,005	9,369 ± 0,009
39 – Espírito Santo do Pinhal - SP	8 ± 1	13,08 ± 0,03	8,8 ± 0,3
40 – Espírito Santo do Pinhal - SP	9,2 ± 0,1	13,2 ± 0,1	8,48 ± 0,02
41 – Santo Antônio do Jardim - SP	9 ± 1	13,4 ± 0,2	8,69 ± 0,06
42 – Espírito Santo do Pinhal - SP	9 ± 2	12,5 ± 0,2	7,91 ± 0,06
43 – Santo Antônio do Jardim - SP	10,0 ± 0,1	14,1 ± 0,3	8,7 ± 0,1
44 – Águas da Prata - SP	10,0 ± 0,6	13,29 ± 0,08	8,5 ± 0,4
45 – Espírito Santo do Pinhal - SP	8,9 ± 0,2	13,975 ± 0,006	7,45 ± 0,09
46 – Espírito Santo do Pinhal - SP	10 ± 1	13,9 ± 0,2	8,37 ± 0,08
47 – Mogi Guaçu - SP	9,4 ± 0,7	13,8 ± 0,2	8,27 ± 0,09
48 – Franca – SP	8,1 ± 0,2	13,5 ± 0,2	11,60 ± 0,05
49 – Franca – SP	9,4 ± 0,1	13,01 ± 0,09	10,95 ± 0,06
50 – Barreiras - BA	9,0 ± 0,4	13,05 ± 0,06	10,12 ± 0,04
51 – Barreiras - BA	8,2 ± 0,4	11,09 ± 0,05	11,54 ± 0,06
52 – Barreiras - BA	8,6 ± 0,5	12,73 ± 0,02	9,94 ± 0,06
53 – Marília – SP	9,2 ± 0,7	12,8 ± 0,2	10,8 ± 0,2
54 – Marília – SP	10,1 ± 0,7	12,6 ± 0,1	9,54 ± 0,09
55 – Varginha - MG	9,4 ± 0,2	12,9 ± 0,1	10,76 ± 0,03
56 – Varginha - MG	8,7 ± 0,2	12,8 ± 0,1	10,89 ± 0,05
57 – Varginha – MG	9,3 ± 0,1	13,31 ± 0,03	11,2 ± 0,1

* amostra excluída por contaminação.

8.3. RESULTADOS E DISCUSSÃO

8.3.1. AÇÚCARES TOTAIS

Para a construção dos modelos de regressão foram utilizados os valores médios dos teores de açúcares totais obtidos na determinação empregando o método de referência e 147 espectros das amostras de café cru. Para formar o conjunto de calibração foram selecionados 81 espectros, correspondentes às 27 amostras, de forma a cobrir toda a faixa do teor de açúcar nas amostras. Os 66 espectros restantes, correspondentes às 22 amostras, foram utilizados para formar o conjunto de validação externa.

Devido aos efeitos de espalhamento de luz provocado pela não-homogeneidade das amostras decorrentes principalmente de diferença de granulometria foi necessário utilizar a técnica de correção de espalhamento de luz (MSC) nos espectros transformados pela equação de Kubelka-Munk antes de se proceder à construção dos modelos de regressão. Os resultados das transformações Kubelka-Munk e a correção multiplicativa de sinal (MSC) nos dados dos espectros originais podem ser visualizados na Figura 8.1.

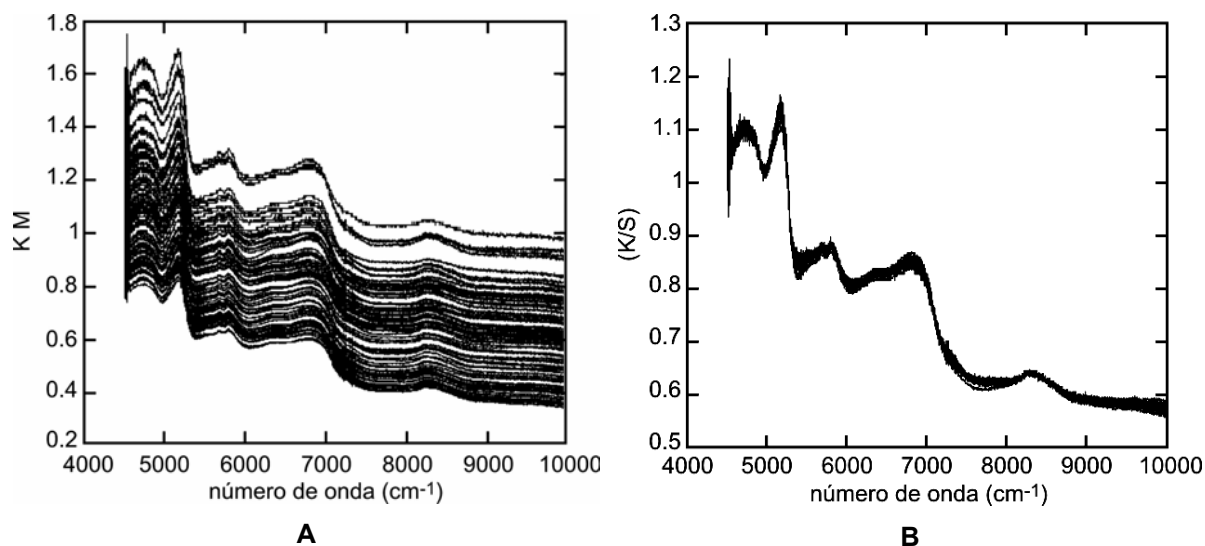


Figura 8.1. **A** – Espectros NIR após a transformação Kubelka-Munk; **B** – Espectros NIR após a correção multiplicativa de sinal (MSC).

A matriz de dados resultante, a partir dos espectros transformados, foi escalada pela variância, ou seja, cada coluna foi dividida pelo respectivo desvio-padrão.

A seleção de variáveis foi obtida através das informações dadas pelo correlograma [Bellon-Maurel *et al.*, 1995]. O correlograma foi gerado calculando-se o coeficiente de correlação linear (r) entre as variáveis independentes, ou seja, os resultados das absorvâncias (bloco-Y) e cada uma das variáveis dependentes, ou seja, os resultados do teor de açúcar obtidos pelo método de referência (bloco-X), de modo que no final do processo tem-se para cada variável independente um valor associado de r que ao serem dispostos em um gráfico geram o correlograma. O coeficiente de correlação linear (r_i) entre estes dois blocos foi utilizado para compor o correlograma em um gráfico no qual as ordenadas são os coeficientes de correlação linear e as abscissas os comprimentos de onda lidos (os números representativos de cada variável).

Para escolher qual o melhor valor de corte a ser empregado, acompanhou-se o valor do PRESS para o conjunto de calibração e de validação externa fazendo-se variar o valor de corte, isto foi feito para modelos PLS com 1 até 10 variáveis latentes (VLs). A Figura 8.2 apresenta estes resultados. Observa-se que o PRESS de calibração aumenta significativamente, quando se utilizam valores de corte superiores a 0,3. Nesta figura observa-se, também, que o PRESS da calibração diminui, à medida que o modelo se torna mais complexo, no entanto, esta tendência não é observada com relação ao PRESS da validação externa. Como o que se pretende é um modelo com uma boa capacidade de predição optou-se pelo modelo que apresentou o menor PRESS no conjunto de validação (assinalado com asterisco na Figura 8.2-A), que corresponde ao modelo que utiliza 6 VLs e um valor de corte igual a 0,1.

O correlograma referente aos dados é mostrado na Figura 8.2-B. Observa-se que a correlação linear entre o bloco-X e o bloco-Y não é muito pronunciada apresentando valores inferiores a 0,6. Com um valor de corte de 0,1 foi possível selecionar 831 variáveis das 2250 variáveis originais. Os intervalos de números de

onda selecionados estão relacionados principalmente com as bandas de combinação e sobretons (estiramento + deformação angular) das ligações C-H de grupos CH_3 (7353; 5900; 5865; 5797 cm^{-1}) e ligação O-H (6757- 6329 cm^{-1}) da glicose.

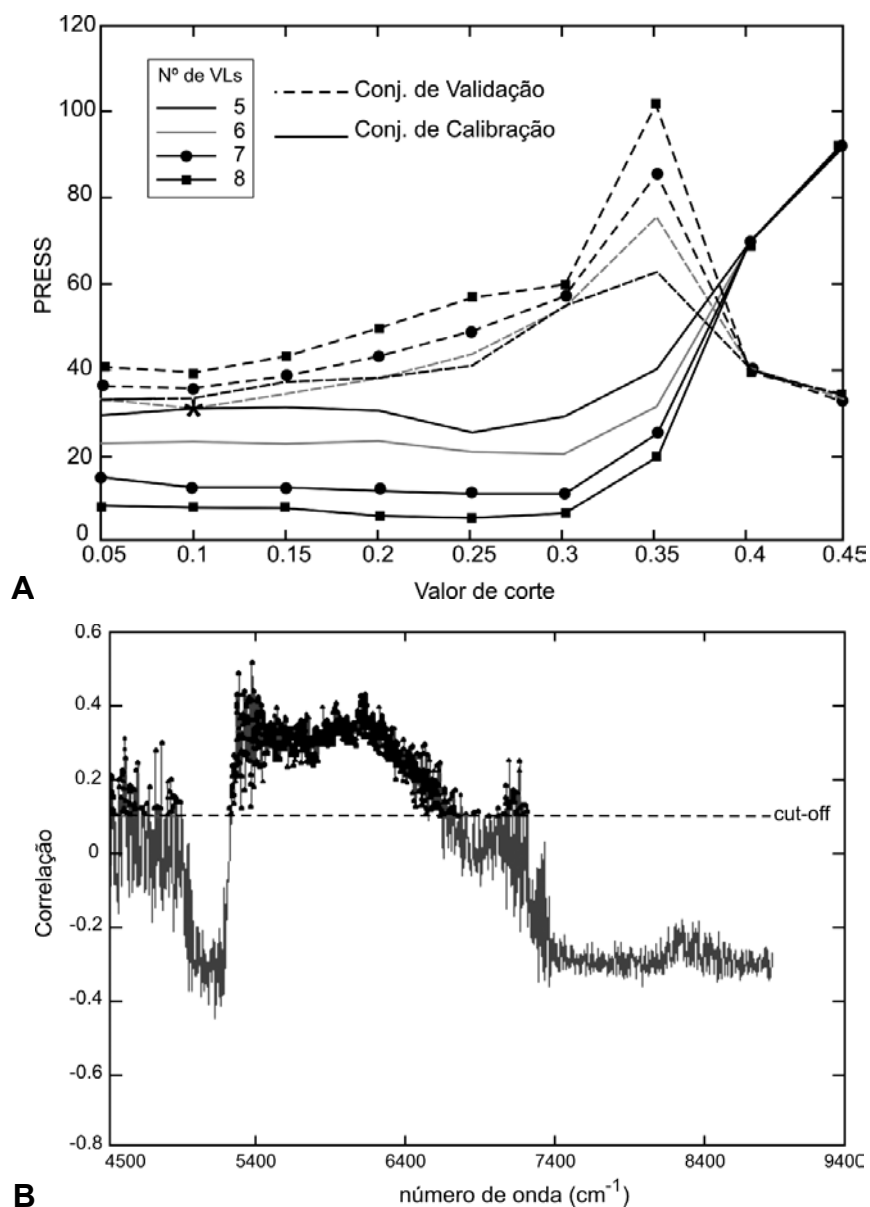


Figura 8.2. **A-** Gráfico do PRESS de validação interna e de validação externa em função do valor de corte do correlograma (selecionando apenas as variáveis positivamente correlacionadas) dos modelos PLS utilizando 5, 6, 7 e 8 VLs. **B-** Gráfico do correlograma apresentando o valor de corte utilizado (0,1) e as variáveis selecionadas.

Ajuste do modelo: O número de variáveis latentes, utilizado no modelo PLS, foi determinado a partir do PRESS exibido no conjunto de validação externa. A Figura 8.3 mostra o comportamento do PRESS para o conjunto de calibração e de validação externa em função do número de variáveis latentes (VL) utilizadas nos modelos. O valor mínimo do PRESS exibido pelo conjunto de validação externa é atingido quando se utiliza um número de VL igual a 6. O comportamento do PRESS da validação indica que utilizando-se modelos com mais de 6 VLs os dados sofrem um sobre-ajuste (*overfit*), que fica evidenciado também no decréscimo monotônico do PRESS do conjunto de calibração com o aumento da complexidade do modelo. Problemas sub-ajuste (*underfit*) ficam evidenciados no aumento dos valores do PRESS de ambos os conjuntos ao se utilizar modelos com um número de VLs inferiores a seis. Levando-se em conta os fatores abordados acima, concluiu-se que o uso de seis variáveis latentes é o mais apropriado na construção do modelo regressão PLS para determinar o teor de açúcar em amostras de café cru.

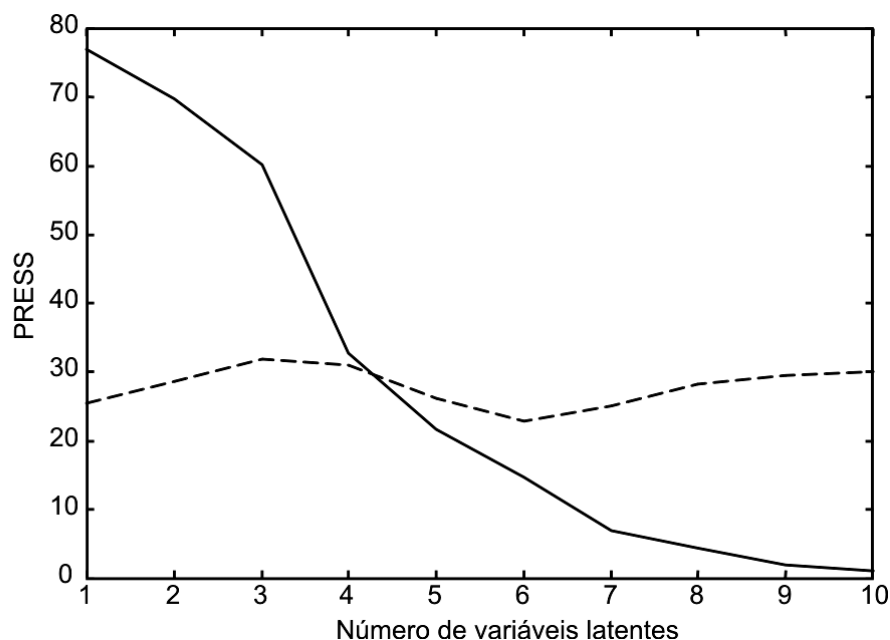


Figura 8.3. Gráfico do PRESS dos conjuntos de validação interna (—) e externa (---) em função do número de variáveis latentes utilizadas nos modelos.

Com o modelo PLS utilizando 6 VL foi possível descrever praticamente toda a variabilidade tanto do bloco – Y como do bloco – X, utilizados na calibração. A Tabela 8.2 traz os valores de variância capturados por cada VL, estes valores mostram que a estrutura dominante dos dados, tanto para o bloco-X quanto para o bloco-Y, é descrita basicamente pela primeira VL.

Tabela 8.2. Porcentagem de variância capturada para cada variável latente do modelo PLS.

VL*	Bloco-X		Bloco-Y	
	Variância (%)	Variância Acumulada (%)	Variância (%)	Variância Acumulada (%)
1	99,99	99,99	98,71	98,71
2	0,01	99,99	0,12	98,83
3	0,00	99,99	0,16	98,99
4	0,00	99,99	0,46	99,45
5	0,00	100	0,19	99,64
6	0,00	100	0,12	99,76

*VL = Variável latente.

Os resultados da calibração podem ser visualizados na Figura 8.4, que exhibe os valores experimentais versus os valores preditos pelo modelo PLS com 6 VLs. Os resíduos da calibração obedecem a uma distribuição normal como se pode ver pelo histograma apresentado na Figura 8.5. Os parâmetros de erros (g/100g) e ajuste relativos ao conjunto de calibração encontrados foram os seguintes: PRESS = 16,41; MSE = 0,197; RMSE = 0,444; REV = 0,826 e $r = 0,907$.

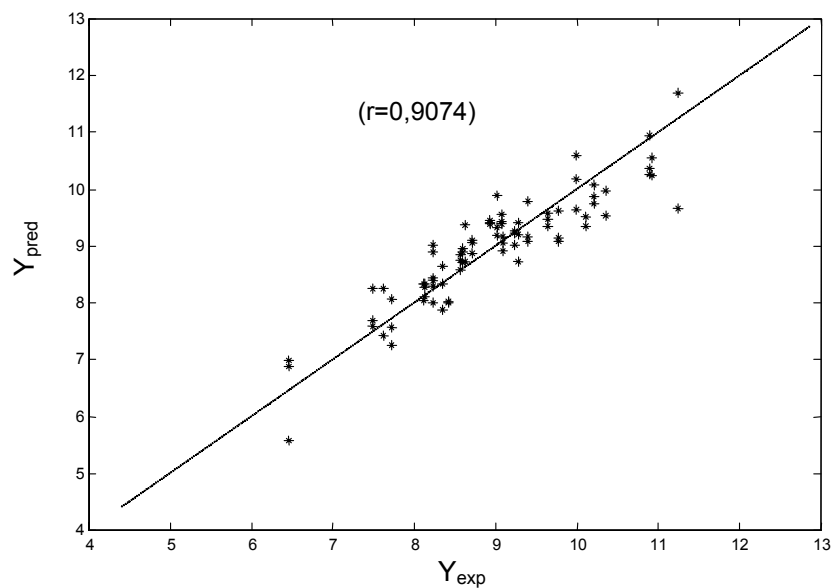


Figura 8.4. Valor experimental (Y_{exp}) versus valor predito (Y_{pred}) (em g de açúcares totais/100g) para o conjunto de calibração do modelo de regressão usando 6 variáveis latentes.

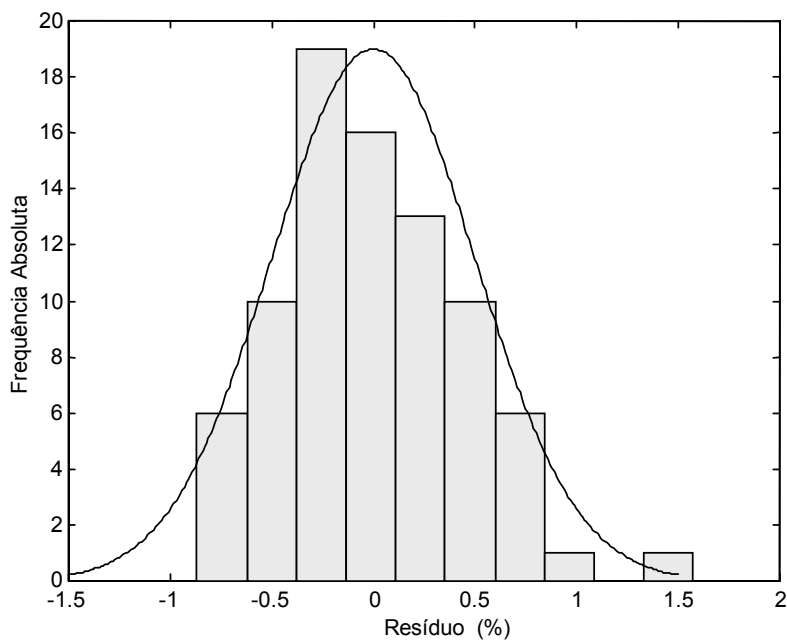


Figura 8.5. Histograma dos resíduos da calibração (%) do modelo PLS com 6 variáveis latentes.

Validação do modelo: O modelo de calibração para a predição do teor de açúcar total foi validado por validação externa, utilizando um conjunto 66 espectros, relativos a 22 amostras tomadas em triplicata. A Tabela 8.3 apresenta os valores médios preditos pelo modelo, o resíduo e o erro percentual relativo para as amostras do conjunto de validação externa. O histograma para estes dados, mostrado na Figura 8.6, indica que o erro percentual para este conjunto fica na maioria dos casos confinado entre os valores de ± 5 %. O erro máximo obtido para o teor de açúcar total foi 7,7 % para as 22 amostras do conjunto de validação, o menor erro foi 0 % e a média 3,4 %. Os parâmetros de erro obtidos foram: PRESS = 3,354; MSE = 0,152; RMSE = 0,390; $r = 0,645$. Embora o coeficiente de correlação linear (r) para este conjunto seja relativamente baixo, os resultados obtidos para o modelo superaram as expectativas estipuladas arbitrariamente em trabalhar com desvios inferiores a 10%.

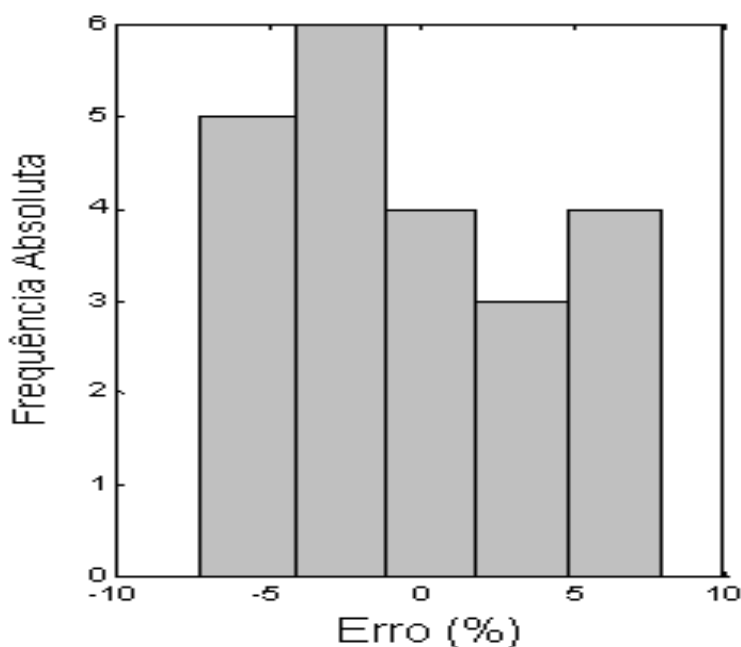


Figura 8.6. Representação gráfica da frequência do erro percentual para o conjunto de validação externa.

Tabela 8.3. Resultados de predição do teor de açúcares totais para as amostras do conjunto de validação.

Resultados obtidos pelo método referência (g/100g)	Resultados obtidos pelo método NIR/PLS (g/100g)	Resíduo (g/100g)	Erro (%)
8,5 ± 0,1	8,6 ± 0,8	-0,1	-1,2
8,9 ± 0,2	8,8 ± 0,4	0,1	1,1
9,50 ± 0,09	9,0 ± 0,6	0,5	5,3
9,6 ± 0,4	9,0 ± 0,5	0,6	6,2
9,0 ± 0,2	9 ± 1	0	0
8,9 ± 0,2	9,3 ± 0,5	-0,4	- 4,5
9 ± 1	8,8 ± 0,5	0,2	2,2
9,2 ± 0,5	9,7 ± 0,2	-0,5	- 5,4
9,6 ± 0,1	9,4 ± 0,5	0,2	2,1
9,2 ± 0,3	9,0 ± 0,3	0,2	2,2
9,3 ± 0,3	8,9 ± 0,6	0,4	4,3
9,2 ± 0,1	9,7 ± 0,9	-0,5	-5,4
9 ± 1	9,2 ± 0,4	-0,2	-2,2
9 ± 2	9,1 ± 0,3	-0,1	-1,1
10,0 ± 0,1	9,5 ± 0,1	0,5	5,0
10 ± 1	9,8 ± 0,4	0,2	2,0
8,1 ± 0,2	8,7 ± 0,3	-0,6	-7,4
9,4 ± 0,1	8,9 ± 0,5	0,5	5,3
9,0 ± 0,4	9,1 ± 0,7	-0,1	-1,1
9,2 ± 0,7	9,3 ± 0,7	-0,1	-1,1
9,4 ± 0,2	9,1 ± 0,5	0,3	3,2
10,4 ± 0,2	9,6 ± 0,7	0,8	7,7

Resíduo: valor obtido pelo método de referência – valor obtido pelo método NIR/PLS;
 erro (%) = (resíduo/ valor obtido pelo método de referência) x 100.

8.3.2. PROTEÍNAS

Para a construção dos modelos de regressão foram utilizados os valores médios do teor de proteína obtido na determinação empregando o método de Kjeldhal e os 159 espectros das amostras de café, sendo que 87 espectros, correspondentes a 29 amostras, foram utilizados na etapa de calibração e os 72 espectros restantes, correspondentes a 24 amostras, foram usados na validação externa dos modelos.

Aos espectros NIR foram aplicados a transformação Kubelka-Munk; a correção multiplicativa de sinal; o alisamento com SPLINE; a derivada primeira e a média móvel com o objetivo de reduzir o número de variáveis. As absorvâncias referentes a cada comprimento de onda foram escaladas pela variância.

Para a determinação do número de variáveis latentes (VL) e do tamanho da janela móvel foram testados modelos construídos com uma até dez variáveis latentes. Para cada um destes modelos variou-se o tamanho da janela móvel de 2 até 40. Estes resultados estão apresentados na Figura 8.7. Observa-se que o PRESS de validação interna diminui com o aumento da complexidade do modelo, ou seja, com o aumento do número de variáveis latentes utilizado. Este comportamento foi observado também em relação aos modelos construídos, utilizando um número de variáveis latentes inferior a 3 e superior a 5. Esta tendência, apresentada no PRESS de validação interna, não é observada com relação ao PRESS do conjunto de validação externa.

O modelo que apresentou o menor PRESS de validação foi escolhido. Este modelo foi construído com 4 VL e utiliza no cálculo da média móvel uma janela de tamanho 20. O gráfico do PRESS do conjunto de validação externa é apresentado na Figura 8.8 e confirma o uso de 4 VL. O uso de uma janela com estas dimensões permitiu reduzir o número de variáveis de 2250 (que são os comprimentos de onda lidos) para 69 variáveis. Destas 69 variáveis foram selecionadas 62 para compor o modelo final. Esta seleção foi feita a partir do valor dos *leverage* [Sharaf *et al.*, 1986] das variáveis. Foram descartadas as variáveis que apresentaram um valor de *leverage* inferior a 0,023 (valor de corte). Como *leverage* em relação às variáveis está intimamente relacionado com o poder de modelagem que estas apresentam, a seleção foi feita descartando as variáveis que apresentavam um baixo valor de *leverage*, ou seja, as variáveis que contribuíam pouco na construção do modelo.

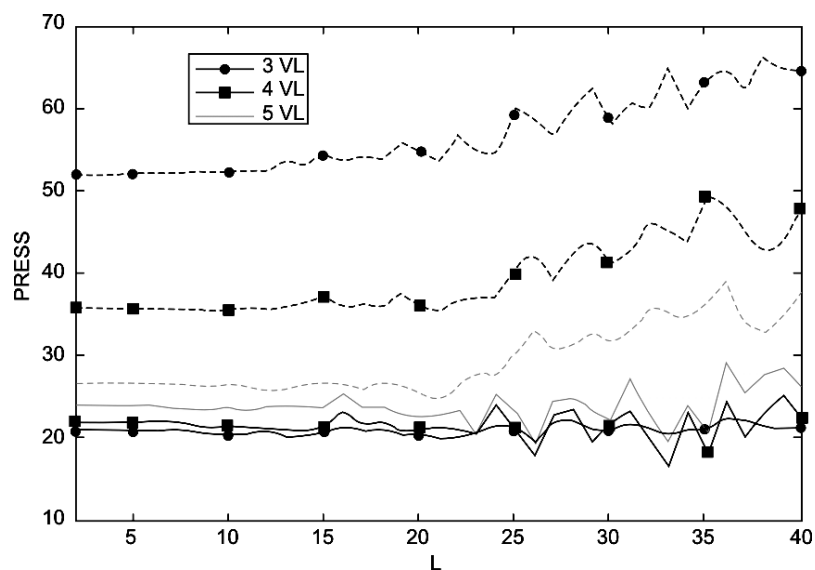


Figura 8.7. Representação gráfica do PRESS de validação interna (- -) e de validação externa (—) em função do tamanho da janela (L), para modelos com 3, 4 e 5 variáveis latentes (VL).

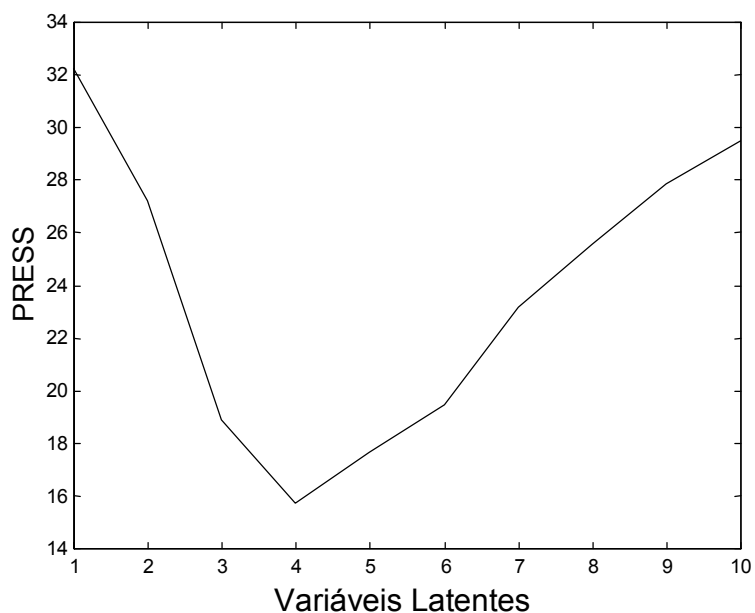


Figura 8.8. Representação gráfica do PRESS do conjunto de validação externa em função do número de VL consideradas nos modelos.

Segundo Nousiainen *et al.* [2004], Soares *et al.* [1998] e Büning-Pfaue [2003] as regiões dos espectros no infravermelho próximo para estudo de proteínas estão associadas principalmente às ligações N-H de amins (9804, 6667, 6536, 5051,

4938, 4690, 4395, 3180, 3330 cm^{-1}) e ligações C=O (4926, 4730, 4587, 4359 cm^{-1}). As regiões espectrais selecionadas neste trabalho para compor os modelos de regressão correspondem principalmente às bandas de combinação (estiramento + deformação angular) dos grupos N-H da proteína e estão compreendidas entre 6623 e 5051 cm^{-1} .

A Tabela 8.4 mostra os parâmetros de erro e ajuste para os modelos com e sem seleção de variáveis. Os menores valores de erro e um pequeno aumento no valor do coeficiente de correlação linear (r) apresentados no conjunto de calibração para o modelo II mostram que com a seleção de variáveis foi possível obter um modelo ligeiramente mais ajustado. O valor superior do parâmetro REV evidencia um melhor conjunto de variáveis com melhor capacidade de predição para este modelo. A habilidade de predição do teor de proteína do modelo II pode ser constatada nos baixos valores de erro no conjunto de validação externa, quando comparados com os valores do modelo I.

Tabela 8.4. Parâmetros de ajuste e erro dos modelos obtidos utilizando 4 variáveis latentes (g/100g) na determinação do teor de proteína.

Modelo PLS	Conjunto de calibração					Conjunto de validação externa			
	PRESS	MSE	RMSE	REV	r	PRESS	MSE	RMSE	r
I (69 variáveis)	19,836	0,225	0,475	0,481	0,690	15,984	0,222	0,471	0,626
II (62 variáveis)	19,316	0,219	0,469	0,495	0,701	15,706	0,218	0,467	0,633

Press = soma dos erros quadráticos médios; MSE = erro quadrático médio; RMSE = raiz quadrado do erro quadrático médio; REV = variância relativa; r = coeficiente de correlação.

Com a utilização de 4 variáveis latentes no modelo PLS foi possível descrever praticamente toda a variação dos dados, tanto do bloco X quanto do bloco Y, conforme pode ser observado na Tabela 8.5.

A Figura 8.9-A mostra o gráfico dos valores experimentais versus valores preditos para as amostras do conjunto de calibração. A distribuição normal dos resíduos de calibração (Figura 8.9-B) é um indicativo de que o modelo é robusto e

apresenta um bom ajuste. Apesar dos baixos valores de erro, observa-se no modelo final uma tendência de superestimar o teor de proteína para as amostras de baixo teor protéico e subestimar os valores de teor de proteína para as amostras de alto teor protéico.

Tabela 8.5. Porcentagem de variância capturada para cada variável latente do modelo PLS.

VL	Bloco-X		Bloco-Y	
	Variância (%)	Variância acumulada (%)	Variância (%)	Variância acumulada (%)
1	98,75	98,75	99,64	99,64
2	0,20	98,94	0,15	99,79
3	0,20	99,15	0,05	99,83
4	0,06	99,20	0,04	99,88

VL = variável latente

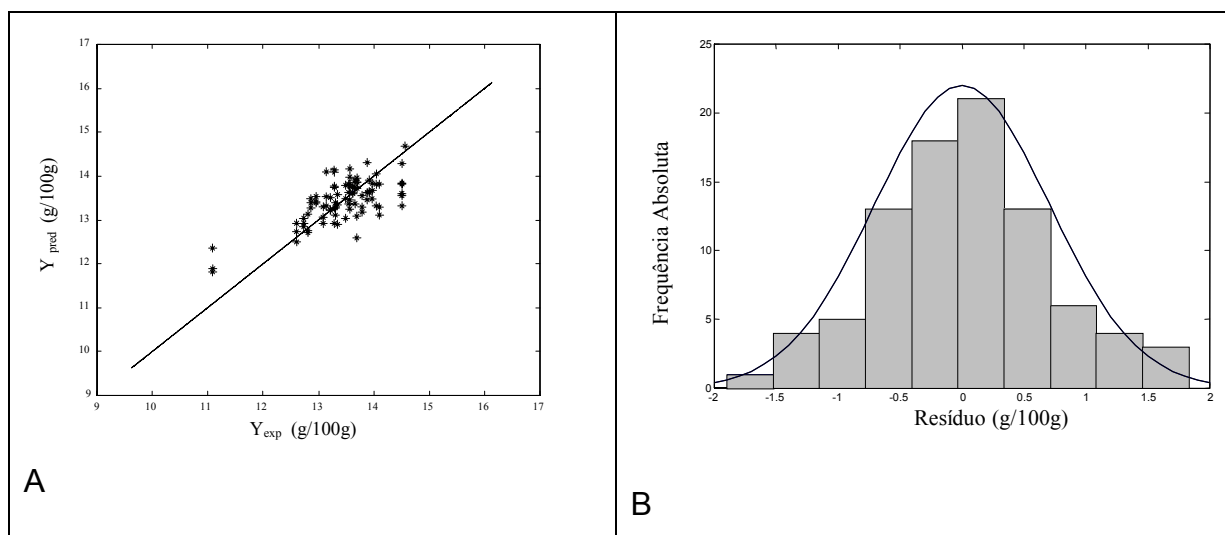


Figura 8.9. **A** - Valores experimentais (Y_{exp}) *versus* valores preditos (Y_{pred}) do conjunto de calibração para o modelo usando 4 variáveis latentes. **B** - Histograma dos resíduos de calibração.

Validação do modelo. A validação do modelo contou com um conjunto de 24 amostras tomadas em triplicatas. A Figura 8.10-A mostra os valores experimentais *versus* valores preditos pelo modelo PLS (modelo II) para o conjunto de validação externa.

Na Figura 8.10-B, é mostrado o gráfico dos valores experimentais versus resíduos. Pode-se observar que a mesma tendência apresentada no conjunto de calibração foi encontrada também no conjunto de validação externa, em que as amostras que apresentam alto teor de proteína têm os seus valores subestimados pelo modelo. Já as amostras que apresentam baixo teor de proteína têm os seus valores superestimados.

Os parâmetros de erro (g/100g) para as amostras do conjunto de validação externa foram: PRESS = 4,611; MSE = 0,192; RMSE = 0,438; ER (%) = 2,864 e $r = 0,694$. Os baixos valores dos erros apresentados neste conjunto revelam uma boa habilidade de predição do modelo.

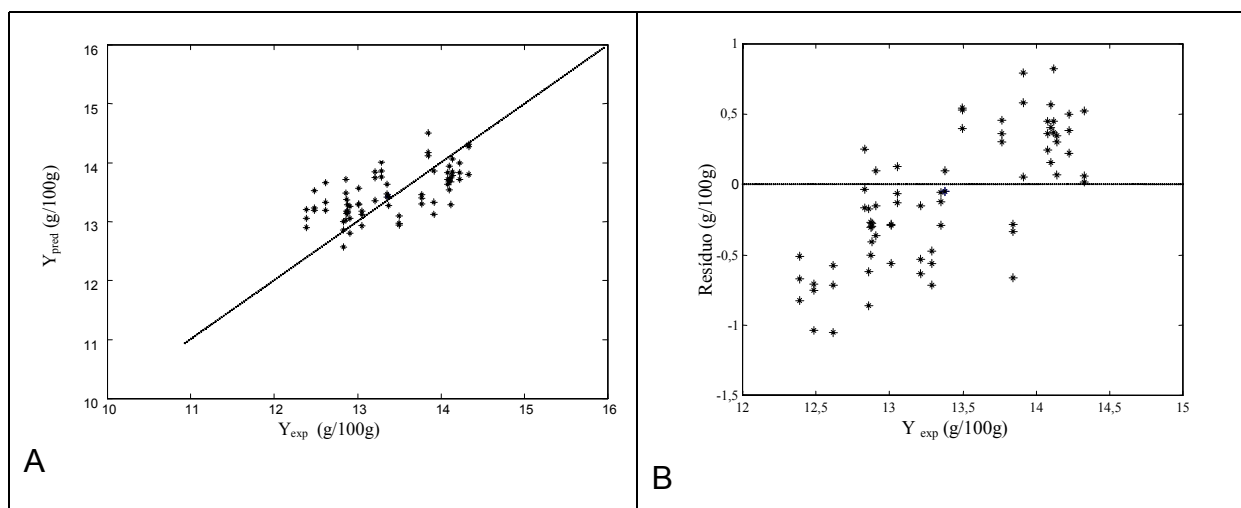


Figura 8.10. A- Curva de calibração dos valores experimentais (Y_{exp}) versus valores preditos (Y_{pred}) para o conjunto de validação externa usando 4 variáveis latentes. **B-** Gráfico dos valores experimentais (Y_{exp}) versus resíduos para o conjunto de validação externa.

A Tabela 8.6 apresenta, para fins comparativos, os valores experimentais, os valores médios preditos, os resíduos e o erro percentual para as amostras do conjunto de validação externa. A maior discordância entre o valor experimental e o valor predito pelo modelo encontrado neste conjunto foi de 6,4 %, a menor foi de 0% e a média para as 24 amostras foi de 2,8 %. A melhor habilidade de predição do

modelo também pode ser avaliada pelo baixo valor do erro relativo percentual apresentado pelo conjunto de validação externa.

Tabela 8.6. Comparação dos resultados de proteína total (g/100g) para as amostras do conjunto de validação externa obtidos pelos métodos de Kjeldhal e NIRS / PLS.

Valor experimental (Método de Kjeldhal)	Valor médio predito (Método NIRS-PLS)	Resíduo	Erro (%)
13,84 ± 0,06	14,3 ± 0,4	-0,5	- 3,6
14,3 ± 0,2	14,1 ± 0,6	0,2	1,4
14,1 ± 0,1	13,7 ± 0,2	0,4	2,8
13,29 ± 0,06	13,9 ± 0,2	-0,6	- 4,5
14,1 ± 0,1	13,7 ± 0,4	0,4	2,8
13,35 ± 0,07	13,5 ± 0,2	-0,1	- 0,7
12,61 ± 0,07	13,4 ± 0,5	-0,8	- 6,3
12,9 ± 0,1	13,4 ± 0,7	-0,5	- 3,9
13,77 ± 0,05	13,4 ± 0,1	0,4	2,9
14,22 ± 0,09	13,8 ± 0,3	0,4	2,8
14,14 ± 0,09	13,9 ± 0,3	0,2	1,4
12,9 ± 0,2	13,0 ± 0,5	-0,1	- 0,8
12,9 ± 0,4	13,2 ± 0,1	-0,3	- 2,3
13,2 ± 0,1	13,6 ± 0,5	-0,4	- 3,0
13,4 ± 0,2	13,4 ± 0,2	0,0	0,0
12,5 ± 0,2	13,3 ± 0,3	-0,8	- 6,4
14,1 ± 0,3	13,6 ± 0,5	0,5	3,5
13,9 ± 0,2	13,4 ± 0,8	0,5	3,6
13,5 ± 0,2	13,0 ± 0,2	0,5	3,7
13,01 ± 0,09	13,4 ± 0,3	-0,4	- 3,1
13,06 ± 0,06	13,1 ± 0,3	-0,0	- 0,0
12,8 ± 0,2	12,8 ± 0,4	0,0	0,0
12,9 ± 0,1	13,2 ± 0,2	-0,3	- 2,3
12,4 ± 0,1	13,0 ± 0,3	-0,6	- 4,8

Resíduo: valor obtido pelo método de referência – valor obtido pelo método NIR/PLS; erro (%) = (resíduo/ valor obtido pelo método de referência) x 100.

A precisão das medidas em triplicata obtida através do modelo PLS foi comparada com a precisão das medidas do teor de proteína para os métodos de referência. Para comparar a precisão entre estes dois conjuntos de dados foi aplicado o Test F (INMETRO, 2003). Excetuando uma medida, todas as outras medidas mostram que ao nível de 95% de probabilidade não há diferença

significativa entre a precisão dos dados obtidos pelo modelo e a precisão dos dados obtidos a partir do método de referência.

8.3.3. UMIDADE

O modelo PLS para a quantificação do teor de umidade foi construído utilizando 85 espectros no conjunto de calibração e validado por validação externa utilizando um conjunto de 72 espectros. As seguintes transformações matemáticas foram usadas: transformação Kubelka-Munk e correção multiplicativa de sinal. Como os espectros eram muito ruidosos, os mesmos foram alisados, utilizando SPLINE e aplicação de média móvel.

Dos modelos testados os que apresentaram melhor performance de predição foram os modelos construídos usando 7 variáveis latentes e tamanho de janela 15 (Figura 8.11).

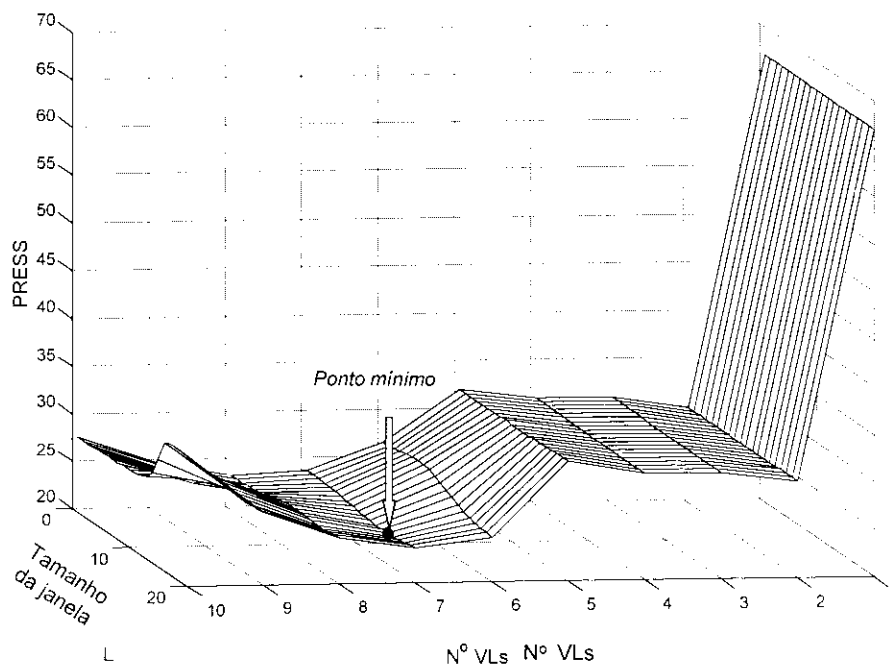


Figura 8.11. Comportamento do PRESS x número de variáveis latentes (Nº VL) x tamanho da janela da média móvel.

Usando as informações do correlograma, apresentado na Figura 8.12, foi possível reduzir o número de variáveis do modelo para 15 (número de variáveis selecionadas no correlograma) e o número de VL também reduziu de 7 para 3. Estas reduções no número de variáveis gerou um modelo com um menor valor de MSE para o conjunto de validação externa, ou seja, uma pequena melhora de habilidade de predição, como mostra os parâmetros de erro e ajustes presentes na Tabela 8.7, porém o coeficiente de correlação teve uma leve redução.

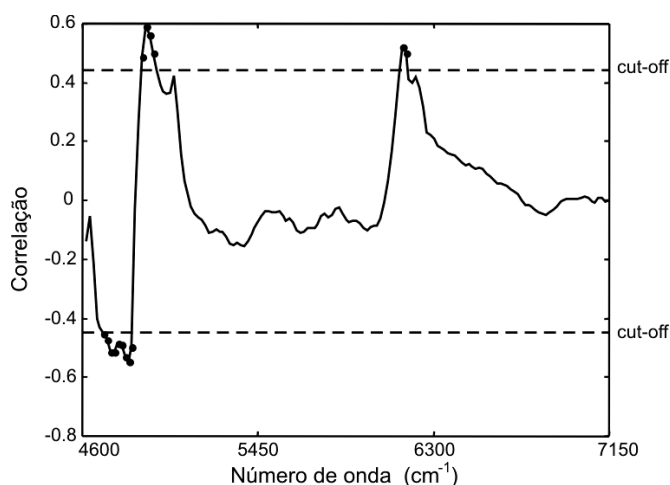


Figura 8.12. Correlograma dos dados mostrando o valor de corte utilizado (*cut-off*) e as variáveis selecionadas (em preto) (nº onda) na determinação da umidade.

Tabela 8.7. Parâmetros de ajuste e erro dos modelos PLS (g/100g) sem seleção (modelo I) e com seleção de variáveis (modelo II) usados na construção dos modelos para a determinação de umidade.

Modelo*	Conjunto de calibração					Conjunto de validação externa			
	PRESS	MSE	RMSE	REV	r	PRESS	MSE	RMSE	r
I	33,240	0,391	0,625	0,665	0,813	21,785	0,303	0,550	0,821
II	48,323	0,569	0,754	0,513	0,712	21,495	0,298	0,546	0,818

*Modelo I: Modelo com 7 VL e 150 variáveis decorrentes da filtragem dos espectros transformados utilizando média móvel com uma janela de tamanho 15.

Modelo II: Modelo com 3 VL e 15 variáveis escolhidas pelo correlograma a partir dos espectros transformados e filtrados por média móvel com uma janela de tamanho 15.

Press = soma dos erros quadráticos médios; MSE = erro quadrático médio; RMSE = raiz quadrada do erro quadrático médio; REV = variância relativa; r = coeficiente de correlação.

A utilização de 3 VLs no modelo II é suficiente para capturar praticamente toda estrutura do bloco-X (variáveis espectrais transformadas) e do bloco-Y (teor de umidade obtido pelo método de referência) utilizados na construção do modelo, como pode-se ver pelos dados apresentados na Tabela 8.8.

Tabela 8.8. Porcentagem de variância capturada para cada variável latente (VL) do modelo PLS.

VL	Bloco-X		Bloco-Y	
	Variância (%)	Variância acumulada (%)	Variância (%)	Variância acumulada (%)
1	99,99	99,99	98,81	98,81
2	0,01	100	0,49	99,30
3	0,01	100	0,13	99,44

A água absorve fortemente no infravermelho e como a maioria dos alimentos possui água na sua constituição, uma forte banda identifica a presença de umidade; amostras secas possuem bandas menos intensas. Em amostras de alimentos, as bandas de interesse nos espectros podem sofrer algumas variações dependendo do alimento, devido ao diferente teor de umidade presente em cada tipo de alimento e a forma como a água se apresenta ligada aos diferentes componentes dos alimentos [Villamarín *et al.*, 2002; Watson, 1977].

No trabalho de Büning-Pfaue [2003] a região compreendida entre 5159 a 4990 cm^{-1} foi usada para a determinação de umidade em diferentes matrizes de alimentos e foi atribuída às ligações do tipo O-H e ligações de H. As regiões espectrais selecionadas neste trabalho para a construção dos modelos de regressão estão relacionadas às bandas de combinação (estiramento + deformação angular) das ligações O-H e ligações C-O compreendidas entre 4650 a 4900 e 5004 a 5062; e região entre 6170 a 6180 que correspondem as ligações C-H.

Os resultados da previsão do teor de umidade pelo modelo PLS (modelo II) para as amostras do conjunto de calibração podem ser vistos no gráfico mostrado na Figura 8.13.

Os resultados da previsão do modelo PLS para o teor de umidade para as amostras da validação externa comparados com os valores do teor de umidade do método de referência estão presentes no gráfico da Figura 8.14. Os parâmetros de erro (g/100g) para as amostras do conjunto de validação externa foram: Press = 21,495; MSE = 0,298; RMSE = 4,504 e $r = 0,818$. As 24 amostras do conjunto de validação externa foram tomadas em triplicata.

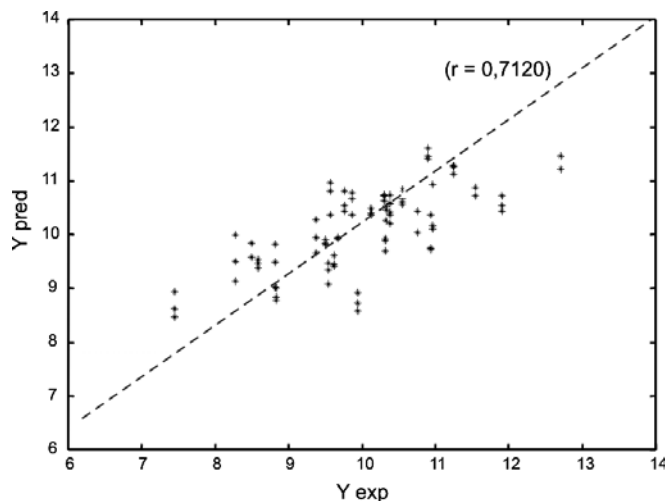


Figura 8.13. Valores experimentais (Y_{exp}) versus valores preditos (Y_{pred}) da determinação de umidade para o conjunto de calibração externa usando 3 variáveis latentes.

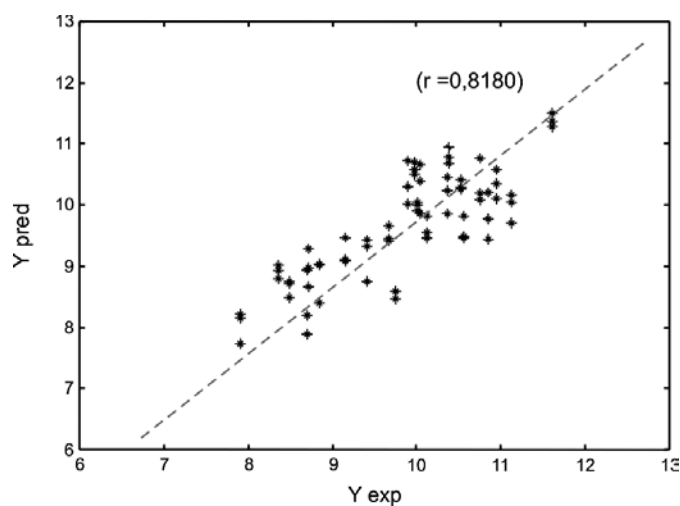


Figura 8.14. Valor experimental versus valor predito para as amostras do conjunto de validação externa.

A Tabela 8.9 apresenta os valores experimentais e os valores preditos pelo modelo PLS para as 24 amostras do conjunto de validação externa. A maior discordância entre o valor experimental e o valor predito pelo modelo encontrado neste conjunto foi de 10,2 %, a menor foi de 0 % e a média para as 24 amostras foi de 4,1 %.

Tabela 8.9. Comparação dos resultados do teor de umidade (g/100g) para as amostras do conjunto de validação externa obtidos pelos métodos de referência e NIRS – PLS, aplicando o modelo de regressão de melhor ajuste obtido.

Valor obtido pelo método de referência	Valor obtido pelo método NIR/PLS	Resíduo	Erro (%)
10,36 ± 0,03	10,3 ± 0,7	0,1	0,6
9,75 ± 0,01	8,7 ± 0,3	1,0	10,2
10,05 ± 0,05	10,4 ± 0,5	- 0,3	- 3,0
9,98 ± 0,05	10,7 ± 0,2	- 0,7	- 7,0
10,52 ± 0,01	10,4 ± 0,2	0,1	0,9
10,01 ± 0,03	10,1 ± 0,2	- 0,1	1,0
11,1 ± 0,2	10,1 ± 0,2	1,0	9,0
10,55 ± 0,01	9,7 ± 0,3	0,8	7,6
9,41 ± 0,01	9,3 ± 0,3	0,1	1,1
8,8 ± 0,1	8,9 ± 0,3	- 0,1	- 1,1
10,4 ± 0,2	10,9 ± 0,2	- 0,5	- 4,8
9,90 ± 0,06	10,5 ± 0,4	0,6	6,1
9,67 ± 0,02	9,7 ± 0,5	0	0
8,48 ± 0,02	8,9 ± 0,4	- 0,4	4,7
8,69 ± 0,06	8,6 ± 0,6	0,1	1,1
7,91 ± 0,06	8,3 ± 0,4	- 0,4	- 5,0
8,7 ± 0,1	9,2 ± 0,3	- 0,5	- 5,7
8,37 ± 0,08	9,1 ± 0,3	- 0,7	- 8,4
11,60 ± 0,05	11,50 ± 0,08	0,10	0,09
10,95 ± 0,06	10,5 ± 0,3	0,4	3,6
10,12 ± 0,04	9,82 ± 0,09	0,30	3,0
10,8 ± 0,2	10,0 ± 0,5	0,8	7,4
10,76 ± 0,03	10,5 ± 0,4	0,3	2,8
9,15 ± 0,02	9,45 ± 0,02	- 0,3	- 3,3

Resíduo: valor obtido pelo método de referência – valor obtido pelo método NIR/PLS; erro (%) = (resíduo/valor obtido pelo método de referência) x 100.

8.4. CONCLUSÃO

As principais conclusões obtidas nesta etapa do trabalho foram:

- A partir de modelos de regressão PLS pode-se inferir o teor de proteína, açúcar total e umidade em amostras de grãos de café cru com os seguintes erros: MSE ($\text{g } 100\text{g}^{-1}$): 0,22; 0,15 e 0,30; r: 0,63; 0,64 e 0,82 e erro relativo médio para as amostras do conjunto de validação (%): 2,8; 3,4 e 4,1, respectivamente.
- O alisamento dos espectros utilizando SPLINE possibilita, exceto para a determinação do teor de açúcar total, a criação de modelos com uma maior habilidade de predição.
- A filtragem dos espectros utilizando média-móvel e a seleção de variáveis via correlograma e *leverage* das variáveis são importantes na construção dos modelos.
- O método proposto reduz o tempo requerido de análise, pois em cerca de 1 minuto se obtém um espectro da amostra e, também, proporciona a mínima manipulação da amostra, sendo necessário apenas a homogeneização em moinho.
- Possibilita a análise simultânea de várias propriedades do café.

9. MINERAIS EM SUCOS DE FRUTAS

Nesta etapa do trabalho, dois métodos de preparação de amostra para determinação de Ca, P, Na, K, Mg, Zn, Fe, Mn e Cu em suco de frutas foram comparados: extração com ácido clorídrico a frio com agitação e digestão em sistema fechado, sob pressão, assistida por microondas. As concentrações dos elementos minerais foram obtidas empregando-se a técnica de espectrometria de emissão óptica com fonte de plasma indutivamente acoplado (ICP OES). A precisão e exatidão dos métodos foram avaliadas usando uma amostra de suco de uva. A metodologia de extração com ácido clorídrico foi empregada para a determinação dos teores de minerais em 7 diferentes variedades de sucos processados de abacaxi, acerola, caju, goiaba, manga, maracujá e uva. Para facilitar a interpretação dos teores dos minerais encontrados, foram utilizados os métodos de análise por componentes principais (PCA) e análise de agrupamento por método hierárquico (HCA). Foi, também, avaliada a contribuição nutricional dos sucos para a dieta de adultos baseando-se nos valores de ingestão diária recomendada para os minerais estudados.

9.1. ESPECTROMETRIA ÓPTICA EM PLASMA INDUTIVAMENTE ACOPLADO

A investigação da composição mineral dos alimentos requer métodos de análises capazes de avaliar teores maiores, menores e traços de uma grande variedade de elementos em diferentes matrizes de alimentos. Uma boa técnica de análise de elementos metálicos deve ter elevada sensibilidade, tolerância a interferentes ou fácil correção das mesmas e a preparação da amostra deve ser realizada com o mínimo manuseio e mínimo uso de reagentes possíveis [López *et al.*, 1996].

Para a determinação de elementos minerais em alimentos, a técnica empregando a espectrometria de emissão óptica em plasma indutivamente acoplado

(ICP OES) vem sendo cada vez mais utilizada [Bou *et al.*, 2005; Gupta *et al.*, 2003; Ozcan, 2005 e Pennington e Young, 1990]. Uma das vantagens desta técnica é a baixa interferência entre elementos, devido a altas temperaturas usadas. Bons espectros de emissão são obtidos em um único conjunto de excitação; como resultado, espectros de diversos elementos podem ser registrados simultaneamente. Outra vantagem é que as fontes de plasma são mais energéticas e permitem a determinação em baixas concentrações de elementos refratários, como óxidos de boro, fósforo e zircônio, entre outros. Finalmente, os métodos baseados em fontes de plasma oferecem amplas faixas de concentração linear [Giné, 1998].

Na técnica de ICP OES a amostra é submetida a elevadas temperaturas, variando entre 10000 a 6000 K, e os seus componentes são convertidos a átomos gasosos ou íons elementares, podendo ocorrer, também, excitação das espécies a estados eletrônicos mais elevados. A relaxação das espécies excitadas é acompanhada pela produção de linhas espectrais que são usadas nas análises qualitativas bem como nas análises quantitativas. As altas temperaturas podem ser conseguidas por um plasma de argônio. O plasma é um gás parcialmente ionizado, onde a carga negativa total das partículas é aproximadamente igual à soma das cargas positivas e, portanto, eletricamente neutro e um bom condutor de corrente elétrica. O plasma de argônio é formado por um fluxo de gás argônio em um campo magnético de alta frequência que é criado em uma tocha de quartzo, formada de três tubos concêntricos. A vantagem do uso de um plasma como fonte de excitação, quando comparado com uma chama, é a obtenção de temperaturas mais altas e formação de ambientes quimicamente menos reativos [Skoog *et al.*, 1998].

Um espectrômetro de emissão óptica em plasma indutivamente acoplado constitui-se basicamente de um gerador de rádio-frequência, um sistema para a introdução da amostra, uma tocha, um sistema de gás argônio, um sistema de detecção do sinal analítico e um sistema computacional para o controle do equipamento [Boss e Fredeen, 1999].

O gerador de rádio-frequência (RF) é um dispositivo elétrico empregado como fonte de potência e tem como função sustentar o plasma [Skoog *et al.*, 1998].

O sistema de nebulização é composto, em geral, por um nebulizador e uma câmara de nebulização. O nebulizador produz um aerossol da amostra que é conduzido ao plasma pela câmara, a qual favorece a introdução apenas de gotas de menor volume ao plasma. Os tipos mais comuns de nebulizadores são os pneumáticos (concêntricos, *cross-flow* e *cone spray*) e os ultra-sônicos. O sistema ideal de introdução de amostra é aquele que distribui toda a amostra no plasma, de forma que este possa reproduzir a dessolvatação, vaporização, atomização, ionização e excitação [Boss e Fredeen, 1999].

Os nebulizadores pneumáticos são os mais usados devido à sua maior simplicidade de construção, robustez, boa reprodutibilidade e baixo custo. Normalmente emprega-se uma bomba peristáltica para a introdução da amostra, mantendo-se assim constante a taxa de aspiração da amostra para uma mesma matriz. As principais desvantagens deste tipo de nebulizador são: pode ocorrer variação na taxa de aspiração da amostra pela viscosidade da amostra e bloqueios provocado por sólidos em suspensão.

A tocha é um dispositivo de quartzo onde se forma o plasma. É constituída por três tubos concêntricos por onde passam os fluxos do argônio principal, auxiliar e nebulizador. O fluxo do argônio principal é responsável pela manutenção do plasma e proteção das paredes da tocha contra a fusão; o fluxo nebulizador introduz a amostra no plasma e o fluxo auxiliar tem a função de direcionar o aerossol da amostra para dentro do plasma [Skoog *et al.*, 1998].

A parte mais complexa do espectrômetro de emissão é o sistema óptico. Existem dois tipos de sistemas: o de medida de sinal seqüencial e o simultâneo, sendo que os seqüenciais embora sejam mais lentos, são mais convenientemente empregados para análises de amostras de composição variada. As linhas espectrais são selecionadas, dependendo dos elementos a serem determinados, da composição e impurezas da amostra. A observação do sinal pode ser feita em duas formas: na configuração radial ou axial da tocha. Na configuração radial são captadas as emissões que ocorrem no raio da tocha, em uma dada altura de observação, enquanto as emissões que ocorrem no eixo da tocha são captadas na configuração axial. Na configuração axial toda a região ao longo do eixo do plasma é

observada e obtém-se, assim, maiores sensibilidades e melhores limites de detecção em relação à configuração radial. Porém, essa configuração é susceptível a interferências devido aos processos de auto-absorção e recombinação que podem ocorrer na extremidade do plasma [Giné, 1998].

As maiores vantagens do uso da espectrometria de emissão em plasma indutivamente acoplado aplicada à análise de alimentos é devido à sua capacidade multielementar, sua rapidez analítica, sua facilidade de operação, curvas analíticas lineares podem ser obtidas até seis ordens de grandeza, bons limites de detecção podem ser conseguidos, na faixa de ng L^{-1} e excelente precisão e exatidão.

9.2. PARTE EXPERIMENTAL

9.2.1. AMOSTRAS

Amostras de 7 diferentes variedades de sucos de frutas foram obtidas em duas datas diferentes de processamento e uma amostra de suco de uva foi usada para a avaliação das metodologias de preparação das amostras. Foram analisados sucos de frutas de: abacaxi (11° Brix), acerola (6° Brix), caju (11° Brix), goiaba (6° Brix), manga (11° Brix), maracujá (11° Brix) e uva (15° Brix).

As determinações dos minerais cálcio, cobre, sódio, manganês, magnésio, ferro, potássio, fósforo e zinco foram realizadas com 5 repetições analíticas na avaliação do método (suco de uva) e, após a escolha do método de preparação, as determinações foram feitas com 3 repetições analíticas.

9.2.2. MÉTODO PARA PREPARAÇÃO DAS AMOSTRAS

Para a avaliação do método de preparo de amostra, foram realizadas as determinações dos minerais cálcio, fósforo, manganês, potássio, cobre, magnésio,

ferro, sódio e zinco em uma amostra de suco de uva, usando dois procedimentos diferentes, conforme descrito a seguir:

– **Extração dos minerais com HCl a frio e agitação:** Uma alíquota de 5,00 mL de suco foi transferida para um balão volumétrico de 25,00 mL, seguido de adição de 5,0 mL de ácido clorídrico concentrado, grau analítico. A mistura foi agitada por 3 horas com um agitador mecânico orbital e, posteriormente, completou-se o volume do balão com água bidestilada. Dois brancos foram preparados da mesma forma.

– **Digestão ácida em sistema fechado assistido por microondas:** Foi utilizado um digestor de microondas com sistema fechado, modelo MDS 2000 (CEM, Matthews, USA). Utilizaram-se 7 copos sendo cinco amostras e dois brancos. Uma alíquota de 5 mL de suco foi transferida para cada copo de digestão, adicionando-se em seguida 5 mL de ácido nítrico concentrado. Cinco estágios foram aplicados para a digestão: (1) 220 MHz de potência, 20 psi de pressão por 40 minutos; (2) 285 MHz de potência, 40 psi de pressão por 5 minutos; (3) 345 MHz de potência, 60 psi de pressão por 50 minutos; (4) 410 MHz de potência, 60 psi de pressão por 10 minutos; (5) 470 MHz de potência, 90 psi de pressão por 15 minutos. Após o término da digestão, os copos foram abertos e as amostras foram transferidas quantitativamente para balões volumétricos de 25 mL completando-se o volume com ácido clorídrico 20 % (v/v).

9.2.3. INSTRUMENTAÇÃO (ICP OES)

Para a determinação do teor dos minerais, foi utilizado um ICP OES simultâneo, modelo ICP 2000 (BAIRD, Massachusetts, USA). As condições de operação encontram-se na Tabela 9.1. As condições ótimas para determinação multielementar foram estabelecidas utilizando-se o elemento manganês para o canal de alinhamento do espectrômetro de emissão, conforme recomendação do fabricante.

As soluções-estoque foram preparadas a partir de soluções-padrão dos metais (Titrisol, Merck), água bidestilada e ácido nítrico grau analítico, em concentrações de $1000 \mu\text{g mL}^{-1}$. As soluções dos metais de trabalho multielementares foram preparadas em diferentes concentrações dependendo do metal a ser determinado em solução de ácido clorídrico 20% (v/v), por diluição das soluções-estoque. Um branco foi preparado com solução de ácido clorídrico 20% (v/v) em água bidestilada.

Tabela 9.1. Condições de operação do espectrômetro ICP OES otimizadas no laboratório.

Potência do plasma	1,0 kW
Gás refrigerante (Ar)	15 L min^{-1}
Gás auxiliar (Ar)	$1,5 \text{ L min}^{-1}$
Vazão de amostra	2 mL min^{-1}
Altura de observação vertical	19 mm^a
Pressão do nebulizador ^b	3 bar

^a acima da bobina de indução; ^b nebulizador pneumático concêntrico.

Os comprimentos de onda usados para as determinações dos 9 elementos encontram-se na Tabela 9.2. As curvas analíticas definidas para cada elemento foram: fósforo, cálcio, sódio e magnésio de 1 a 50 mg L^{-1} ; ferro, zinco, cobre e manganês de $0,025$ a 3 mg L^{-1} ; potássio de 25 a 500 mg L^{-1} e mostraram-se lineares em toda a faixa de trabalho.

Tabela 9.2. Elementos minerais determinados, comprimento de onda, limite de detecção e parâmetros da curva analítica.

Metal	λ (nm)	LOD ($\mu\text{g L}^{-1}$)	LOQ ($\mu\text{g L}^{-1}$)	A	B	r
Cálcio	317,93	5	16	- 184,421	881,781	0,9999
Magnésio	279,08	58	193	4,623	143,142	1
Fósforo	178,28	31	105	- 3,038	35,840	0,9999
Ferro	259,94	1	5	5,901	3030,593	1
Manganês	257,67	1	2	6,088	2531,797	1
Sódio	589,59	79	262	10,503	4,583	1
Potássio	766,49	52	174	0,689	14,663	1
Zinco	213,86	1	3	15,582	10242,181	0,9999
Cobre	324,75	7	25	1,056	1086,210	1

LOD = limite de detecção (3 vezes o desvio-padrão de 10 leituras em concentração próxima ao primeiro ponto da curva analítica); LOQ = limite de quantificação (10 vezes o desvio-padrão de 10 leituras em concentração próxima ao primeiro ponto da curva analítica); A = coeficiente angular da regressão; B = coeficiente linear da regressão; r = coeficiente de correlação da regressão; número de pontos das curvas analítica = 5.

9.3. RESULTADOS E DISCUSSÃO

Os resultados encontrados para os minerais determinados em suco de uva, usando-se os dois métodos de abertura de amostra descritos anteriormente, podem ser visualizados na Tabela 9.3. Aplicando-se o Test F (INMETRO, 2003), verificou-se que não existe diferença significativa entre os métodos quanto à sua precisão, ao nível de 95% de confiança e 4 graus de liberdade, pois os valores de $F_{4,4}$ calculado (Tabela 9.3) são inferiores ao valor crítico $F_{4,4}$ (tabelado) = 6,4.

A precisão em termos de repetitividade dos métodos empregados foi avaliada através do coeficiente de variação (em %) para as cinco determinações dos minerais, empregando-se o método de extração com HCl e o de digestão em microondas, respectivamente: Ca = 1,2 e 0,9; P = 1,2 e 2,2; Mn = 0,7 e 0,4; K = 1,0 e 1,0; Cu = 1,0 e 0,7; Mg = 1,4 e 0,9; Fe = 1,3 e 0,5; Na = 1,0 e 0,5; Zn = 4,6 e 5,7. O método de extração empregando microondas apresentou melhor precisão nas medidas dos minerais com exceção para P, K e Zn, onde o método de extração com HCl forneceu melhores resultados.

Tabela 9.3. Comparação dos teores dos minerais em suco de uva obtidos pelo método de extração com HCl e agitação e pelo método de digestão usando energia de microondas.

Mineral	Método de extração com HCl e agitação		Método de digestão por microondas		F _{4,4} calculado (Teste F)
	Média* (mg L ⁻¹)	DP (mg L ⁻¹)	Média* (mg L ⁻¹)	DP (mg L ⁻¹)	
Cálcio	131	1	127	1	1
Fósforo	99	1	97	2	4
Manganês	2,32	0,02	2,28	0,01	4
Potássio	333	3	325	3	1
Cobre	3,90	0,04	3,82	0,03	2
Magnésio	96	1	93	1	1
Ferro	1,98	0,02	1,91	0,01	4
Sódio	166	2	165	1	4
Zinco	0,64	0,03	0,65	0,04	2

*5 repetições analíticas. DP = desvio-padrão. F_{4,4} (tabelado) = 6,4.

Para avaliar a exatidão dos métodos de preparo de amostra foi realizada a recuperação dos minerais na amostra de suco de uva ao nível dos minerais encontrados na amostra, ou seja, antes da digestão foram adicionadas quantidades de solução padrão dos metais estudados. Os resultados obtidos na recuperação pelo método de extração com HCl e microondas estão apresentados na Tabela 9.4.

Tabela 9.4. Avaliação da exatidão dos métodos usando cálculo de recuperação para uma amostra fortificada de suco de uva.

Minerais	Adicionado (mg L ⁻¹)	Extração com HCl		Microondas	
		Recuperado (mg L ⁻¹)	% Recuperação	Recuperado (mg L ⁻¹)	% Recuperação
Ca	130	131,3	101	128,7	99
Cu	3	2,85	95	2,28	76
Fe	2	1,84	92	1,74	87
Mg	100	96	96	85	85
Mn	2	1,92	96	1,68	84
Na	150	151,5	101	115,5	77
P	100	107	107	109	109
K	300	279	93	219	73
Zn	0,5	0,475	95	0,545	109

O método de extração com ácido clorídrico apresentou melhores resultados para todos os metais, com níveis de recuperação variando entre 92 e 107 %, enquanto com o método empregando a energia de microondas os níveis de recuperação ficaram entre 73 e 109 %. Os menores níveis de recuperação encontrados para Cu, Mg, Mn, Na, e K podem provavelmente estar associados à adsorção destes metais na parede do frasco de digestão (Teflon PFA), pois, estes já tinham um bom tempo de uso e se apresentavam deformados.

O método de extração com ácido clorídrico foi utilizado para a determinação dos teores dos minerais nas amostras de sucos de frutas devido à sua facilidade de aplicação, rapidez, baixo custo e por fornecer resultados com variações máximas inferiores a 3 mg L⁻¹. Os resultados obtidos podem ser visualizados na Tabela 9.5.

Foi observado que no suco de abacaxi ocorreu variação de 42 % no teor de ferro de um lote para o outro e no suco de maracujá o teor de cálcio apresentou uma variação de 26 %. Nos demais sucos, os minerais apresentaram pequenas variações, sendo que na maioria das vezes os teores dos minerais são similares de um lote para o outro. O suco de abacaxi apresenta teores mais elevados de cálcio, magnésio, manganês e ferro; o suco de maracujá apresenta maior teor de fósforo, potássio e zinco. O elemento cobre aparece em maior concentração no suco de uva. Porém, estes resultados são para os sucos sem considerar as diluições recomendadas pelo fabricante. Foram feitos os cálculos das concentrações dos minerais, seguindo as recomendações de diluições do fabricante para os sucos de fruta e, a seguir, aplicados os métodos de análise por componentes principais e análise de agrupamento por método hierárquico, para melhor visualização dos resultados.

O método de análise por componentes principais (PCA) proporciona um estudo multivariado dos dados experimentais obtidos, facilitando a visualização da correlação entre amostras e variáveis. O método de análise por agrupamento hierárquico (HCA) foi usado com o objetivo de verificar as semelhanças entre as amostras com base nos teores de minerais determinados para os diversos sucos de frutas, através do cálculo de similaridade (calculada utilizando a distância Euclideana entre as amostras).

Foram aplicadas a PCA e a HCA à matriz de dados gerada (42x9) com os resultados dos teores de minerais, considerando a diluição recomendada pelo fabricante (uva, manga e goiaba 1,5 vezes; abacaxi e acerola 2 vezes; maracujá 4,5 vezes; caju 5 vezes) com o objetivo de melhor visualização do teor de minerais dos sucos. O pré-processamento usado foi o auto-escalamento dos dados.

TABELA 9.5. Teores de minerais ($\text{mg } 100\text{mL}^{-1}$) para os dois lotes dos diferentes sucos de frutas estudados (Lotes A e B). Para cada lote os valores correspondem à média de três repetições analíticas (estimativa de desvio-padrão).

Suco	Ca	Mg	P	Fe	Mn	Na	K	Zn	Cu
Caju									
Lote A	0,77 (0,00)	8,44 (0,09)	9,5 (0,1)	0,13 (0,01)	0,077 (0,001)	34,3 (0,4)	102 (1)	0,107 (0,002)	0,036 (0,001)
Lote B	0,64 (0,01)	8,2 (0,1)	8,6 (0,4)	0,125 (0,003)	0,071 (0,002)	33,4 (0,4)	103 (1)	0,102 (0,001)	0,038 (0,004)
Maracujá									
Lote A	5,09 (0,02)	10,90 (0,09)	16,6 (0,3)	0,388 (0,005)	0,088 (0,001)	21,6 (0,5)	222 (1)	0,228 (0,009)	0,061 (0,001)
Lote B	3,79 (0,01)	10,13 (0,04)	16,0 (0,2)	0,30 (0,01)	0,083 (0,000)	16,6 (0,1)	189 (1)	0,213 (0,004)	0,069 (0,001)
Goiaba									
Lote A	2,70 (0,08)	3,35 (0,00)	6,1 (0,2)	0,069 (0,002)	0,061 (0,000)	16,5 (0,4)	114 (1)	0,044 (0,002)	0,025 (0,001)
Lote B	2,78 (0,05)	3,49 (0,04)	6,56 (0,07)	0,059 (0,001)	0,063 (0,001)	16,0 (0,0)	111,4 (0,3)	0,050 (0,001)	0,027 (0,000)
Manga									
Lote A	5,83 (0,04)	7,56 (0,05)	8,19 (0,09)	0,112 (0,003)	0,299 (0,002)	15,5 (0,1)	116,3 (0,6)	0,055 (0,001)	0,066 (0,001)
Lote B	5,44 (0,08)	6,95 (0,10)	7,8 (0,1)	0,139 (0,003)	0,388 (0,007)	15,4 (0,2)	112 (2)	0,059 (0,004)	0,075 (0,001)
Abacaxi									
Lote A	14,65 (0,03)	12,43 (0,08)	8,32 (0,07)	0,425 (0,007)	1,126 (0,005)	17,3 (0,0)	140,8 (0,9)	0,061 (0,002)	0,039 (0,000)
Lote B	16,5 (0,1)	10,22 (0,04)	6,86 (0,03)	0,248 (0,002)	1,081 (0,006)	17,2 (0,1)	101,1 (0,2)	0,057 (0,001)	0,032 (0,001)
Acerola									
Lote A	4,19 (0,06)	6,08 (0,05)	9,1 (0,1)	0,169 (0,004)	0,022 (0,001)	16,1 (0,2)	90,5 (0,7)	0,048 (0,002)	0,030 (0,001)
Lote B	4,14 (0,02)	6,16 (0,06)	9,24 (0,03)	0,160 (0,003)	0,025 (0,000)	22,2 (0,2)	85,3 (0,5)	0,050 (0,001)	0,031 (0,001)
Uva									
Lote A	11,04 (0,08)	7,73 (0,03)	9,9 (0,1)	0,180 (0,002)	0,219 (0,001)	16,3 (0,1)	27,6 (0,1)	0,056 (0,004)	0,394 (0,003)
Lote B	11,1 (0,2)	7,7 (0,1)	10,2 (0,1)	0,180 (0,003)	0,221 (0,004)	16,7 (0,3)	24,6 (0,3)	0,061 (0,003)	0,380 (0,006)

A representação gráfica das componentes principais permite a caracterização dos minerais presentes nas diferentes amostras de suco de fruta estudadas. Os gráficos dos dois primeiros eixos (PC) que estão associados à cada variável (Ca, Mn, Cu, Fe, Mg, Na, K, P e Zn) são mostrados nas Figuras 9.1-A e 9.1-B. A PCA mostrou que a primeira componente principal (PC1) explica 49,6% da variância total dos dados, sendo que a segunda e a terceira componentes principais explicam 19,6 % e 13,8 %, respectivamente.

A primeira componente principal está relacionada principalmente aos minerais Ca, Mg, Fe e P, enquanto a segunda componente principal relaciona-se aos minerais Mn, Cu, P e K, principalmente. Pelos gráficos de *escores* e *pesos* (Figuras 9.1-A e 9.1-B) observa-se que existe a separação entre os sucos de caju, maracujá, goiaba, acerola e manga, na primeira componente principal, enquanto a segunda componente principal separa os sucos de abacaxi e uva dos demais.

Na primeira componente principal, as amostras de sucos de caju são as que têm menores valores de *escores* (Figura 9.1-A) e estão separadas das demais por apresentar os menores teores dos minerais Ca, Mg e Fe que são os metais que mais se destacam nesta componente principal (Figura 9.1-B). Os sucos de abacaxi apresentam valores de *escores* mais positivos nas PC1 e PC2 (Figura 9.1-A). Comparando com os pesos (Figura 9.1-B), observa-se que isto se deve ao fato de que estes sucos são aqueles que têm maiores teores de Mn, Fe, Mg, e Ca. Os sucos de uva têm *escores* altamente positivos na PC1 e negativos na PC2 (Figura 9.1-A). Para os demais sucos, goiaba, acerola, maracujá e manga, apesar de estarem satisfatoriamente separados na PC1, eles se encontram próximos ao zero das PC1 e PC2 e, portanto, não foi possível atribuir quais minerais foram importantes para suas separações.

A terceira componente principal (13,8% da variância total) tem alta contribuição de zinco e potássio, sendo responsável pela discriminação dos sucos de maracujá. É interessante notar que os diferentes sucos, foram agrupados entre si, de acordo com as diferenças nos teores de minerais encontrados. Os sucos de acerola e goiaba foram aqueles que se mantiveram próximos nestas três componentes principais, indicando que seus teores de minerais diferenciam menos.

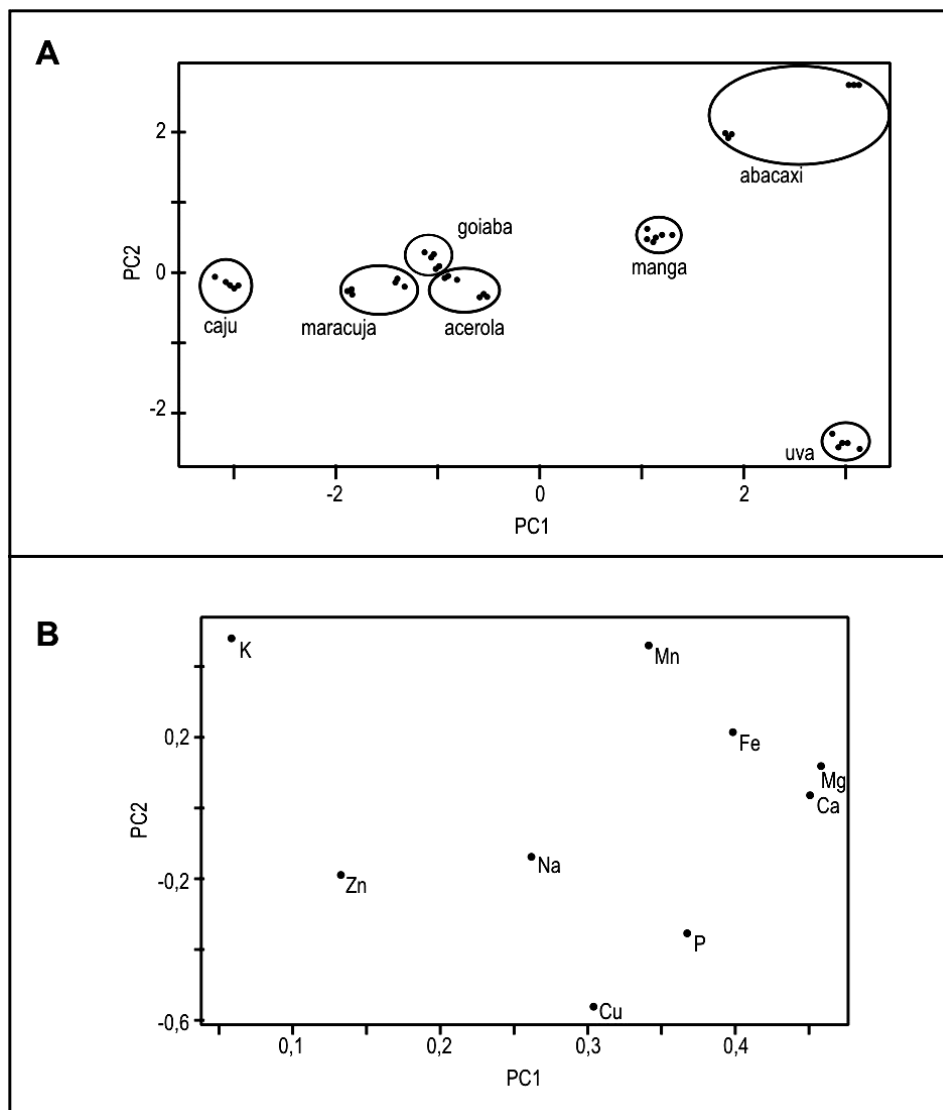


Figura 9.1. A - Gráfico de escores. **B** - Gráfico de pesos.

A análise hierárquica de agrupamentos (HCA) complementa a análise de componentes principais (PCA), sendo uma outra forma de visualizar as semelhanças e as diferenças na composição mineral dos diferentes sucos. O grau de similaridade entre as amostras varia entre 0 e 1 (quando amostras são idênticas). Com um grau de similaridade 0,7, destacam-se 6 grupos de sucos: uva, abacaxi, maracujá, caju, manga e um grupo com acerola e goiaba. Estes grupos correspondem aos mesmos grupos formados pelas 3 componentes principais, evidenciando a importância de usarmos estes dois tipos de análises em conjunto para a interpretação dos resultados. É interessante notar também na Figura 9.2, que as diferentes variedades

de sucos formam cada uma dois grupos distintos correspondendo aos dois lotes de sucos de frutas estudados.

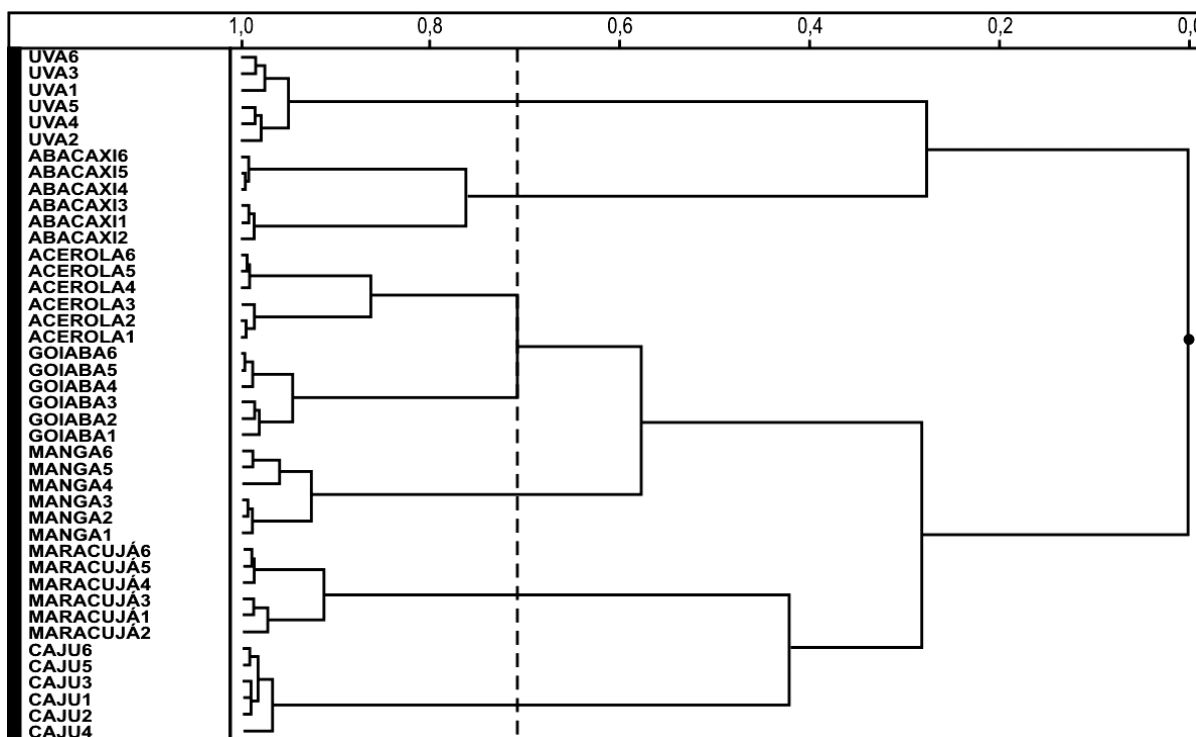


Figura 9.2. Dendrograma obtido pela HCA para as amostras de sucos de frutas.

Os resultados obtidos para os elementos minerais presentes nas amostras de sucos de frutas estudados foram avaliados quanto à Ingestão Diária Recomendada (IDR) [NRC, 1989; ANVISA, 2003], considerando a diluição para cada suco e uma ingestão de um copo de 300 mL (Tabela 9.6). O suco de abacaxi apresenta uma contribuição nutricional elevada de manganês (144 % da IDR), seguido pelo suco de manga (45 % da IDR) e pelo suco de uva (29 % da IDR). Para o elemento cobre, um copo de suco de uva fornece 129 % da IDR, seguida pelo suco de manga (24 % da IDR) e de maracujá (22 % da IDR). Para o cálcio e sódio nenhuma das amostras de sucos estudadas podem ser considerados como fonte deste mineral, pois as contribuições para a ingestão diária recomendadas são baixas. O suco de abacaxi foi aquele que apresentou a maior contribuição para a IDR para o elemento zinco (9 % da IDR) e os sucos de maracujá e abacaxi apresentaram a maior contribuição para a IDR de ferro (7 % da IDR). Com relação ao potássio as maiores contribuições

nutricionais são provenientes dos sucos de maracujá (18 % da IDR), seguido pelos sucos de goiaba, manga e abacaxi com 10 % da IDR. Quanto ao magnésio, os sucos de abacaxi, maracujá e caju são os com maior contribuição nutricional (13, 12 e 10% da IDR, respectivamente). Estes cálculos não levam em consideração a biodisponibilidade, a qual é afetada por fatores individuais, presença de inibidores ou compostos com ação sinérgica no alimento, solubilidade, concentração e estado de oxidação do mineral [Ekmekcioglu, 2000].

Tabela 9.6. Contribuição nutricional média estimada para adultos de um copo de 300 mL de suco, após a diluição recomendada pelo fabricante.

Elemento Mineral	IDR* (mg)	% da Ingestão Diária Recomendada						
		Caju	Maracujá	Goiaba	Manga	Abacaxi	Acerola	Uva
Ca	1000	0	1	1	2	5	1	3
Mg	260	10	12	4	8	13	7	9
P	700	4	7	3	3	3	4	4
Fe	14	3	7	1	3	7	4	4
Mn	2,3	10	11	8	45	144	3	29
Na	2400	4	2	2	2	2	2	2
K	3500	9	18	10	10	10	8	2
Zn	7	4	9	2	2	3	2	3
Cu	0,9	12	22	9	24	12	10	129

* NRC, 1989; ANVISA, 2003.

9.4. CONCLUSÃO

- O método de extração com ácido clorídrico mostrou ser adequado para a determinação de minerais em amostras de suco de frutas pelos melhores níveis de recuperação obtidos para os minerais determinados, pelo maior número de amostras que podem ser extraídas simultaneamente e pela sua simplicidade. O método empregando a energia de microondas apresentou melhor precisão para alguns

metais, porém os níveis de recuperação foram inferiores aos obtidos pelo método de extração com ácido clorídrico.

- A composição de minerais nas diferentes variedades de sucos de frutas estudados diferiram entre si. A análise de componentes principais proporcionou uma interpretação multivariada dos resultados, mostrando que as amostras são agrupadas de acordo com as diferenças nos teores dos minerais. A primeira componente principal identifica os sucos de manga e caju. A segunda, os sucos de abacaxi e uva. A terceira componente principal separa dos restantes, o suco de maracujá. Os sucos de goiaba e acerola são aqueles que têm mais características semelhantes.
- O suco de abacaxi apresentou os maiores teores de cálcio e magnésio; o suco de uva apresentou os maiores níveis de cobre; o suco de maracujá os maiores níveis de fósforo, potássio e zinco.
- A análise de agrupamento por método hierárquico, utilizada para confirmar os resultados da PCA, produziu resultados semelhantes, inclusive identificando os dois lotes de sucos estudados.
- A contribuição nutricional para adultos de um copo dos sucos de frutas estudados são pobres em cálcio e sódio e, em geral, ricas em cobre e manganês. Os sucos de abacaxi, de manga e de uva têm grande contribuição para a ingestão diária de manganês e os sucos de uva, de manga e de maracujá têm grande contribuição de cobre.

10. CONCLUSÕES GERAIS

- As técnicas de espectroscopia no infravermelho médio e próximo por refletância, empregada para determinar o teor de sacarose em açúcar cristal e os teores de umidade, açúcar total e proteínas apresentam algumas vantagens sobre os métodos de referência (polarimetria e AOAC), considerando o tempo de análise, a não geração de resíduos, além do baixo custo, justificando sua utilização em análises de rotina no controle de qualidade.
- O uso da técnica de espectroscopia no infravermelho com transformada de Fourier (FT-IR) e do acessório de ATR associada ao método PLS de regressão forneceu resultados erro relativo médio de 0,5% para a determinação do teor de sacarose em amostras de açúcar cristal.
- A técnica usando espectroscopia NIR associada ao método PLS de regressão para a quantificação de algumas propriedades químicas do café cru, como proteínas, açúcares totais e proteínas também podem ser usadas com boa precisão e exatidão.
- Técnicas de correção de sinal, variações na linha de base e normalização dos dados são necessárias para remover as interferências originadas pela heterogeneidade na superfície das amostras de café, melhorando significativamente a análise quantitativa (precisão dos modelos).
- Baseando-se nos estudos pode-se concluir que a combinação entre calibração multivariada e a espectroscopia NIR é uma maneira apropriada para análises quantitativas rápidas e estimulando o desenvolvimento de métodos que não agredem o meio ambiente.

- De forma geral, verificou-se a existência de algumas amostras anômalas de café cru, isto pode ser atribuído à heterogeneidade das amostras obtidas de diferentes procedências, variedades, solo, clima, fertilização, etc.
- Os resultados obtidos demonstram que provavelmente outras propriedades do café também podem ser modeladas, de forma a se obter modelos de regressão capazes de prever suas concentrações nas amostras.
- A determinação dos teores de minerais em sucos de frutas é importante para garantir a qualidade e autenticidade dos mesmos. A técnica analítica usando ICP OES para a quantificação dos elementos minerais mostrou-se viável, por ser uma técnica multielementar e possibilitar a leitura com mínimo de tratamento das amostras, reduzindo de forma considerável o tempo e custo às análises químicas.
- A análise de componentes principais evidenciou as características comuns e discrepantes entre os diferentes sucos de frutas, importantes para o controle de qualidade dos produtos. A análise hierárquica por agrupamentos complementou a análise de componentes principais sendo uma outra forma de visualizar as diferenças e semelhanças na composição mineral dos diferentes sucos de frutas, já que ambas podem ser usadas para a diferenciação e caracterização de amostras de sucos de frutas baseadas nos teores dos elementos minerais estudados.

REFERÊNCIAS BIBLIOGRÁFICAS

AMORIM, H.V. Chemistry of Brazilian green coffee and quality of the beverage: V – multiple linear regression analysis. **Turrialba**, v. 25, n. 1, p. 25-28, 1975.

AMORIM, H.V.; SMUCKER, R.; PFISTER, R. Some physical aspects of Brazilian green coffee beans and quality of the beverage. **Turrialba**, v. 26, n. 1, p. 24-27, 1976.

AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA (ANVISA). **Resolução RDC nº 360**, de 23 de dezembro de 2003. [Internet]. Disponível em http://www.anvisa.gov.br/legis/resol/2003/rdc/360_03rdc.ftm. Acesso em 05 de janeiro de 2004.

BAYER, S.; MCHARD, J.A.; WINEFORDNER, J.D. Determination of the geographical origins of frozen concentrated orange juice via pattern recognition. **Journal of Agricultural Food Chemistry**, v. 28, p. 1306-1307, 1980.

BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. **Chemometrics: A Practical Guide**, Wiley, New York, 1998.

BEEBE, K.B.; KOWALSKI, B.R. An introduction to multivariate calibration and analysis. **Analytical Chemistry**, v. 59, p. 1007A, 1987.

BELLON-MAUREL, V.; VALLAT, C.; GOFFINET, D.; Quantitative analysis of individual sugars during starch hydrolysis by FT-IR/ATR spectrometry. 2. Influence of external factors and wavelength. **Applied Spectroscopy**, v. 49, n. 5, p. 563-568, 1995.

BELLON-MAUREL, V.; VALLAT, C.; GOFFINET, D. Quantitative analysis of individual sugars during starch hydrolysis by FT-IR/ATR spectrometry. Part I:

multivariate calibration study—repeatability and reproducibility. **Applied Spectroscopy**, v. 49, n. 5, p. 556-562, 1995.

BIRTH,G.S.; HECHT,H.G. **The physics of near-infrared reflectance in near-infrared technology in the agricultural and food industries**. By Phil Williams & Karl Norris, St. Paul: American Association of Cereal Chemists, 1987.

BOSS, C.B.; FREDEEN, K.J. **Concepts, instrumentation and techniques in inductively coupled plasma optical emission spectrometry**. 2nd ed. Perkin Elmer Corp., USA, 1999.

BOU, R.; GUARDIOLA, F.; PADRÓ, A.; PELFORT, E.; CODONY, R. Validation of mineralisation procedures for the determination of selenium, zinc, iron and copper in chicken, meat and feed samples by ICP-AES and ICP-MS. **Journal of Analytical Atomic Spectrometry**, v. 19, n. 10, p. 1361-1369, 2004.

BOUHSAIN, Z.; GARRIGUES, J.M.; GARRIGUES, S.; DE LA GARDIA, M.; Flow injection Fourier transform infrared determination of caffeine in coffee. **Vibrational Spectroscopy**, v. 21, p. 143-150, 1999.

BRIANTED, R.; KEMSLEY, E.K.; WILSON, R.H.; Discrimination of Arabica and Robusta in instant coffee by Fourier transform infrared spectroscopy and chemometrics. **Journal of Agricultural Food Chemistry**, v. 44, n. 1, p. 170-174, 1996.

BRIANTED, R.; KEMSLEY, E.K.; WILSON, R.H.; Approaches to adulteration in instant coffees using infrared spectroscopy and chemometrics. **Journal of Agricultural Food Chemistry**, v. 44, n. 3, p. 359-366, 1996.

BÜNING-PFAUE, H. Analysis of water in food by near infrared spectroscopy, **Food Chemistry**, v. 82, p. 107-115, 2003.

CAFÉ DAMASCO - **História do Café**. [Internet]. Disponível em: http://www.cafedamasco.com.br/historia_composicao.asp. Acesso em: 15 de agosto de 2005.

CÁMARA, M.; DIEZ, C.; TORIJA, E. Chemical characterization of pineapple juices and nectars. Principal components analysis. **Food Chemistry**, v. 54, p. 93-100, 1995.

CARVALHO, V.D. de; CHAGAS, S.J. de R.; SOUZA, S.M.C. Fatores que afetam a Qualidade do Café. **Informe Agropecuário**, v. 18, n. 187, p. 5-20, 1997.

CADET, F.; BERTRAND, D.; ROBERT, P.; MAILLOT, J.; DIEUDONNÉ, J.; ROUCH, C. Quantitative determination of sugar cane sucrose by multidimensional statistical analysis of their mid-infrared attenuated total reflectance spectra. **Applied Spectroscopy**, v. 45, n. 2, p. 166-172, 1991.

COATES, J. Vibrational Spectroscopy: Instrumentation for infrared and Raman spectroscopy. **Applied Spectroscopy Review**, v. 33, p. 267-425, 1998.

COLEMAN, P.B. **Practical sampling techniques for infrared analysis**. CRC Press, London, 1993.

CONSECANA-PR, Conselho dos produtores de cana-de-açúcar, açúcar e álcool do estado do Paraná. **Normas operacionais de avaliação da qualidade da cana-de-açúcar**. [Internet]. Disponível em: <http://www.faepr.com.br/consecana/normasop.htm>. Acesso em: 04 de outubro de 2005.

COSTA FILHO, P.A.; POPPI, R.J. Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio. Determinação simultânea de glicose, frutose e maltose. **Química Nova**, v. 25, n. 1, p. 46-52, 2002.

CURRIE, L.A.; SVEHLA, G. Nomenclature for the presentation of results of chemical analysis. **International Union of Pure and Applied Chemistry**, v. 66, p. 595-608, 1994.

DAVIES, A.M.C. Progress in human food analysis by near-infrared. **Analytical Proceedings**, v. 21, p. 488-491, 1984.

DEVAUX, M.F.; DUFOUR, N.N.; ROBERT, P.; BERTRAND, D. Effects of particle size on the near-infrared reflectance spectra of wheat and rape seed meal mixtures. **Applied Spectroscopy**, v. 49, p. 84-91, 1995.

DOWNEY, G.; SPENGLER, B.; Compositional analysis of coffee blends by near infrared spectroscopy. **Irish Journal Agricultural Food Research**, v. 35, n. 2, p. 179-188, 1996.

EKMEKCIOGLU, C. Intestinal bioavailability of minerals and trace elements from milk and beverage in humans. **Nahrung**, v. 44, p. 390-397, 2000.

FARIA, C.G.; FERREIRA, M.M.C. **QUIMIOPAC, LQTA**, Instituto de Química, Universidade Estadual de Campinas, lqta.iqm.unicamp.br., 2002.

FERRÃO, M. F.; FURTADO, J.C.; NEUMANN, L.G.; KONZEN, P.H.A.; MORGANO, M.A.; BRAGAGNOLO, N.; FERREIRA, M.M.C. Técnica não destrutiva de análise de tanino em café empregando espectroscopia no infravermelho e algoritmo genético. **Tecno Lógica**, v. 7, n. 1, p. 9-26, 2003.

FERRÃO, M.F.; CARVALHO, C.W.; MULLER, E.I.; DAVANZO, C.U. Determinação simultânea dos teores de cinza e proteína em farinha de trigo empregando NIR-PLS e DRIFT-PLS. **Ciência e Tecnologia de Alimentos**, v. 24, n. 3, p. 333-340, 2004.

FERRÃO, M.F.; DAVANZO, C.U. Horizontal attenuated reflection applied to simultaneous determination of ash and protein contents in commercial wheat flour. **Analytica Chimica Acta**, v. 540, p. 411-415, 2005.

FERREIRA, M.M.C.; MORGANO, M.A.; QUEIROZ, S.C.N.; MANTOVANI, D.M.B. Relationships of the Mineral and Fatty Acid Contents in Processed Turkey Meat. **Food Chemistry**, v. 69, n. 3, p. 259-265, 2000.

FILLI, S.P.; OLIVEIRA, E.; OLIVEIRA, P.V. On-line digestion in a focused microwave-assisted oven for elements determination in orange juice by inductively coupled plasma optical emission spectrometry. **Journal of Brazilian Chemical Society**, v. 14, n. 3, p. 435-441, 2003.

GELADI, P.; KOWALSKI, B.R. Partial least-squares regression – A tutorial. **Analytica Chimica Acta**, v. 185, p. 1-17, 1986.

GINÉ, M.F. **Espectrometria de emissão atômica com plasma acoplado indutivamente**, série didática, v. 3, CENA/USP, Piracicaba, 1998, cap. 2.

GISLUM, R.; MICKLANDER, E.; NIELSEN, J. P., Quantification of nitrogen concentration in perennial ryegrass and red fescue using near-infrared reflectance spectroscopy (NIRS) and chemometrics. **Field Crops Research**, v. 88, p. 269-277, 2004.

GRAEFF, S.; CLAUPEIN, W, Quantifying nitrogen status of corn (*Zea mays* L.) in the field by reflectance measurements. **European Journal of Agronomy**, v. 19, p. 11-618, 2003.

GUPTA, K.K.; BHATTACHARJEE, S.; KAR, S.; CHAKRABARTY, S. THAKUR, P.; BHATTACHARYYA, G.; SRIVASTAVA, S.C. Mineral compositions of eight common

spices. **Communications in Soil Science and Plant Analysis**, v. 34, n. 5-6, p. 681-693, 2003.

GUTTERES, M. V.; VOLPE, P. O. L.; FERREIRA M. M. C., Multiway Calibration for Creatinine Determination in Human Serum Using the Jaffé Reaction, **Applied Spectroscopy**, v. 58, p. 54-60, 2004.

HAALAND, D.M. Multivariate calibration methods applied to quantitative FT-IR analysis practical Fourier transform infrared spectroscopy. **Industrial and Laboratory Chemical Analysis**. In: FERRARO, J.R.; KRISMAN, K. San Diego, Academic Press, 1990, 395-469.

HERMIDA, M.; GONZALEZ, J. M.; SANCHEZ, M.; RODRIGUEZ, O.; JOSE, L. Moisture, solids-non-fat and fat analysis in butter by near infrared, **International Dairy Journal**, v. 11, p. 93-98, 2001.

HORWITZ, W. (Ed.). **Official Methods of Analysis of the Association of Official Analytical Chemists**. 17th ed. Gaithersburg, Maryland: AOAC, 2000. v. 2, chapter. 44, method 44.1.6, p.9.

HORWITZ, W. (Ed.). **Official Methods of Analysis of the Association of Official Analytical Chemists**. 17th ed. Gaithersburg, Maryland: AOAC, 2000. v. 2, chapter. 30, method 968.11, p. 4.

IAL - INSTITUTO ADOLFO LUTZ. **Normas analíticas do Instituto Adolfo Lutz: métodos químicos e físicos para análise de alimentos**. 3. ed. São Paulo, 1985. v. 1. MÉTODO 4.12 CAP. 4, P. 44.

IKEGAYA, K.; IWAMOTO, M.; UOZUMI, J.; NISHINARI, K.; Near infrared spectra of caffeine and its related compounds and their application to determination of caffeine content in green tea. **Nippon Shokuhin Kogyo Gakkaishi**, v. 34, p. 254-258, 1987.

ILLY, A; VIANI, R. **Espresso Coffee: The Chemistry of Quality**. Academic Press, London, 1998.

INMETRO - Instituto Nacional de Metrologia, Normalização e Qualidade Industrial; **Orientação sobre validação de métodos de ensaios químicos**, DOQ-CGCRE-008, 2003. p. 32.

KONZEN, P.H.A.; FERRÃO, M.F.; FURTADO, J.C.; MORGANO, M. A.; BRAGAGNOLO, N.; FERREIRA, M.M.C. Determinação não destrutiva de proteína por DRIFTS/PLS/AG. In: **XVIII Congresso Brasileiro de Ciência e Tecnologia de Alimentos**, Porto Alegre, 04 a 07 de agosto de 2002, 2310-2313. 1 CD-ROM.

KORTE, E.H.; OTTO, A. Infrared diffuse reflectance accessory for local analysis on bulky samples. **Applied Spectroscopy**, v. 42, p. 38-43, 1988.

LINDGREN, F; GELADI, P.; WOLD, S. Kernel-based PLS regression; Cross validation and applications to spectral data. **Journal of Chemometrics**, v. 8, p. 377-389, 1994.

LÓPEZ, M.A.A.; ROJAS, R.M.; SEGARRA, P.J.S.; COSANO, G.Z. Tratamiento de las matrices lácteas para la determinación de trazas metálicas por absorción atómica. Revisión. **Alimentaria**, v. 34, n. 271, p. 71-78, 1996.

LUNA, F.R.T.; ERIN, M.; FRANÇA, R.S.; BELAI DI, H. Medidas de dispersão angular em placas de quartzo e aplicações na indústria do açúcar. **Metrologia-2003**, set. 2003, p. 01-05.

MACKENZIE, M.W. **Advances in applied Fourier transform infrared spectroscopy**. Chichester, John Wiley & Sons, 1988.

MARTENS, H.; NAES, T. **Multivariate Calibration**. John Wiley & Sons, New York 1996.

MATLAB – The Language of Technical Computing. **MATLAB User's guide**, Ver.5.1. The MathWorks Inc.,1997.

MILLER, J.N.; MILLER, J.C. **Statistics and chemometrics for analytical chemistry**. Chinchester: Prentice Hall, 4nd, 2000.

MILLER-IHLI, N.J. Atomic absorption and atomic emission spectrometry for the determination of the trace element content of selected fruits consumed in the united states. **Journal Food Composition of Analysis**, v. 9, p. 301-311, 1996.

MIRABELLA, F.M. Internal reflection spectroscopy. **Applied Spectroscopy**, v. 21, p. 45-178, 1985.

MIRALBÉS C., Quality control in the milling industry using near infrared transmittance spectroscopy. **Food Chemistry**, v. 88, p. 621-628, 2004.

MONTEIRO, M.C.; TRUGO, L.C. Determinação de compostos bioativos em amostras comerciais de café torrado. **Química Nova**, v. 28, n. 1, p. 637-641, 2005.

MORGANO, M.A.; FARIAS, C.G.; FERRÃO, M.F.; BRAGAGNOLO, N.; FERREIRA, M.M.C. Determinação de proteína em café cru por espectroscopia NIR e Regressão PLS. **Ciência e Tecnologia de Alimentos**, v. 25, n. 1, p. 25-31, 2005.

MORGANO, M. A.; CAMARGO, C.; PAGEL, A.P.; FERRÃO, M.F.; BRAGAGNOLO, N.; FERREIRA, M.M.C. Determinação simultânea dos teores de cafeína, trigonelina e ácido clorogênico em amostras de café cru por análise multivariada (PLS) em dados de espectroscopia por reflexão difusa no infravermelho próximo. In: **II Simpósio de Pesquisa dos Cafés do Brasil**, Vitória, 24 a 27 de setembro de 2001, 1502-1510. 1 CD-ROM.

MORGANO, M.A.; QUEIROZ, S.C.N.; FERREIRA, M.M.C. Aplicação da Análise exploratória na diferenciação de vegetais. **Brazilian Journal of Food Technology**, v. 2, n. 1/2, p. 73-79, 1999.

MORGANO, M.A.; MORIYA, C.; FERREIRA, M.M.C. Determinação quantitativa do teor de sacarose em açúcar por espectroscopia FT-IR/ATR e regressão PLS. **Brazilian Journal of Food Technology**, v. 6, p. 77-83, 2003.

NATIONAL RESEARCH COUNCIL. **Recommended dietary allowances**. National Academy Press. Washington, D.C., 1989.

NIEMEYER, J.; CHEN, Y.; BOLLAG, J.M. Characterization of humic acids, composts, and peat by diffuse reflectance Fourier-transform infrared spectroscopy. **Soil Science Society of America Journal**, v. 56, p. 130-135, 1992.

NOUSIAINEN, J.; AHVENJÄRVI, S.; RINNE, M.; HELLÄMÄKI M.; HUHTANEN, P. Prediction of indigestible cell wall fraction of grass silage by near infrared reflectance spectroscopy. **Animal Feed Science and Technology**, v. 115, p. 295-311, 2004.

OLINGER, J.M.; GRIFFITHS, P.R. Effects of sample dilution and particle size/morphology on diffuse reflection spectra of carbohydrate systems in the near and mid-Infrared. part I: Single analytes. **Applied Spectroscopy**, v. 47, p. 687-694, 1993.

OLINGER, J.M. & GRIFFITHS, P.R. Effects of sample dilution and particle size/morphology on diffuse reflection spectra of carbohydrate systems in the near and mid-Infrared. part II: Durum wheat. **Applied Spectroscopy**, v. 47, p. 695-701, 1993.

OZCAN, M. Determination of mineral contents of Turkish Herbal tea (*Salvia aucheri* var. *canescens*) at different infusion periods. **Journal of Medicinal Food**, v. 8, n. 1,

p. 110-112, 2005.

PENNINGTON, J.A.T.; YOUNG, B. Iron, zinc, manganese, selenium and iodine in foods from the United States total diet study. **Journal of Food Composition and Analysis**, v. 3, p. 166-184, 1990.

PIROUETTE, Multivariate Data Analysis for IBM PS Systems, version 3.1, Infometrix, Seattle, WA, USA, 2001.

PORTAL DE EMBALAGENS (FURG) - **Embalagens para sucos**. [Internet]. Disponível em: <http://www.furg.br/portaldeembalagens/quatro/sucos.html>. Acesso em: 17 de agosto de 2005.

PROCANA – **Ligado no dia a dia do setor** [Internet]. Disponível em: <http://www.procana.com.br/Conteudo/Conheca%20%20Setor.asp>. Acesso em: 20 de abril de 2005.

QUAST, D. G. Tipos de açúcar produzidos no Brasil. **Alimentos e Tecnologia**, 10ªed. 49-52, jul/ago 1986.

REEVES III, J.B.; ZAPF, C.M. Mid-infrared diffuse reflectance spectroscopy for discriminant analysis of food ingredients. **Journal of Agricultural and Food Chemistry**, v. 46, p. 3614-3622, 1998.

RIZZON, L.A.; MIELE, A. Características analíticas de sucos de uva elaborados no Rio Grande do Sul. **Boletim da Sociedade Brasileira de Ciência e Tecnologia de Alimentos**, v. 29, p. 129-133, 1995.

RUBAYIZA, A.B.; MEURENS, M. Chemical discrimination of arábica and robusta coffees by Fourier transform Raman spectroscopy. **Journal of Agriculture and Food Chemistry**, v. 53, n. 12, p. 4654-4659, 2005.

SALGO, A.; NAGY, J.; MIKÓ, É. Application of near infrared spectroscopy in the sugar industry. **Journal of Near Infrared Spectroscopy**, v. 6, p. A101-A106, 1998.

SCHULZ, H.; SCHRADER, B.; QUILTZSCH, R.; TEUER, B. Quantitative analysis of various citrus oils by ATR/FT-IR and NIR-FT Raman spectroscopy. **Applied Spectroscopy**, v. 56, n. 1, p. 117-124, 2002.

SCHUSTER, J.H. Semi-automated proximate analysis. **Cereal Food World**, v. 23, p. 180-182, 1978.

SCHWAB, E. C.; SHAVER, R. D.; LAUER, J. G.; COORS, J. G. Estimating silage energy value and milk yield to rank corn hybrids. **Animal Feed Science and Technology**, v. 109, p. 1-18, 2003.

SKOOG, D.A.; HOLLER, F. J.; NIEMAN, T.A. **Principles of instrumental analysis**. 5th ed., Harcourt Brace College Publishers, Philadelphia, 1998, p. 230 a 252.

SENA, M.M.; TREVISAN, M.G.; POPPI, R.J. Parafac: Uma ferramenta quimiométrica para tratamento de dados multidimensionais. Aplicações na determinação direta de fármacos em plasma humano por espectrofluorimetria. **Química Nova**, v. 28, n. 5, p. 910-920, 2005.

SGARBIERI, W.C. **Alimentação e nutrição: Fator de saúde e desenvolvimento**. Unicamp/Almed, Campinas/São Paulo, 1987.

SHARAF, M.A.; ILLMAN, D.L.; KOWALSKI, B.R. **Chemometrics**, 1^a ed., John Wiley & Sons, New York, 1986, 332p.

SCHINDLER, R.; WATKINS, M.; VONACH, R.; LENDL, B.; KELLNER, R. Automated multivariate calibration in sequential injection-Fourier transform infrared spectroscopy for sugar analysis. **Analytical Chemistry**, v. 70, n. 2, p. 226-231, 1998.

SIMPKINS, W.A.; LOUIE, H.; WU, M.; HARRISON, M.; GOLDBERG, D. Trace elements in Australian orange juice and other products. **Food Chemistry**, v. 71, p. 423-433, 2000.

SINGH, B.R.; WECHTER, M.A.; HU, Y.; LAFONTAINE, C.; Determination of caffeine content in coffee using Fourier transform infra-red spectroscopy in combination with attenuated total reflectance technique: a bioanalytical chemistry experiment for biochemists. **Biochemical Education**, v. 26, p. 243-247, 1998.

SOARES, L.M.V; SHISHIDO, K.; MORAES, A.M.M; MOREIRA, V.A. Composição mineral de sucos concentrados de frutas brasileiras. **Ciência e Tecnologia de Alimentos**, v. 24, n. 2, p. 202-206, 2004.

SOARES, A. M. B.; MURRAY, I.; PATERSON, R. M.; ABREU, J. M. F., Use of near infrared reflectance spectroscopy (NIRS) for the prediction of the chemical composition and nutritional attributes of green crop cereals. **Animal Feed Science and Technology**, v. 75, p. 15-25,1998.

TEVRUCHT, M.L.E.; GRIFFITHS, P.R. Quantitative investigation of matrices for diffuse reflectance infrared Fourier transform spectrometry. **Talanta**, v. 38, p. 839-849, 1991.

TEWARI, J.; MEHROTRA, R.; IRUDAYARAJ, J. Direct near infrared analysis of sugar cane clear juice using a fibre-optic transmittance probe. **Journal of Near Infrared Spectroscopy**, v. 11, n. 5, p. 351-356, 2003.

THOMAS, E.V. A primer on multivariate calibration. **Analytical Chemistry**, v. 66, n. 15, p. A795-A804, 1994.

THOMAS, E.V.; HAALAND, D.M.; Comparison of multivariate calibration methods for quantitative spectral-analysis. **Analytical Chemistry**, v. 62, n. 10, p. 1091 - 1099, 1990.

TKACHUK, R. Analysis of whole grains by near-infrared reflectance. In: **Near-infrared technology in the agricultural and food industries**. Phil Williams & Karl Norris, St. Paul: American Association of Cereal Chemists, 1987.

TURRA, C. **Elementos químicos como critério de discriminação de citros orgânicos de convencionais**. Piracicaba, 2005. 82 f. Dissertação (Mestrado em Ecologia de Agroecossistemas), Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo.

ÚNICA – **União da Agroindústria Canavieira de São Paulo** [Internet]. Disponível em: http://www.unica.com.br/pages/unica_perfil1.asp. Acesso em: 20 de abril de 2005.

VALDERRAMA, P. **Avaliação de figuras de mérito em calibração multivariada na determinação de parâmetros de controle de qualidade em indústria alcooleira por espectroscopia no infravermelho próximo**. Campinas, 2005. 128 f. Dissertação (Mestrado em Química Analítica), Instituto de Química, Universidade Estadual de Campinas.

VALE, C.M.; ZAMBIAZI, R.C. Previsão de estabilidade de óleos vegetais através da rede de inteligência artificial. **Ciência e Tecnologia de Alimentos**, v. 20, n. 3, p.342-348, 2000.

VAN DE VOORT, F.R.; ISMAIL, A.A. Proximate analysis of foods by mid-FTIR spectroscopy. **Trends in Food Science and Technology**, v. 2, p. 13-17, 1991.

VAN DE VOORT, F.R. Fourier transform infrared spectroscopy applied to food analysis. **Food Research International**, v. 25, p. 397-403, 1992.

VETTORAZI, G.; MACDONALD, I. **Sacarose: aspectos nutricionais e de segurança no uso do açúcar**. HUCITEC, São Paulo, p. 35-64, 1989.

VILLAMARÍN, B.; FERNÁNDEZ, E.; MENDEZ, J.; COEERN, S. C. L., Analysis of grass silage from Northwestern Spain by near-infrared reflectance spectroscopy, **Journal of Association of Official Analytical Chemists International**, v. 85, p. 541-545, 2002.

WATSON, C. A., Near-infrared reflectance spectrophotometric analysis of agricultural products. **Analytical Chemistry**, v. 49, p. 835A-840A, 1977.

WETZEL, D.L. Near-infrared reflectance analysis sleeper among spectroscopic techniques. **Analytical Chemistry**, v. 55, p. 1165A-1176A, 1983.

WILLIAMS, P.C.; STEVENSON, S.G. Near-infrared reflectance analysis: Food industry applications. **Trends in Food Science and Technology**, v. 1, p. 44-48, 1990.

WILLIAMS, P.; NORRIS, K. **Near-Infrared technology in the agricultural and food industries**. Minnesota: AACC, 1998.

WILSON, R.H.; TAPP, H.S. Mid-infrared spectroscopy for food analysis: Recent new applications and relevant developments in sample presentation methods. **Trends in Analytical Chemistry**, v. 18, p. 85-93, 1999.

WILSON, R.H. Fourier transform mid-infrared spectroscopy for food analysis. **Trends in Analytical Chemistry**, v. 9, p.127-131, 1990.

WOJCIECHOWSKI, C.; DUPUY, N.; TA, C. D.; HUVENNE, J. P.; LEGRAND, P. Quantitative analysis of water-soluble vitamins by ATR-FTIR spectroscopy. **Food Chemistry**, v. 63, n. 1, p. 133-140, 1998.