



Universidade Estadual de Goiás

Unidade Universitária de Ciências Exatas e Tecnológicas (UnUCET)

MESTRADO EM CIÊNCIAS MOLECULARES

DETERMINAÇÃO DA TENSÃO INTERFACIAL DE ÓLEOS ISOLANTES USANDO ANÁLISE DE IMAGENS E ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO

Dissertação de Mestrado

Aluna: Mariana da S. Godinho

**Orientador: Prof. Dr. Marcelo Martins de Sena
Co-Orientador: Prof. Dr. Anselmo E. de Oliveira (IQ/UFG)**

**Anápolis
Outubro/2009**

Catálogo na Fonte
Biblioteca UnUCET – UEG
Helenir Freire Batista Machado – CRB 1704

611:543.38

G585d Godinho, Mariana da Silva

Determinação da tensão Interfacial de Óleos Isolantes usando Análise de Imagens e Espectroscopia no Infravermelho Próximo. / Mariana da Silva Godinho - Anápolis: Universidade Estadual de Goiás- Mestrado em Ciências Moleculares, 2009

96 p. ; 31 cm. – Dissertação/ Universidade Estadual de Goiás, Unidade Universitária de Ciências Exatas e Tecnológicas – Mestrado Ciências Moleculares, 2009.

I. Godinho, Mariana da Silva

II. Título

III. Dissertação

1. Transformadores.

2. Óleos Isolantes.

3. Quimiometria.

4. N- PLS

5.

PARAFAC

DETERMINAÇÃO DA TENSÃO INTERFACIAL DE ÓLEOS ISOLANTES USANDO
ANÁLISE DE IMAGENS E ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO

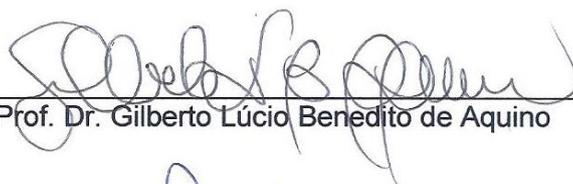
MARIANA DA SILVA GODINHO

Dissertação apresentada ao corpo docente do Programa de Mestrado em Ciências Moleculares da Universidade Estadual de Goiás apresentada como parte dos requisitos necessários para obtenção do título de Mestre em Ciências Moleculares.

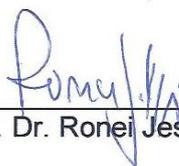
Aprovada por:



Prof. Dr. Marcelo Martins de Sena



Prof. Dr. Gilberto Lúcio Benedito de Aquino



Prof. Dr. Ronei Jesus Poppi

ANÁPOLIS, GO- BRASIL
SETEMBRO 2009

*À minha primeira mestra, minha
mãe Edná, pelo amor, apoio,
exemplo e carinho.*

Ao Alex ...

AGRADECIMENTOS

Trata-se de uma tarefa muito difícil, porém muito boa de ser escrita, pois é o momento de lembranças de todos aqueles que diretamente e indiretamente contribuíram para que o trabalho fosse concluído. Devido à dificuldade de separar a vida pessoal da acadêmica, os agradecimentos serão divididos em etapas.

Aos Colaboradores:

- À CELG pela doação das amostras de óleos e papéis.
- Aos químicos Marcos e Estevão do Laboratório de Ensaio Químicos da CELG, pela recepção no laboratório e informações de grande importância na elaboração deste trabalho.
- À Rede Transformadores pela doação de papel tipo Kraft novo.
- Ao LQTC/UFG pela estrutura física que foi fundamental na realização desse trabalho.
- Ao Prof. Ronei J. Poppi pela utilização do espectrofotômetro.
- Aos Prof^{os}. José Divino, Hamilton e Marcelo M. Sena pelo apoio e incentivo durante as disciplinas ministradas.
- A todos integrantes do Colégio Estadual Chico Mendes pela compreensão, apoio e amizade ao longo desse trabalho.

Um agradecimento muito especial aos professores:

- Dr. Marcelo Martins de Sena pela orientação, amizade, apoio e maravilhosas discussões.
- Dr. Anselmo Elcana de Oliveira pela co-orientação, pelo aprendizado decorrido da convivência dos últimos anos, pelo apoio e amizade.

Aos familiares:

- Aos meus pais Oride e Edna por sempre perdoar minhas falhas, compreender minhas imperfeições e sempre estarem dispostos a ajudar-me.
- À minha tia Divina por ser minha segunda mãe e sempre estar presente em minha vida.

Aos amigos:

- À Cristina Bernardes pela companhia (minha grande parceira), pela amizade e apoio, e discussões a respeito de Quimiometria e demais assuntos.
- À Helen Graci, mesmo distante fisicamente nunca deixou de estar presente.
- À Pollyana Godoy pela eterna amizade e paciência.
- Ao GC (grupo formado por: Guilherme Aires, João, Fabrício e Carolina Diomara) pela amizade de 8 anos.
- Ao Guilherme Aires por ser tão amigo e paciente e estar sempre com o celular ligado.
- À amiga Jussânia por ter ensinado ser uma pessoa melhor, mas principalmente pela amizade, apoio e compreensão com meus atrasos.
- Aos amigos Lucimeire, Udson, Serginho, Lélia, Milena, Lívia, Flávia, Kleber, Silvânia, Andréa, e Eliete.

Aos colegas de grupo (LQTC/UFG) e UEG:

- Deângelis Damasceno pelas longas discussões, companhia, paciência, lealdade e amizade.
- NilsonTavares pelo café e paciência.
- Diogo, River, Maycon e Claudimeire pela convivência no LQTC.
- Marcos Reis pela amizade e grande ajuda na disciplina de Quântica.
- Deomar e Emiret pela grande ajuda na disciplina de Espectroscopia.
- Aos colegas e amigos do ônibus: Josué, Evandro e Camila por ter tornados algumas tardes tão divertidas e inesquecíveis.

CURRICULUM VITAE

DADOS PESSOAIS

Nome: Mariana da Silva Godinho

Nascimento: 06/10/1983 – Itapuranga - GO

E-mail: marrygodinho@yahoo.com.br

FORMAÇÃO ACADÊMICA

Graduação em Curso Superior de Tecnologia em Química Industrial - 2002/2005
Centro Federal de Educação Tecnológica de Goiás, CEFETGO, Brasil

Graduação em Ciências - Habilitação em Química – 2002/2005
Universidade Estadual de Goiás, UEG, Anápolis, Brasil

ATUAÇÃO PROFISSIONAL

Professora concursada – nível médio

Colégio Estadual Garavelo Park - 2006/Atual

Professora da disciplina de Química – nível médio

Colégio Academia – 2005/2006

ATIVIDADES ACADÊMICAS

Iniciação científica – Instituto de Química da UFG, 2003-2006.

Orientador: Prof. Dr. Anselmo Elcana de Oliveira.

Monitoria - Disciplina de Análise Química Qualitativa (1º semestre/2004).

Palestra Ministrada (IQ/UFG)- Análise por Agrupamento Hierárquico

ARTIGOS PUBLICADOS

1. GODINHO, M. S., PEREIRA, R. O., KEYSE,, SCHIMIDT, Fernando, OLIVEIRA, S. B., OLIVEIRA, A. E.

Classificação de Refrigerantes Através de Análise de Imagens e Análise de Componentes Principais (PCA). Química Nova. , v.31, p.1485 - 1489, 2008.

TRABALHOS COMPLETOS PUBLICADOS EM ANAIS DE EVENTOS

1. Damasceno, D., GODINHO, M. S., GONCALVES, M. I., MESQUITA, N. A. S., SOARES, M. H. F. B.

Aplicação de Inseticida Natural em Aulas Experimentais no Ensino de Química XIV Encontro Nacional de Ensino de Química, 2008.

APRESENTAÇÃO ORAL DE TRABALHOS EM EVENTOS NACIONAIS

1. GODINHO, M. S., SILVA, T. C. O., OLIVEIRA, A. E.

Classificação de cerveja por análise de imagens e PCA In: 30º Reunião Anual da Sociedade Brasileira de Química, 2007.

RESUMOS EXPANDIDOS PUBLICADOS EM ANAIS DE EVENTOS

1. DAMASCENO, D., GODINHO, M. S., SOARES, M. H. F. B., OLIVEIRA, A. E. Caracterização Acadêmica dos Professores de Química do Ensino Médio do Estado de Goiás utilizando Análise de Componentes Principais In: V Congresso de Pesquisa, Ensino e Extensão – CONPEEX, 2008.
2. DAMASCENO, D., GODINHO, M. S., SOARES, M. H. F. B., OLIVEIRA, A. E. Uso da Análise de Componentes Principais para estudar a evolução do quadro de professores de Química no Ensino Médio de Goiás entre os anos de 2003 e 2007 XV, Encontro Centro-Oeste de debates de Ensino de Química, 2007.

TRABALHOS RESUMIDOS PUBLICADOS EM ANAIS DE EVENTOS

1. GODINHO, M. S., LEMES, M. A., OLIVEIRA, A. E., SENA, M. M. Previsão da Tensão interfacial de Óleos Isolantes usando Análise de Imagens e métodos de Calibração de ordem superior - 15^o ENQA e 3^o CIAQA, Salvador, 2009.
2. GODINHO, M. S., LEMES, M. A., OLIVEIRA, A. E., SENA, M. M. Previsão da Tensão interfacial de Óleos Isolantes usando Espectroscopia no Infravermelho Próximo e Calibração Multivariada – 32^a Reunião Anual da Sociedade Brasileira de Química, Fortaleza, 2009.
3. GODINHO, M. S., DAMASCENO, D., SENA, M. M., OLIVEIRA, A. E. Avaliação da degradação de papel tipo Kraft através do tratamento quimiométrico de imagens digitalizadas - 31^a Reunião Anual da Sociedade Brasileira de Química, 2008.
4. DAMASCENO, D., GODINHO, M. S., SOARES, M. H. F. B., OLIVEIRA, A. E. Comparação do Perfil Profissional dos Professores de Química do Estado de Goiás com os Resultados do Vestibular usando PCA - XIV Encontro Nacional de Ensino de Química, 2008
5. DAMASCENO, D., GODINHO, M. S., SOARES, M. H. F. B., OLIVEIRA, A. E. Perfil dos professores de Química do Ensino Médio do estado de Goiás utilizando Análise de Componentes Principais - 31^a Reunião da Sociedade Brasileira de Química, Águas de Lindóia, 2008.
6. GODINHO, M. S., OLIVEIRA, A. E., SÁ de, A. L. C. R. Avaliação da degradação de papel tipo Kraft através de análise de imagens In: 30^a Reunião da Sociedade Brasileira de Química, Águas de Lindóia, 2007.
7. GODINHO, M. S. Análise de Componentes Principais no estudo de dados estatísticos em Ensino de Química In: 29^a Reunião Anual da Sociedade Brasileira de Química, Águas de Lindóia, 2006.
8. GODINHO, M. S. Aplicação de Métodos de Reconhecimento de Padrões na Diferenciação entre Soja Transgênica e Convencional In: XII ENEQ, III ECOQ, XIII ECODEQ, V SEMANA DO QUÍMICO (UFG, UEG, UnB, CEFET-GO), Goiânia, 2004.

RESUMO

DETERMINAÇÃO DA TENSÃO INTERFACIAL DE ÓLEOS ISOLANTES USANDO ANÁLISE DE IMAGENS E ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO

Autora: Mariana da Silva Godinho

Orientador: Prof. Dr. Marcelo Martins de Sena

Co-orientador: Prof. Dr. Anselmo Elcana de Oliveira

Transformadores são máquinas de grande importância na transmissão e distribuição de energia elétrica. Um dos seus componentes mais importantes é o sistema isolante, constituído pelo papel tipo Kraft imerso em óleo mineral, e um dos parâmetros medidos para avaliar o estado de degradação desse sistema é a tensão interfacial do óleo. O objetivo deste trabalho foi determinar a tensão interfacial em amostras de óleos isolantes usando Espectroscopia no Infravermelho Próximo (NIR) e Análise de Imagens, juntamente com métodos de calibração de primeira (PLS, mínimos quadrados parciais) e segunda ordem (PARAFAC, análise de fatores paralelos, e NPLS, PLS multilinear). Foram obtidas 48 amostras, cujos valores de tensão interfacial já haviam sido determinados. Em seguida, foram registrados espectros na faixa de 780 a 2500 nm. As amostras foram divididas em um conjunto de calibração (38) e um de validação (10). O melhor modelo PLS foi construído usando a faixa espectral de 1360 a 1430 nm e 8 variáveis latentes, produzindo um erro médio quadrático de previsão (RMSEP) de 2,1 dina/cm. Foram obtidas imagens dos mesmos conjuntos de amostras em um *scanner* comum, as quais foram convertidas para escala de cinzas e agrupadas formando arranjos de dados tridimensionais, os quais foram tratados com os métodos PARAFAC e NPLS. O melhor modelo PARAFAC foi construído com 6 fatores e forneceu um RMSEP de 3,9 dina/cm, enquanto o melhor modelo NPLS foi construído com 3 variáveis latentes e mostrou um RMSEP de 3,0 dina/cm. O modelo NIR-PLS forneceu os melhores resultados, com erros de previsão entre -7,9 e 10,2%, considerados aceitáveis pelas empresas de energia elétrica. Os métodos propostos são rápidos e não destrutivos, mostrando grandes vantagens sobre os métodos tradicionais, os quais demandam muito tempo, além de gerar uma grande quantidade de resíduos. Este trabalho abre perspectivas para um monitoramento direto (*on-line*) do sistema isolante

ABSTRACT

DETERMINATION OF INTERFACIAL TENSION IN INSULATING OILS USING IMAGE ANALYSIS AND NEAR INFRARED SPECTROSCOPY

Author: Mariana da Silva Godinho

Adviser: Dr. Marcelo Martins de Sena

Co-Adviser: Dr. Anselmo Elcana de Oliveira

Power transformers are of great importance in the transmission and distribution of electrical energy. One of their most important parts is the insulating system, consisting of Kraft paper immersed in insulating oil. One of the parameters used for evaluating the degradation of this system is the oil interfacial tension. The aim of this study was to determine the interfacial tension in samples of insulating oils by using Near Infrared Spectroscopy (NIRS) and Image Analysis, combined with first-order (PLS, partial least squares) and second-order (PARAFAC, parallel factor analysis, and NPLS, N-way PLS) calibration methods. Forty eight oil samples were obtained, whose values of interfacial tension had already been determined. These samples were divided in calibration (38) and validation (10) sets, and NIR spectra were recorded from 780 to 2500 nm. The best PLS model was built using the spectral range from 1360 to 1430 nm and eight latent variables, providing a root mean square error of prediction (RMSEP) of 2.1 dyne/cm. Scanner images were also obtained for these same sets of samples, converted for gray-scale and stacked to three-way data arrays, before applying PARAFAC and NPLS. The best PARAFAC model was built with six factors and provided a RMSEP of 3.9 dyne/cm, while the best NPLS model was obtained with three latent variables and provided a RMSEP of 3.0 dyne/cm. NIRS-PLS was considered the best model, providing individual prevision errors between -7.9 and 10.2%, which were acceptable for electric energy companies. The proposed methods are rapid and non-destructive, showing great advantages over the traditional ones, which are slow and produce chemical waste. This work opens perspectives for developing an on-line monitoring of the insulating transformer systems.

SUMÁRIO

CAPÍTULO 1 TRANSFORMADORES.....	1
1. TRANSFORMADORES.....	2
1.1. PAPEL TIPO KRAFT.....	3
1.2. ÓLEO MINERAL ISOLANTE.....	5
1.3. MECANISMO DE DEGRADAÇÃO DO SISTEMA ISOLANTE.....	6
1.4. ENSAIOS FÍSICO-QUÍMICOS.....	9
1.4.1. Teor de água.....	9
1.4.2. Rigidez dielétrica.....	9
1.4.3. Fator de potência.....	10
1.4.4. Índice de Neutralização.....	10
1.4.5. Tensão interfacial.....	11
1.5. OBJETIVO.....	12
CAPÍTULO 2 TÉCNICAS ANALÍTICAS.....	13
2. TÉCNICAS ANALÍTICAS.....	14
2.1. ANÁLISE DE IMAGENS.....	14
2.1.1. Cor.....	15
2.1.2. Sistema de cor.....	15
2.1.2.1. Sistema RGB.....	15
2.1.2.2. Sistema HSI.....	16
2.1.3. Tratamento de imagens.....	17
2.2. ESPECTROSCOPIA NO INFRAVERMELHO (IV).....	19
2.2.1. Fundamentos da espectroscopia no infravermelho próximo.....	20
2.2.2. Características da espectroscopia no infravermelho próximo.....	23

CAPÍTULO 3 MÉTODOS QUIMIOMÉTRICOS.....	25
3. MÉTODOS QUIMIOMÉTRICOS.....	26
3.1. ORGANIZAÇÃO DOS DADOS.....	26
3.2. PRÉ-TRATAMENTO DOS DADOS ESPECTRAIS.....	28
3.2.1. Centrar os dados na média.....	29
3.2.2. Aplicação de Derivadas.....	30
3.2.3. Correção de espalhamento multiplicativo (MSC).....	30
3.2.4. Alisamento Savitzky-Golay.....	31
3.3. ANÁLISE DE COMPONENTES PRINCIPAIS.....	31
3.4. CALIBRAÇÃO MULTIVARIADA.....	34
3.4.1. Regressão por mínimos quadrados parciais – PLS.....	35
3.4.2. Escolha do Número de Variáveis Latentes.....	37
3.4.3. Previsões do Modelo.....	38
3.5. MÉTODOS DE SEGUNDA ORDEM.....	39
3.5.1. PARAFAC.....	39
3.5.2. NPLS.....	41
CAPÍTULO 4 PARTE EXPERIMENTAL.....	44
4. PARTE EXPERIMENTAL.....	45
4.1. AMOSTRAGEM.....	45
4.2. OBTENÇÃO DAS MEDIDAS.....	46
4.2.1. Espectroscopia no Infravermelho Próximo de óleos isolantes.....	46
4.2.2. Análise de imagens de óleos isolantes.....	47
4.3. TRATAMENTO DOS DADOS.....	49
CAPÍTULO 5 RESULTADOS E DISCUSSÃO.....	52

5. RESULTADOS E DISCUSSÃO.....	53
5.1 ESPECTROSCOPIA NIR.....	53
5.2. ANÁLISE DE IMAGENS.....	57
5.2.1. Modelo PARAFAC.....	58
5.2.2. Modelo NPLS.....	60
5.3. OUTROS MÉTODOS DE TRATAR IMAGENS.....	63
5.3.1. Histogramas em escalas de cinza.....	63
5.3.2. Histogramas RGB.....	64
5.3.3. PLS desdobrado.....	66
CAPÍTULO 6 CONCLUSÕES E PERSPECTIVAS.....	67
6.1. CONCLUSÃO.....	68
6.2. PERSPECTIVAS.....	69
REFERÊNCIAS BIBLIOGRÁFICAS.....	70
APÊNDICE.....	77

LISTA DE TABELAS

Tabela 1	–	Composição dos óleos parafínicos e naftênicos.....	6
Tabela 2	–	Divisão das regiões espectrais do Infravermelho.....	19
Tabela 3	–	Variância explicada pelos componentes principais.....	34
Tabela 4	–	Tensão interfacial quantificada nas amostras de óleos e usada para construir os modelos de calibração para dados espectrais e de imagens.....	53
Tabela 5	–	Valores de tensão interfacial previstos pelo modelo PLS e erros percentuais.....	56
Tabela 6	–	Valores de tensão interfacial previstos pelo modelo PARAFAC e erros percentuais.....	58
Tabela 7	–	Variância explicada pelas matrizes \underline{X} e \underline{Y} para o modelo NPLS para 3 fatores.....	61
Tabela 8	–	Valores de tensão interfacial previstos pelo modelo NPLS e erros percentuais.....	61
Tabela 9	–	Valores de tensão interfacial previstos pelo modelo PLS dos histogramas de freqüências para as imagens em escala de cinza e erros percentuais.....	63
Tabela 10	–	Valores de tensão interfacial previstos pelo modelo PLS dos histogramas de freqüências para os canais RGB e erros percentuais.....	65

LISTA DE FIGURAS

Figura 1 – Vista interna de um transformador de potência (A); núcleo de um transformador e o isolante sólido (papel tipo Kraft) (B).....	2
Figura 2 – Estrutura molecular da celulose.....	4
Figura 3 – Representação das cadeias dos diferentes grupos de hidrocarbonetos que compõem o óleo mineral isolante.....	5
Figura 4 – Esquema de oxidação do óleo mineral isolante.....	8
Figura 5 – Cubo de cores RGB, em que cada eixo varia de 0 a 255 índices de cor.....	16
Figura 6 – Representação de cor no sistema HSI.....	17
Figura 7 – Representação do perfil da energia potencial do modelo do oscilador harmônico.....	21
Figura 8 – Representação do perfil da energia potencial para o modelo do oscilador anarmônico.....	22
Figura 9 – Ilustração das transições entre os níveis de energia em um oscilador anarmônico.....	23
Figura 10 – Dados espectrais arranjados na forma de uma matriz X	27
Figura 11 – Arranjo de dados de imagens.....	27
Figura 12 – Representação tridimensional de dados de imagens.....	28
Figura 13 – Representação gráfica dos dados centrados na média. (A) dados antes da centralização; (B) dados centrados na média.....	30
Figura 14 – Representação das duas primeiras componentes principais em um sistema multidimensional.....	31
Figura 15 – Representação da decomposição da matriz de dados X em escores e pesos.....	32
Figura 16 – Representação dos eixos das componentes principais (A); projeção dos pesos (B); projeção dos escores (C).....	33
Figura 17 – Decomposição das matrizes X e Y em variáveis latentes para	

	o modelo PLS.....	36
Figura 18 –	Esquema da decomposição trilinear do modelo PARAFAC em tríades.....	40
Figura 19 –	Representação esquemática da decomposição de um arranjo multidimensional para o modelo NPLS.....	42
Figura 20 –	Representação esquemática da decomposição da matriz Y para o modelo NPLS.....	42
Figura 21 –	Tensiômetro de torção Krüss.....	46
Figura 22 –	Espectrofotômetro utilizado para a obtenção dos espectros das amostras de óleo.....	47
Figura 23 –	Imagens das 48 amostras de óleos digitalizadas usadas na construção do modelo.....	48
Figura 24 –	<i>Scanner</i> e a placa de petri utilizados na obtenção das imagens de óleos.....	49
Figura 25 –	Esquema do tratamento das imagens e montagem dos arranjos tridimensionais.....	51
Figura 26 –	Gráfico de escores para os dados espectrais.....	54
Figura 27 –	Espectros das 48 amostras de óleos na região NIR (a). Faixa espectral selecionada de 1360 a 1430 nm (b).....	55
Figura 28 –	Valores experimentais versus valores previstos pelo modelo PLS.....	57
Figura 29 –	Gráfico de pesos para o modelo PARAFAC.....	59
Figura 30 –	Valores experimentais versus valores previstos pelo modelo PARAFAC.....	60
Figura 31 –	Valores experimentais versus valores previstos pelo modelo NPLS.....	62
Figura 32 –	Valores experimentais versus valores previstos pelo modelo PLS para os histogramas de freqüência das imagens em escala de cinza.....	64
Figura 33 –	Valores experimentais versus valores previstos pelo modelo PLS para os histogramas de freqüência dos canais RGB.....	65

Capítulo 1

Transformadores



1. TRANSFORMADORES

Transformadores são máquinas elétricas estáticas que transferem energia elétrica de um circuito para outro mantendo a mesma frequência e normalmente variando os valores de corrente e de tensão [1]. Esses dispositivos são de grande importância, pois qualquer que seja a matriz energética (nuclear, hídrica ou fóssil), a transmissão e a distribuição da energia gerada requer seu uso. Essas máquinas são constituídas por materiais como aço, cobre, vernizes, tintas e, principalmente, papel e óleo mineral, responsáveis pelo seu sistema isolante [2]. A parte ativa dessas máquinas é composta pelo núcleo, enrolamento, comutadores e condutores de ligação. O núcleo é um circuito magnético fechado, constituído de chapas de aço e isolado por camadas de papel isolante. Os enrolamentos são constituídos por bobinas de cobre ou alumínio também isoladas por camadas de papel isolante. O conjunto desses elementos ativos é colocado dentro de um tanque contendo óleo mineral isolante [3]. A Figura 1 mostra a parte interna de um transformador e o seu núcleo, onde se encontra parte do sistema isolante.

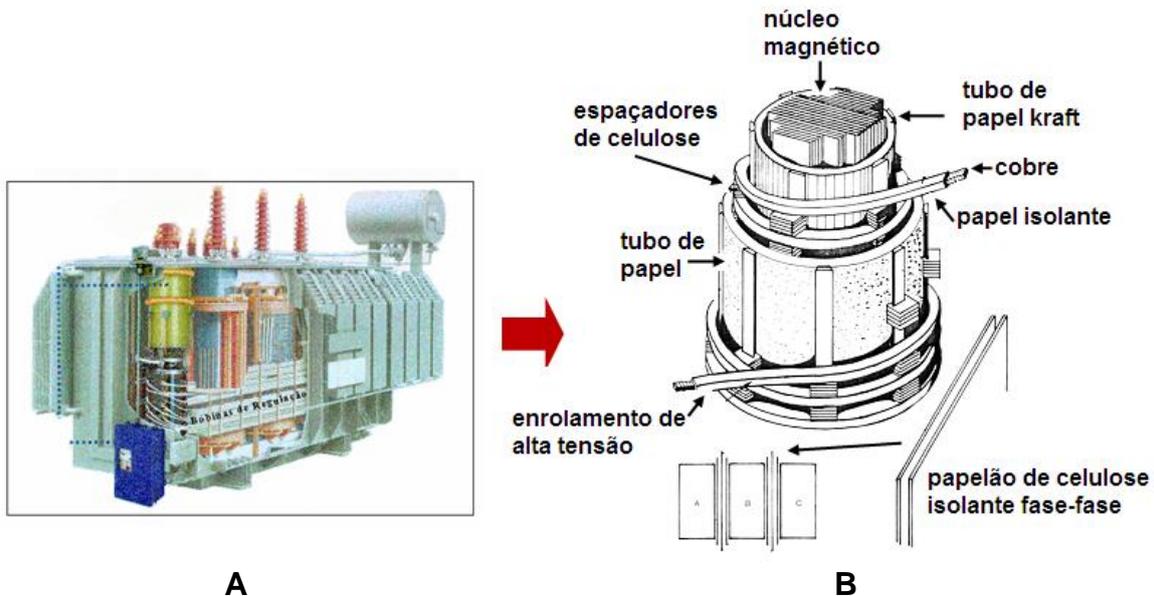


Figura 1 – Vista interna de um transformador de potência (A); núcleo de um transformador e o isolante sólido (papel tipo Kraft) (B).

Além de possuírem um preço elevado, essas máquinas são os componentes críticos de qualquer subestação, pois seu desligamento causa uma série de problemas ao sistema elétrico e à produção industrial. A demanda crescente por energia tem acarretado sobrecarga nos sistemas de potência, levando à necessidade de ferramentas adequadas de monitoramento e diagnóstico do estado operativo desses sistemas, com o objetivo de aumentar sua confiabilidade e reduzir o número de paradas programadas ou eventuais, otimizando dessa forma o fornecimento de energia elétrica.

Durante o funcionamento dos transformadores, vários processos de desgaste e envelhecimento ocorrem em seu sistema de isolamento. Desses processos decorrem a fadiga térmica, química, elétrica e mecânica, tais como pontos quentes, sobre-aquecimentos, sobre-tesões e vibrações, sendo estes responsáveis por alterações no sistema isolante (tanto papel, quanto óleo). Dessa forma, o sistema isolante deve ser monitorado para garantir a eficiência do equipamento e prolongar seu tempo de vida útil. Assim, o tempo de operação de um transformador é determinado pelo estado de degradação desse sistema isolante, podendo ultrapassar 40 anos de vida útil [4].

1.1. PAPEL TIPO KRAFT

O principal isolante sólido usado nos transformadores é de natureza celulósica, o papel extraído da madeira ou de outros vegetais, sendo que os mais empregados no isolamento dos transformadores são:

- Papel e papelão Kraft – feito de fibra de madeira;
- Papel Manilha – feito de fibras de madeira e cânhamo;
- *Pressboard* – papelão feito de fibras de algodão.

O papel tipo Kraft [4-7] é o principal isolante sólido, sendo constituído por fibras de celulose, hemicelulose e ligninas. A celulose é um polímero natural constituído por anéis de glicose unidos através de interações covalentes envolvendo os grupos hidroxila, conforme pode ser observado na Figura 2. As hemiceluloses são polímeros ramificados e de cadeias menores constituídos por vários tipos de unidades de açúcares. As ligninas são polímeros amorfos de composição química

um pouco mais complexa, diferindo da celulose e da hemicelulose pela alta reatividade frente a agentes oxidantes.

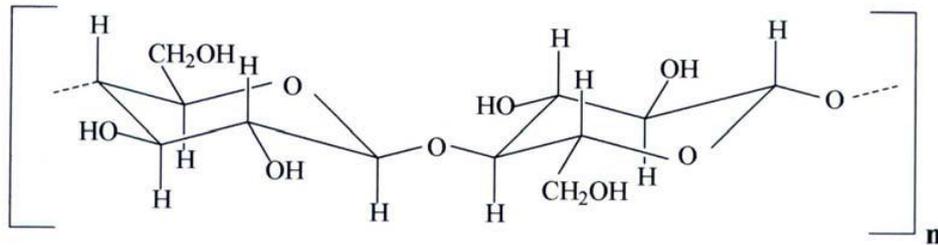


Figura 2 – Estrutura molecular da celulose.

O papel tipo Kraft é produzido usualmente pelo processo Kraft, cuja função é retirar o máximo de ligninas para que o papel se torne mais resistente. O comprimento médio das moléculas de celulose é determinado pelo de grau de polimerização (GP), o qual exprime a média do número de anéis de glicose por molécula de celulose, situando-se na faixa de 1000 a 1400 para o papel novo [8,9]. A secagem e o condicionamento dos enrolamentos do transformador reduzem o GP da celulose a um valor entre 900 e 1000. Para valores de GP superiores a 500 a resistência mecânica do papel é praticamente independente de seu valor, mas à medida que o GP decresce na faixa de 500 a 200, a resistência mecânica é reduzida até cerca de 50% (algumas vezes até 20%) do valor inicial. Para valores de GP abaixo de 200 a resistência mecânica do papel é quase nula, e este se torna frágil e desagrega em pequenos fragmentos, que originam as indesejáveis “lamas” (resíduos sólidos provenientes da degradação do papel), usualmente encontradas no óleo em situações de degradação avançada do papel [10].

O envelhecimento de um transformador está diretamente relacionado com a resistência mecânica do papel isolante, a qual tem a característica de se alterar em função do tempo.

1.2. ÓLEO MINERAL ISOLANTE

Os óleos isolantes são oriundos da destilação de petróleo natural, da fração de 300 a 400°C, sendo constituídos por hidrocarbonetos e podem ser classificados em três grupos: parafínicos (hidrocarbonetos saturados de cadeia aberta, linear ou ramificada), naftênicos (hidrocarbonetos saturados de cadeia fechada) e aromáticos (hidrocarbonetos que contém um ou mais anéis aromáticos) [3,4]. A Figura 3 representa exemplos desses três tipos de cadeias.

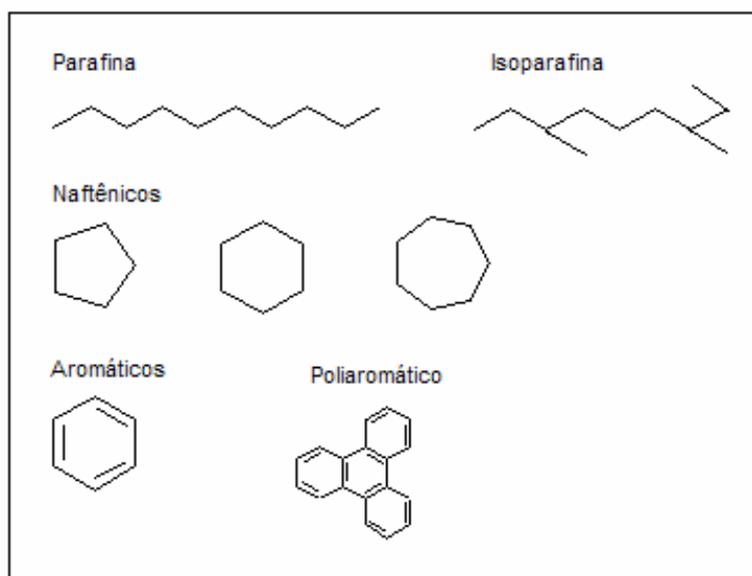


Figura 3 – Representação das cadeias dos diferentes grupos de hidrocarbonetos que compõem o óleo mineral isolante [6].

Além de hidrocarbonetos, no óleo isolante encontram-se compostos orgânicos de enxofre termicamente estáveis que são inibidores naturais do processo de oxidação, podendo ser adicionados inibidores sintéticos como o di-terc-butil-*p*-cresol (DBPC) em teores que variam de 0,08 a 0,33%. A classificação dos óleos isolante em parafínicos ou naftênicos é feita através de medidas de viscosidade, índices de refração e pela correlação desses resultados através de um diagrama ternário. Os resultados indicam o percentual de carbono existente em cadeias parafínicas, naftênicas e aromáticas. Na Tabela 1 pode ser observada a composição dos óleos parafínicos e naftênicos [4].

Tabela 1 – Composição dos óleos parafínicos e naftênicos [4].

Tipo de óleo	Compostos Aromáticos (%)	Compostos Naftênicos (%)	Compostos Parafínicos (%)
Naftênico	11	47	42
Parafínico	8	29	63

Outra forma de classificação dos óleos isolantes é através da espectroscopia na região do infravermelho.

Os óleos isolantes parafínicos e naftênicos apresentam algumas propriedades físicas muito diferentes entre si, dentre elas, densidade, ponto de anilina e fluidez. Embora os óleos parafínicos apresentem ponto de fluidez inferior aos de base naftênica, ambos apresentam bons desempenhos, pois o que é importante para o óleo é que ele possua boas condições dielétricas, independente de sua natureza. Portanto, os óleos isolantes nos transformadores exercem duas funções distintas: uma de natureza isolante e a outra é a transferência do calor produzido pela parte ativa do tanque para as paredes. O óleo isolante mineral exerce muito bem essas funções, pois apresenta uma elevada capacidade térmica, bom coeficiente de transmissão de calor, ótimas propriedades isolantes e rigidez dielétrica quase cinco vezes maior que o ar.

1.3. MECANISMO DE DEGRADAÇÃO DO SISTEMA ISOLANTE

Durante a operação do transformador o óleo e o papel passam por um processo de envelhecimento, que pode ser afetado por variáveis como, temperatura, umidade e presença de oxigênio [11]. O processo de degradação do papel envolve o rompimento das ligações glicosídicas que mantêm unidos os anéis de glicose. Como produtos dessa degradação são produzidos compostos furânicos (5-metil-2-furfuraldeído, 2-acetilfurano, 5-hidroximetilfur-furaldeído, álcool furfurílico e 2-furfuraldeído), dentre os quais o 2-furfuraldeído é o mais importante, pelo fato de ser o composto mais estável e solúvel no óleo, além de ser um produto específico da degradação do papel [10,12]. Em temperaturas baixas de operação, ocorre a

desidratação (formação de água) da celulose, enquanto que em temperaturas elevadas há produção de levoglucosano, que se decompõe formando os compostos furânicos.

Além da produção desses compostos, a degradação da celulose também gera compostos gasosos como hidrogênio (H_2), metano (CH_4), etileno (C_2H_2) e principalmente monóxido (CO) e dióxido de carbono (CO_2) além de água (H_2O). A produção desses compostos é explicada através de complexas seqüências de reações químicas, que provocam cisão das cadeias poliméricas [13].

A degradação do óleo isolante ocorre com o comprometimento de suas propriedades isolantes e os produtos formados aceleram a degradação da celulose. O mecanismo que rege a oxidação dos hidrocarbonetos é a peroxidação que gera como produtos hidroperóxidos. Estes, de acordo com sua espécie irão formar alcoóis, aldeídos e cetonas, os quais são chamados de produtos intermediários, pois na presença de oxigênio irão dar origem aos ácidos carboxílicos. Estes produtos são moléculas polares que agredem o papel isolante e os demais materiais integrantes dos transformadores. Na Figura 4 é mostrado o esquema de oxidação do óleo mineral isolante.

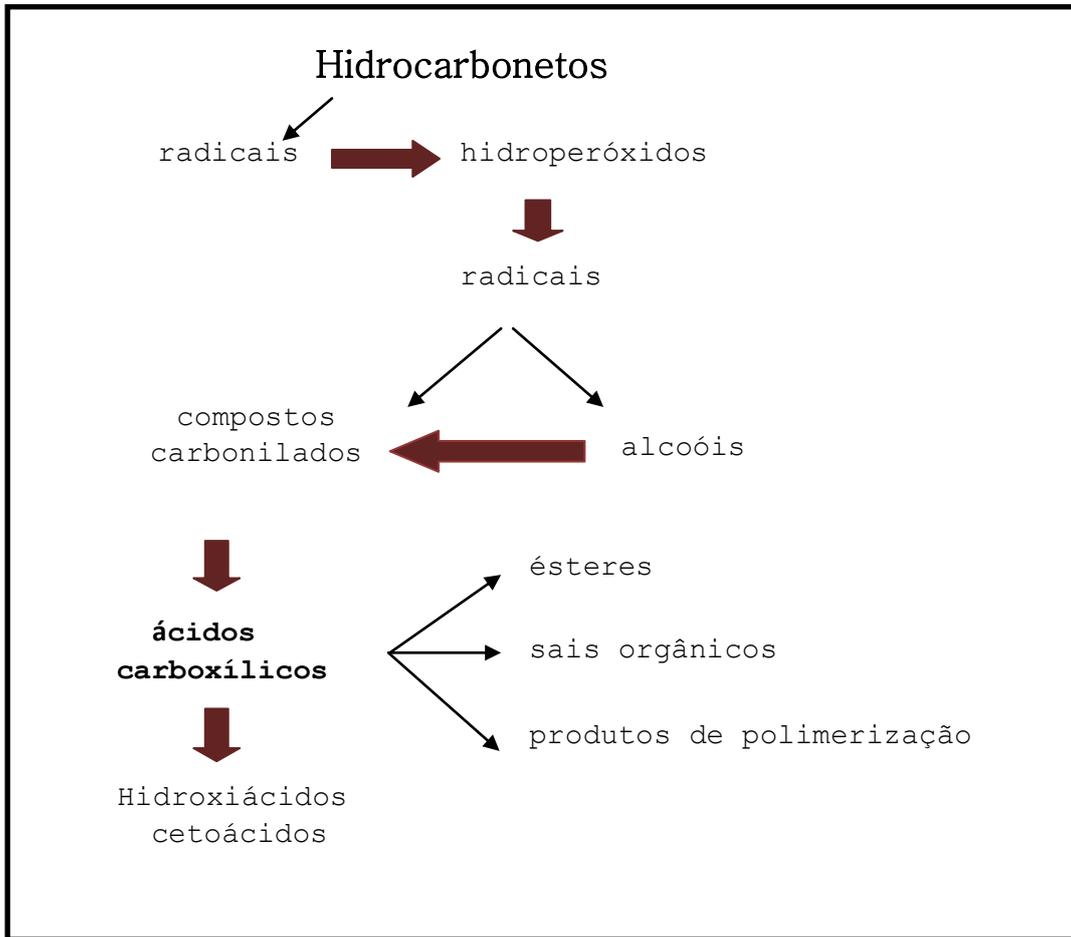


Figura 4 – Esquema de oxidação do óleo mineral isolante.

No monitoramento do sistema isolante são utilizadas técnicas analíticas para avaliar as condições de degradação do sistema e ensaios que medem as propriedades físico-químicas do óleo. As técnicas mais utilizadas para avaliar o estado de degradação do sistema são: análise de gás dissolvido (DGA, *Dissolved Gas Analysis*), medida do grau de polimerização (DP, *Degree of Polymerization*) e análise de furanos [14-16].

As propriedades físico-químicas do óleo mineral são alteradas em decorrência do processo contínuo de oxidação ao qual o óleo encontra-se submetido, sendo que esse processo pode ser acelerado pela presença de compostos gasosos e líquidos que se dissolvem no óleo, ocasionando alterações que podem comprometer o funcionamento do sistema. Essas alterações devem ser acompanhadas por ensaios físico-químicos periódicos. O objetivo desses ensaios é

avaliar a cor, umidade, rigidez dielétrica, fator de potência, tensão interfacial, dentre outros parâmetros [17]. Estes ensaios, apresentados a seguir, traduzem as condições imediatas do óleo em uso, sendo complementados com outras análises.

1.4. ENSAIOS FÍSICO-QUÍMICOS

1.4.1. Teor de água

O teor de umidade presente em sólidos e líquidos isolantes varia de acordo com a estrutura e o funcionamento do transformador. Embora parte da água presente no óleo mineral, seja formada como subproduto de sua oxidação, a maior parte da água existente no óleo é absorvida do ar. O controle do teor de umidade é importantíssimo, uma vez que sua ação acarreta:

- diminuição da rigidez dielétrica do óleo;
- hidrólise da celulose presente no papel;
- corrosão dos metais constituinte do transformador.

Os ciclos de temperatura que ocorrem durante o tempo de operação fazem com que o teor de água em equilíbrio entre óleo e celulose se altere. Quando ocorre o resfriamento devido à queda de demanda ou no período noturno, a água se condensa podendo migrar para o fundo do tanque ou ser absorvida pela celulose. Quando ocorre o aumento de temperatura, a água se dissolve no óleo. Dessa forma, quanto menor o teor de água encontrado nos óleos em operação, mais protegida estará a isolação sólida e melhor serão as características dielétricas do óleo [18].

1.4.2. Rigidez dielétrica

A rigidez dielétrica é uma das propriedades elétricas do óleo isolante. Ela é expressa pela tensão em KV, sendo medida através da aplicação de uma rampa de potencial entre dois eletrodos imersos na amostra e observando o valor no qual ocorre uma descarga elétrica. A determinação dessa propriedade em óleos minerais isolantes tornou-se o primeiro procedimento obrigatório em qualquer laboratório de controle de qualidade desses materiais.

Os resultados dos ensaios de rigidez dielétrica estão sempre associados à presença de água e impurezas sólidas microscópicas, mesmo em pequenas quantidades. O valor dessa propriedade é pouco afetado pela água dissolvida no óleo. Por outro lado, a água livre em suspensão diminui acentuadamente sua rigidez dielétrica. No entanto, a quantificação dessa propriedade não pode ser utilizada para avaliar o processo de envelhecimento do óleo mineral, pois a presença de produtos de oxidação do óleo, como álcoois, aldeídos e cetonas, não afeta os valores de rigidez dielétrica. A quantificação desse parâmetro em um óleo novo é um indicativo das condições de manuseio, transporte e armazenagem a que este foi submetido, enquanto que para óleos em serviço, esta medida reflete as condições de operação do equipamento [19].

1.4.3. Fator de Potência

O fator de potência, ou fator de perdas dielétricas, é outra propriedade importante avaliada no óleo isolante. O aumento do seu valor normalmente está associado à presença de substâncias de alta condutividade, sendo um forte indicativo de contaminação com substâncias metálicas finamente divididas e produtos de degradação da celulose e do óleo propriamente dito.

O fator de potência pode ser designado como o valor da tangente ou seno do ângulo de fase entre a tensão e corrente quando se aplica uma tensão a 60 Hz no óleo a ser analisado. Essa propriedade é uma medida de qualidade dielétrica do óleo e quanto menor o seu valor, mais próximo o sistema estará do ideal, ou seja, um sistema não degradado. Sendo assim, o acompanhamento da variação do fator de potência pode determinar a qualidade do sistema isolante em termos de envelhecimento do óleo [20].

1.4.4. Índice de Neutralização

O índice de neutralização é a quantificação do teor de ácidos formados por oxidação do sistema isolante. A presença desses compostos pode provocar ataque químico aos diversos materiais utilizados na construção dos

transformadores, em especial os metais, ocasionando a formação de produtos que afetam as características dielétricas do óleo. O alto teor de compostos ácidos no óleo é responsável diretamente pela formação de borra (resíduos sólidos). Assim, através desse ensaio é possível acompanhar o envelhecimento do sistema isolante, pois após iniciado o processo de deterioração do transformador a acidez aumenta rapidamente, podendo comprometer a vida útil do equipamento [21].

1.4.5. Tensão Interfacial

Na interface de separação entre dois líquidos imiscíveis ou entre um líquido e um gás forma-se uma força de atração entre as moléculas, denominada de tensão interfacial [22]. Essa propriedade é afetada pela presença de substâncias polares, mesmo em pequenas quantidades, pois estas além de afetar as propriedades elétricas do óleo, possuem centros ativos que aceleram o seu envelhecimento. A medida da tensão interfacial é representada pela força necessária para que um anel de platina rompa a interface água/óleo.

A medida desse parâmetro no sistema água-óleo visa exatamente inferir a quantidade de substâncias polares presentes no óleo, uma vez que tais substâncias tendem a concentrar-se na interface do sistema. Quanto maior a concentração de substâncias polares, menor o valor da tensão interfacial. Dessa forma, a quantificação dessa propriedade é utilizada como um meio de avaliar e monitorar o estado de oxidação do óleo e também avaliar indiretamente o estado de degradação do papel, uma vez que esses compostos polares são oriundos da oxidação tanto do óleo quanto do papel. O valor limite para essa propriedade é de aproximadamente 45 dina/cm para um óleo novo e de 17 dina/cm para um óleo que deve ser descartado, porém a avaliação desses valores limites depende do laboratório em que são feitas as análises e do sistema [23].

1.5. OBJETIVO

A tensão interfacial é o parâmetro físico-químico de maior correlação com o estado de oxidação do óleo e, indiretamente, também do papel, pois nesse ensaio é avaliada a concentração dos subprodutos de degradação do sistema isolante. Dessa forma, o presente trabalho tem como objetivos prever essa propriedade físico-química em amostras de óleos minerais isolantes usando as técnicas de análise de imagens e espectroscopia na região do infravermelho próximo juntamente com métodos de calibração multivariada e multidimensional. O método de referência [22] utiliza um tensiômetro de torção e apresenta as desvantagens de ser relativamente demorado e possuir baixa precisão ($s \cong 10\%$). Os modelos de calibração serão obtidos usando métodos de primeira ordem para os dados de Infravermelho e de segunda ordem para os dados de imagens, propondo uma metodologia mais simples e rápida, que possa ser utilizada diretamente na avaliação do sistema isolante (papel-óleo).

Capítulo 2



Técnicas analíticas

2. TÉCNICAS ANALÍTICAS

2.1. ANÁLISE DE IMAGENS

Imagens são representações de objetos ou de situações reais [24]. Essas representações podem ser expressas como uma função matemática, portanto a imagem pode ser definida em termos de arranjos numéricos de dados representados por uma função bidimensional do tipo $z = f(x, y)$ [25], definida sobre certa região de um plano, onde x e y denotam as coordenadas espaciais, e o valor de f em qualquer ponto (x, y) é proporcional ao brilho da imagem naquele ponto [26]. As imagens ao serem capturadas pelo olho humano ou por qualquer aparelho são de natureza contínua, ou seja, são imagens caracterizadas por ondas eletromagnéticas. Para que estas possam ser armazenadas e manipuladas por um computador é necessário que cada imagem seja transformada em uma seqüência de *bits*, que têm natureza discreta. No entanto, essa função é muito complexa para uma imagem contínua tornando sua obtenção e tratamento tarefas de difícil execução. Dessa forma, essa função contínua que descreve a imagem é convertida em uma função discreta através da digitalização da imagem [27].

No processo de digitalização, uma imagem contínua é transformada em uma imagem digital, que consiste em uma estrutura quadriculada onde cada quadrado recebe o nome de *pixel*, o qual possui um valor de intensidade correspondente.

As imagens podem ser representadas por uma matriz bidimensional, com coordenadas (i, j) e por matrizes tridimensionais de coordenadas (i, j, k) . Imagens representadas por matrizes bidimensionais são chamadas de imagens *gray-scale* (escala de cinzas) e são classificadas como imagens univariadas. Imagens tridimensionais são conhecidas como imagens coloridas e são classificadas como imagens multivariadas [28]. As imagens coloridas dependem de uma informação que é a cor.

2.1.1. Cor

A cor é uma característica muito importante para uma imagem, pois ela é uma manifestação percentual da luz incidida em uma superfície, podendo ser avaliada pela imagem digitalizada e manipulada em um computador [29]. Dois conceitos muito importantes para a compreensão da percepção da cor são a luminância e a crominância. A luminância contém a informação da quantidade de preto e branco que uma cor contém, ou seja, está relacionada com o brilho da cor. Já a crominância contém a informação da tonalidade de cor e está relacionada com a frequência do raio de luz incidente.

A cor é uma propriedade que pode ser representada por um sistema de coordenadas tridimensionais, sendo que o sistema mais utilizado na prática é o sistema RGB (*Red, Green, Blue*) usados em monitores coloridos, *scanners* e câmeras de vídeo. Outros sistemas de cor são: CMY (*Cyan, Magenta, Yellow*) usado em impressoras coloridas, HSI (*Hue, Saturation, Intensity*) e HSV (*Hue, Saturation, Value*), usados freqüentemente em manipulação e processamento de imagens coloridas [30].

2.1.2. Sistemas de cor

2.1.2.1. Sistema RGB

O Sistema de cor RGB [30-32] foi adotado pela CIE (*Commission Internationale de l'Eclairage*) em 1931 e consiste na utilização de um espaço tridimensional em que cada cor é obtida a partir das cores primárias vermelha (R, *Red*) verde (G, *Green*) e azul (B, *Blue*), sendo estas correspondentes respectivamente às frequências baixa, média e alta do espectro visível. Nesse sistema de cor cada *pixel* é representado por três *bytes* de oito *bits* cada, totalizando 24 *bits*. Como um *byte* pode expressar inteiros entre 0 e 255, teremos um total de $256 \times 256 \times 256$, ou seja, $16.777.216 \times 10^6$ combinações possíveis.

A maioria das aplicações usando imagens coloridas, tais como monitores coloridos, câmeras fotográficas e *scanner*, utiliza o sistema RGB, pois é um sistema

matemático mais simples para a representação de cor, podendo ser descrito por um cubo, conforme a Figura 5.

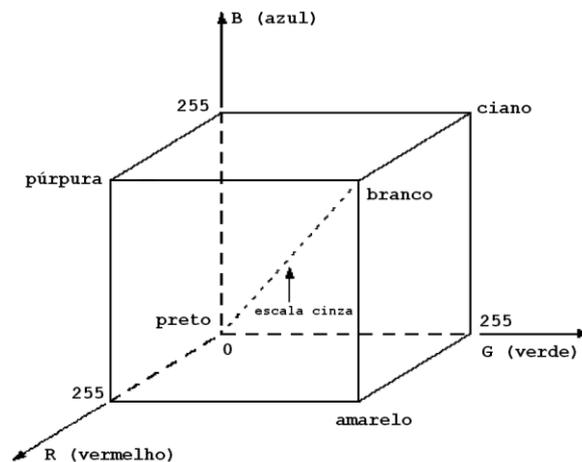


Figura 5 – Cubo de cores RGB, em que cada eixo varia de 0 a 255 índices de cor.

2.1.2.2. Sistemas HSI

Embora o sistema RGB seja muito utilizado em dispositivos de captação de imagens, outros sistemas buscam uma forma mais próxima da manipulação natural de cor utilizada pelo homem. Dentre esses sistemas, o HSI e o HSV trabalham com os conceitos de tonalidade, saturação e luminância. O conceito de tonalidade representa a cor propriamente dita, enquanto a saturação representa a sua pureza, ou seja, reflete o quão forte é a cor e a intensidade identifica a luminância da cor no sistema HSI [30,31].

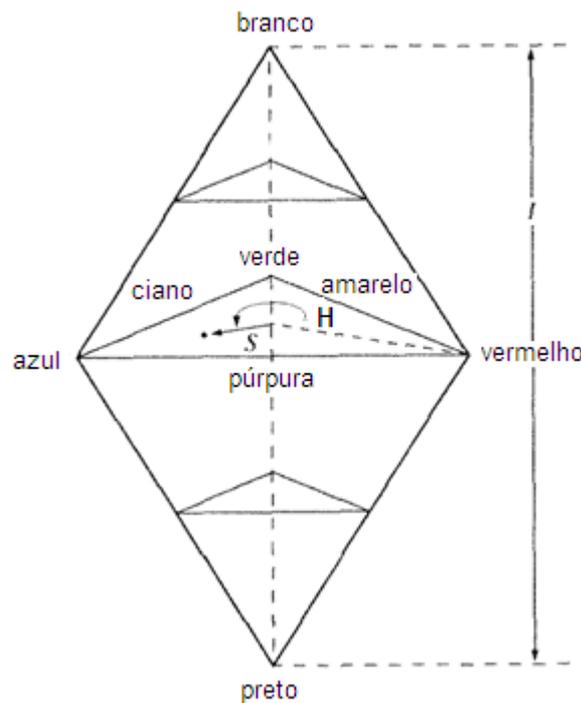


Figura 6 – Representação de cor no sistema HSI.

Nesse sistema o valor de H é dado em graus ($^{\circ}$), dependendo do ângulo θ em relação à cor de partida, o vermelho, como pode ser observado na Figura 6. O valor de intensidade (I) varia no eixo vertical, tendo como extremidades as cores branca e preta, sendo que a saturação depende das misturas entre as cores vermelho, verde e azul. Esse sistema de cor é muito propício em situações em que há a necessidade de separar a cor independente da luminosidade. Entretanto, o grande problema em usar esses sistemas está na necessidade de transformação de cada pixel da imagem captada em RGB para HSI ou HSV, o que demanda tempo de processamento, desestimulando o seu uso.

2.1.3. Tratamento de imagens

As aplicações com imagens têm crescido em diversas áreas, isso devido ao rápido e crescente desenvolvimento de recursos computacionais que permitem gerenciar conjuntos de dados cada vez maiores e também devido ao desenvolvimento de meios acessíveis de captar essas imagens. Esses meios de captação podem ser câmeras de vídeo, câmeras fotográficas, *scanners*, aparelhos de raios X, microscópios eletrônicos, dentre outros [32].

Após a imagem ser captada, ela pode passar por dois tipos de operações, chamadas respectivamente de processamento de imagens e análise de imagens. O processamento de imagens é utilizado com o objetivo de melhorar a definição e ressaltar aspectos e características importantes na imagem. Portanto, trata-se de um conjunto de operações que transformam a imagem original em outra imagem. A análise de imagens é a aplicação de técnicas matemáticas com objetivo de reduzir os dados para extrair informações a respeito de um sistema analisado [33].

As técnicas de processamento de imagens podem ser utilizadas em conjunto com as técnicas de análise de imagens, ou seja, antes do conjunto de imagens ser analisado, ele pode sofrer alterações com o intuito de ressaltar certas características. Dentre as técnicas utilizadas em processamento de imagens, os filtros são as mais conhecidas, onde as transformadas de Fourier e Wavelet têm um grande destaque. Isso devido ao fato dessas transformadas apresentarem vantagens como a utilização de algoritmos muito eficientes, que possibilitam a interpretação de diferentes fatos sobre imagens de maneira unificada [31,34].

Para tratar um conjunto de dados de imagens é necessário recorrer aos métodos estatísticos multivariados. A aplicação desses métodos irá depender do espaço de cor em que as imagens foram digitalizadas e como as matrizes das imagens foram organizadas. Há várias formas de construir uma matriz de dados de imagens. A mais simples é a extração dos histogramas de frequência dos canais de cores R, G e B para imagens coloridas. Outra forma de tratar matrizes de imagens é converter as imagens coloridas para escala de cinza, ou *gray-scale*. Dessa forma, cada matriz tridimensional (imagem colorida) é transformada em uma matriz bidimensional (imagem em escala de cinza) reduzindo uma dimensão do conjunto de dados. As matrizes bidimensionais são agrupadas uma ao lado da outra formando um tensor de terceira ordem, do qual podem ser extraídas informações muito relevantes usando métodos multivariados de segunda ordem.

2.2. ESPECTROSCOPIA NO INFRAVERMELHO (IV)

A região do espectro infravermelho (IV) abrange o intervalo de comprimento de onda de 780 nm a 1.000 μm ou números de ondas de 12800 a 10 cm^{-1} [35]. Essa região pode ser dividida em três regiões distintas, que são classificadas de acordo com sua proximidade do espectro visível, sendo elas: infravermelho próximo (NIR, *Near Infrared*), infravermelho médio (MID, *Middle Infrared*) e infravermelho distante (FIR, *Far Infrared*). A Tabela 2 apresenta as características das três regiões do espectro infravermelho.

Tabela 2 – Divisão das regiões espectrais do Infravermelho.

Região	Comprimento de onda (λ),nm	Número de onda ($\bar{\nu}$), cm^{-1}	Frequência (ν), Hz
Próximo (NIR)	780 a 2500	12.800 a 4.000	$3,8 \times 10^{14}$ a $1,2 \times 10^{14}$
Médio (MID)	2500 a 5000	4.000 a 200	$1,2 \times 10^{14}$ a $6,0 \times 10^{12}$
Distante (FIR)	5000 a 10000	200 a 10	$6,0 \times 10^{12}$ a $3,0 \times 10^{11}$

Normalmente, a radiação no infravermelho não é energética o suficiente para causar transições eletrônicas, portanto sua absorção ocorre em espécies moleculares que apresentam diferenças pequenas de energia em vários estados rotacionais e vibracionais. Para que uma molécula possa absorver a radiação infravermelha é necessário que ocorra uma variação no momento de dipolo em consequência do movimento vibracional ou rotacional, sendo que o momento de dipolo é determinado pela magnitude de diferença de carga e a distância entre os dois centros de carga. Uma molécula que possui essa variação no momento dipolar, ao vibrar provoca uma variação do momento dipolar, surgindo um campo que pode interagir com o campo elétrico associado à radiação infravermelha incidente. Quando a frequência da radiação coincide exatamente com a frequência vibracional natural da molécula, ocorre uma transferência de energia, resultando em uma variação da amplitude da vibração molecular e conseqüentemente, na absorção de radiação [35,36].

2.2.1. Fundamentos da Espectroscopia no Infravermelho Próximo

A Espectroscopia no Infravermelho Próximo [35-37] é um tipo de espectroscopia vibracional que utiliza a radiação eletromagnética caracterizada pelos comprimentos de onda no intervalo de 750 a 2500 nm, sendo descoberta pelo astrônomo Frederick William Herschel em 1800.

Para que seja possível compreender a interação da radiação NIR com uma molécula é necessário estudar o comportamento vibracional das ligações químicas. O modelo do oscilador harmônico é utilizado para ilustrar a vibração dos átomos em uma ligação de uma molécula diatômica. Nesse modelo, as características da vibração de estiramento são aproximadas às de um modelo mecânico representado por duas massas ligadas por uma mola. Dessa forma, as freqüências fundamentais de dois átomos ligados podem ser calculadas assumindo que a energia obedece à Lei de Hooke, conforme a equação 1,

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \quad (1)$$

onde, k é a constante de força e μ é a massa reduzida, a qual pode ser calculada através da equação 2,

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (2)$$

onde m_1 e m_2 são as massas dos átomos envolvidos.

A energia potencial do sistema pode ser calculada em qualquer posição de deslocamento dos átomos envolvidos, de acordo com a equação 3,

$$E = \frac{1}{2} k(r - r_e)^2 = \frac{1}{2} kx^2 \quad (3)$$

onde r é a distância entre os dois núcleos atômicos, r_e é a distância internuclear de equilíbrio e x é o deslocamento dos átomos. Na Figura 7 é mostrada a representação do perfil da energia potencial em um oscilador harmônico quantizado, no qual a vibração de uma molécula apresenta um perfil contínuo de energia.

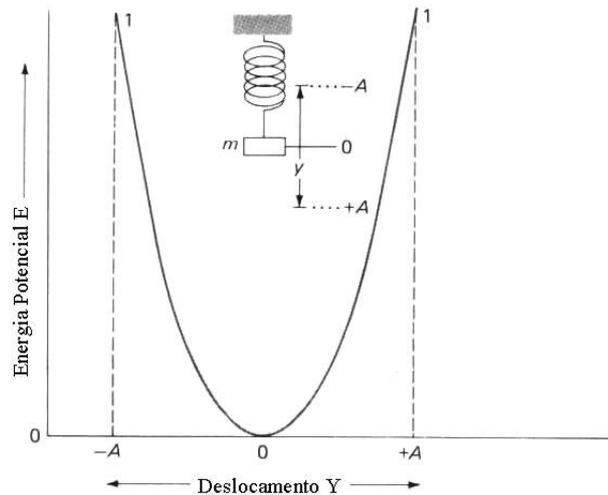


Figura 7 – Representação do perfil da energia potencial do modelo do oscilador harmônico.

Para o cálculo de frequências fundamentais de moléculas diatômicas, esse modelo apresenta bons resultados. Porém, para moléculas poliatômicas, nas quais há um grande número de centros de vibração, ele apresenta falhas, devido a alguns fenômenos moleculares não considerados. Quando dois átomos se aproximam em um movimento de vibração molecular, existem forças de repulsão de Coulomb que fazem com que a energia potencial aumente mais intensamente do que no modelo do oscilador harmônico. Se a distância entre os átomos aumentar de forma significativa, irá ocorrer algumas distorções na molécula, o que também acarreta uma distorção no modelo do oscilador harmônico, originando o modelo anarmônico, conforme pode ser observado pela Figura 8.

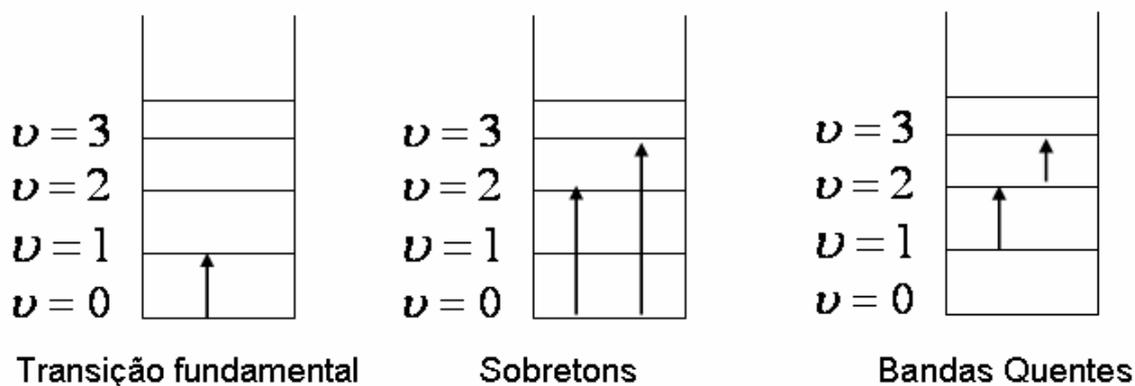


Figura 9 – Ilustração das transições entre os níveis de energia em um oscilador anarmônico.

2.2.2. Características da Espectroscopia no Infravermelho Próximo

Apesar da radiação na região do infravermelho próximo ter sido descoberta antes da radiação na região do infravermelho médio, a espectroscopia NIR acabou sendo ignorada por muito tempo, devido ao fato de existir uma grande dificuldade em extrair as informações relevantes. As bandas de absorção nessa região são produtos de sobretons e bandas de combinação das vibrações de ligações C-H, O-H e N-H, e costumam ser de intensidades fracas e muito sobrepostas. Isso limita as análises qualitativas e quantitativas univariadas nessa região. A maioria dessas aplicações usando NIR em análises quantitativas começou a partir da década de 1980, devido ao salto tecnológico provocado pelo desenvolvimento de ferramentas estatísticas e matemáticas de calibração multivariada, dentro da área de quimiometria [38,39].

A espectroscopia NIR é uma técnica analítica de grande aplicabilidade apresentando inúmeras vantagens em relação às técnicas tradicionais, o que a torna de crescente aplicação em diversas áreas da ciência. Dentre essas vantagens podem-se destacar [35,36]:

- Possibilidade de ensaio sem preparação de amostra;
- Rapidez para a obtenção dos espectros, cálculos e apresentação dos resultados;
- Técnica não destrutiva, permitindo o uso posterior da amostra;

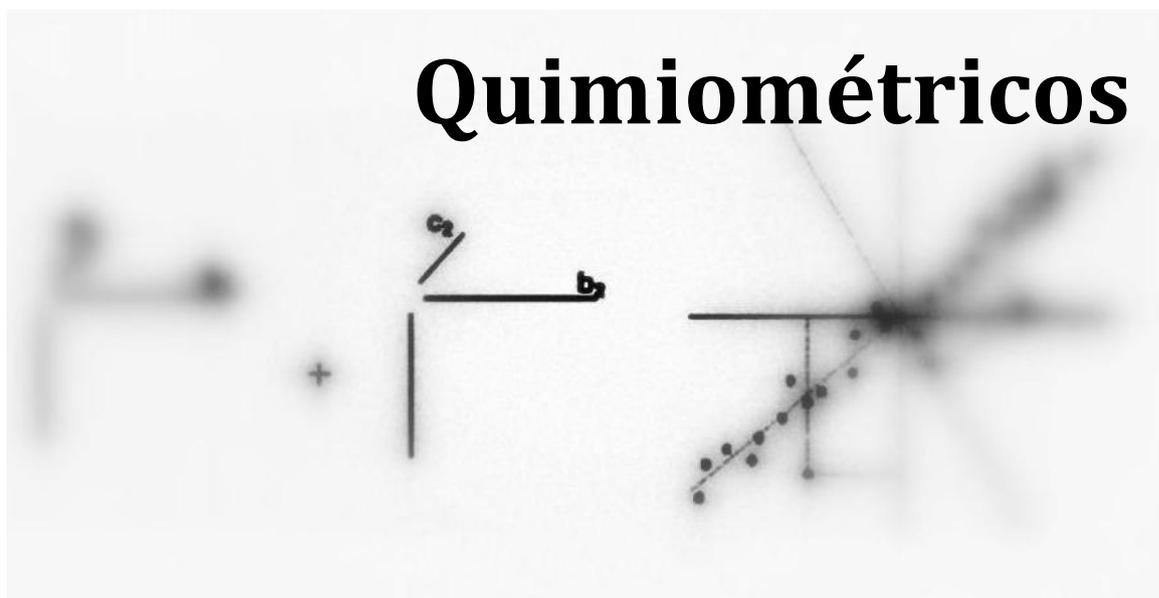
- Técnica não invasiva, com alta penetração do feixe de radiação (cerca de 1 a 3 mm);
- Possibilidade do desenvolvimento de aplicações em linha (analisadores de processo).

As informações a respeito de um determinado sistema usando espectroscopia NIR não são evidentes e não podem ser extraídas por meio de um único ou poucos comprimentos de onda, tornando a análise quantitativa do espectro mais complexa. Para solucionar esse problema, métodos estatísticos multivariados, tais como regressão linear múltipla, regressão por mínimos quadrados parciais e regressão por componentes principais, são aplicados para extrair as informações relevantes do conjunto de dados espectrais e para relacioná-las com a variável de interesse [36,40].

Capítulo 3

Métodos

Quimiométricos



3. MÉTODOS QUIMIOMÉTRICOS

Com o surgimento de métodos instrumentais computadorizados para análise química começou a ser gerada uma grande quantidade de dados e para interpretar essa grande quantidade de informação numérica, os químicos começaram a usar métodos estatísticos e matemáticos que já eram utilizados em outras áreas da ciência, como a geologia e a psicologia [41]. Surgiu então a quimiometria, que pode ser definida como a disciplina que aplica e desenvolve métodos matemáticos e estatísticos para o tratamento de dados químicos [42]. A área da química de maior aplicação de métodos quimiométricos é a química analítica, devido à grande quantidade de metodologias de análise capazes de gerar dados muitas vezes complexos e de natureza variada [43].

As linhas de pesquisa dentro da quimiometria envolvem áreas como reconhecimento de padrões, planejamento e otimização de experimentos, processamento de sinais analíticos, métodos de inteligência artificial e calibração multivariada, sendo esta de grande importância, consistindo na principal linha de pesquisa dentro da quimiometria [44]. Entretanto, antes que sejam construídos os modelos, são fundamentais algumas etapas como a organização dos dados e a etapa de pré-processamento.

3.1. ORGANIZAÇÃO DOS DADOS

Antes de começar a análise dos dados, sejam de primeira ou segunda ordem, é necessário que estes sejam devidamente organizados. Dados de primeira ordem são aqueles em que para cada amostra é obtido um vetor, sendo que para cada análise será obtida uma matriz de dados. Nessa matriz de dados, as linhas correspondem às amostras e as colunas referem-se às variáveis, como pode ser observado na Figura 10. Um exemplo de dados de primeira ordem são as matrizes espectrais.

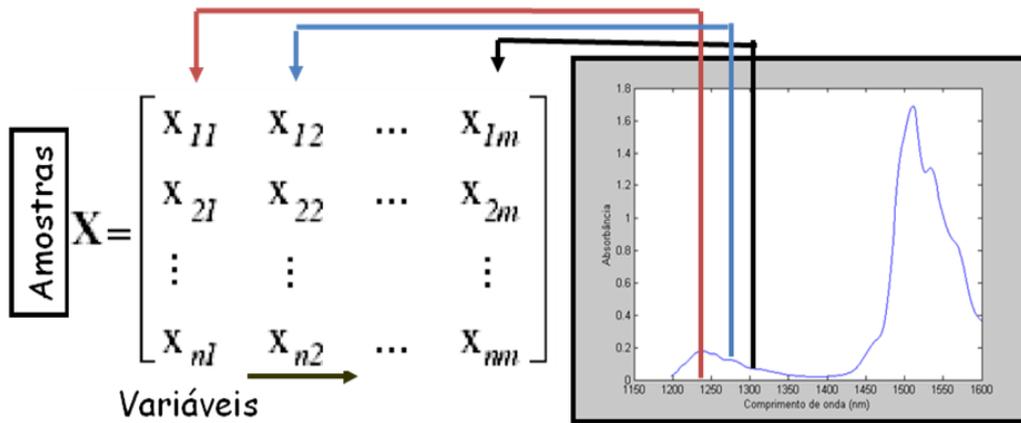


Figura 10 – Dados espectrais arranjados na forma de uma matriz X .

Dados de segunda ordem são aqueles em que para cada amostra é obtida uma matriz, como exemplo dados obtidos através de imagens. Na Figura 11 é mostrado um tensor correspondente aos dados obtidos por imagens.

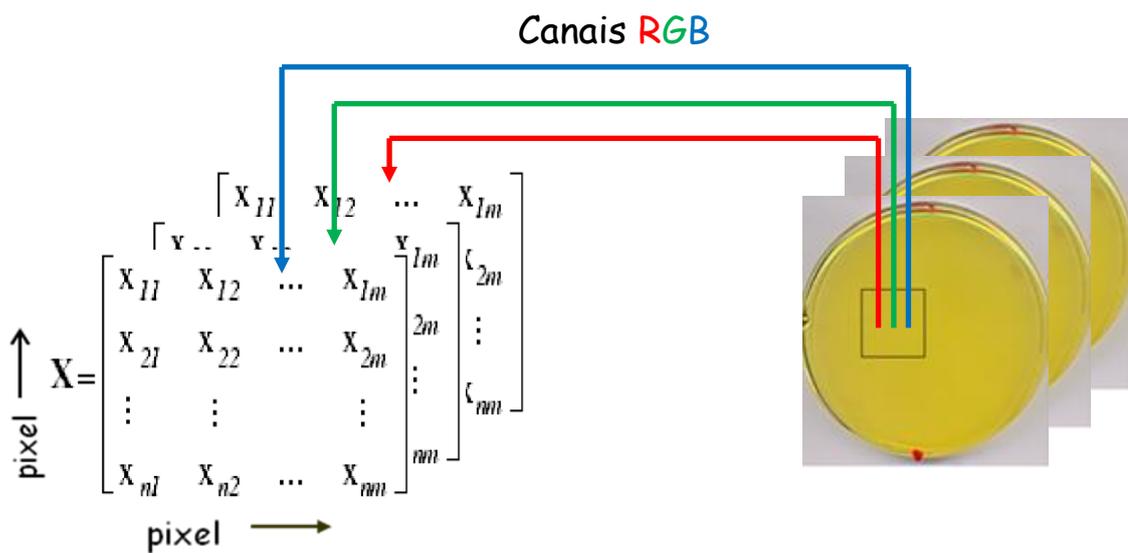


Figura 11 – Arranjo de dados de imagens.

O conjunto de dados de imagens é arranjado conforme a Figura 12, na qual cada matriz referente a uma amostra é colocada justaposta a outra, formando um arranjo cúbico, que será tratado usando métodos de calibração de segunda ordem.

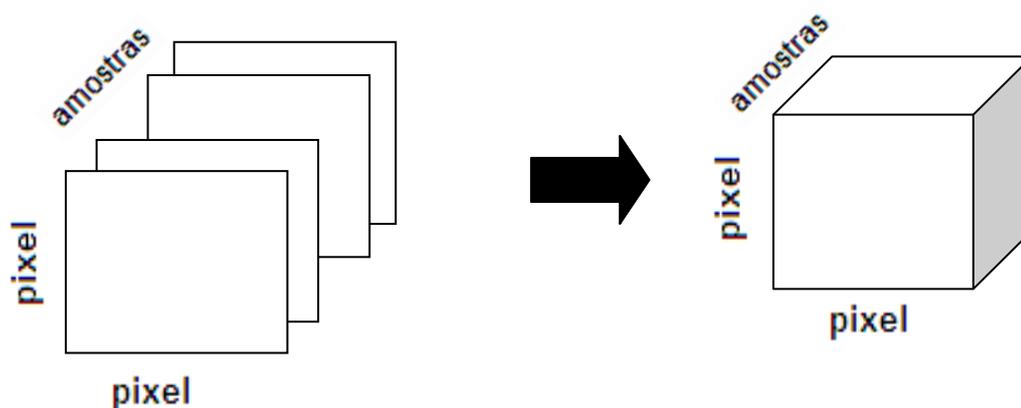


Figura 12 – Representação tridimensional de dados de imagens.

As notações de álgebra linear também se aplicam à quimiometria. Dessa forma, um vetor é uma matriz que tem uma linha ou uma coluna, sendo chamado, respectivamente, de vetor linha ou coluna. Os vetores são representados por letras minúsculas em negrito (**x**) e as matrizes por letras maiúsculas em negrito (**X**). Os elementos das matrizes são representados por letras comuns (x). Os tensores usados para representar dados de segunda ordem são designados por letras maiúsculas em negrito e sublinhadas (**\underline{X}**).

Após a organização dos dados, é efetuada a etapa de pré-processamento, sendo esta de grande importância, pois os dados podem estar expressos em grandezas diferentes ou apresentarem muito ruído comprometendo a análise. Assim, um tratamento prévio resolve esses problemas.

3.2. PRÉ-TRATAMENTO DOS DADOS ESPECTRAIS

A etapa de pré-processamento é muito importante, pois dependendo da natureza e complexidade dos dados originais pode ser determinante na análise. Dados espectrais têm sido muito utilizados em análises quantitativas, porém as técnicas espectroscópicas fornecem uma grande quantidade de informação, sendo que parte desta não se encontra correlacionada diretamente com o analito, podendo ocasionar distorções ao modelo e conseqüentemente conclusões errôneas. Para que esses problemas não ocorram são utilizados métodos de pré-processamento como normalização, correção de espalhamento multiplicativo, cálculo de derivada, escalamento, centrar os dados na média e alisamento *Savitsky-Golay*. No presente

trabalho serão discutidos alguns pré-tratamentos aplicados aos conjuntos de dados espectrais obtidos.

3.2.1. Centrar os dados na média

Centrar os dados na média consiste em fazer com que cada variável tenha média zero. Dessa forma, obtêm-se para cada coluna o valor médio e, em seguida, subtrai-se este valor de cada variável dessa mesma coluna, como pode ser observado na equação 4. Assim, é feita a mudança do sistema de coordenadas para o centro dos dados,

$$\mathbf{X}_{(i,j)cm} = \mathbf{X}_{(i,j)} - \bar{\mathbf{X}}_j \quad (4)$$

onde:

$\mathbf{X}_{(i,j)cm}$ = valor centrado na média para a variável j , na amostra i ;

$\mathbf{X}_{(i,j)}$ = valor da variável j na amostra i ;

$\bar{\mathbf{X}}_j$ = média dos valores da coluna j , calculados pela equação 5, sendo que n corresponde ao número de amostras.

$$\bar{\mathbf{X}}_j = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_{ij} \quad (5)$$

A representação da mudança do sistema de coordenadas para o centro dos dados é mostrada na Figura 13.

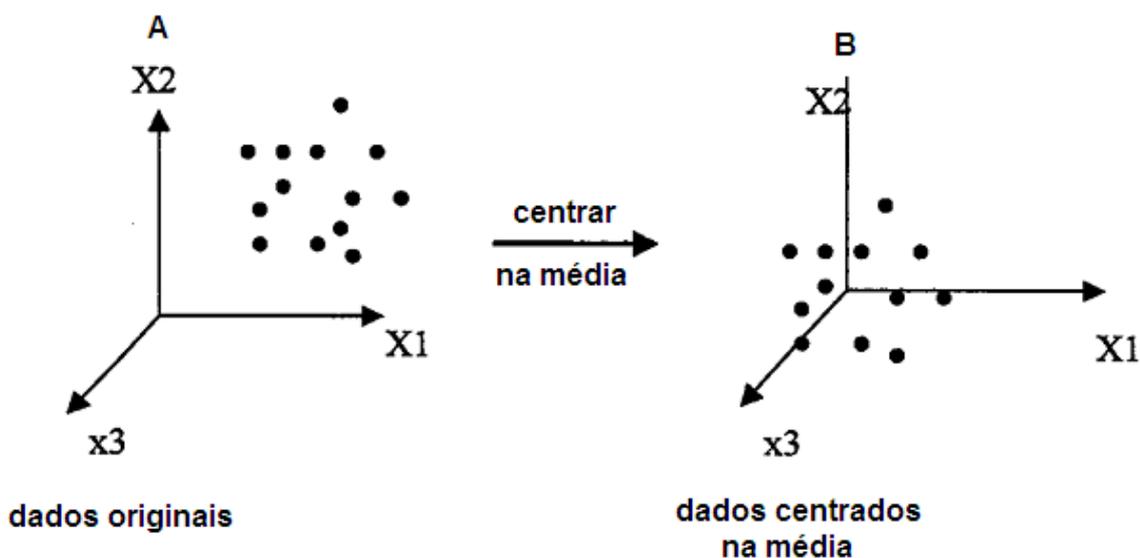


Figura 13 – Representação gráfica dos dados centrados na média. (A) dados antes da centralização; (B) dados centrados na média.

3.2.2. Aplicação de Derivadas

A aplicação da primeira ou da segunda derivada em dados espectrais é uma etapa de pré-processamento que tem por objetivo corrigir desvios lineares na linha base dos espectros. Após a aplicação de derivadas aos espectros, as informações contidas ao longo dos diferentes comprimentos tornam-se acentuadas, não só os sinais espectrais, mas também os ruídos. Portanto, esse tipo de pré-tratamento traz o inconveniente de diminuir a razão sinal/ruído.

3.2.3. Correção de espalhamento multiplicativo (MSC)

A correção de espalhamento multiplicativo (MSC, *Multiplicative Scatter Correction*) é uma técnica matemática que tem por objetivo corrigir desvios de linha base não lineares (*drifts*), causados pelo espalhamento multiplicativo da radiação. Estes efeitos são eliminados fazendo-se a regressão linear das variáveis espectrais nas variáveis do espectro médio. Essa correção faz uma regressão de um espectro medido contra um espectro de referência, geralmente o espectro médio. A vantagem deste pré-tratamento em relação aos demais, é que o espectro corrigido se assemelha ao espectro original, auxiliando na interpretação.

3.2.4. Alisamento Savitzky-Golay

O procedimento de alisamento é feito com o objetivo de minimizar os ruídos espectrais. Para isso, é utilizado o algoritmo de Savitzky-Golay. A característica desse procedimento é a redução dos ruídos por aproximadamente a raiz quadrada do número de pontos usados na convolução. O número de pontos ideal para a convolução é aquele que acarretará a menor distorção em relação aos dados originais.

3.3. ANÁLISE DE COMPONENTES PRINCIPAIS

A base fundamental para a maioria dos métodos estatísticos multivariados é a análise de componentes principais (PCA, *Principal Components Analysis*). Esse método começou a ser utilizado no tratamento de dados químicos no final da década de 60 por Malinowski, ainda com o nome de análise de fatores. O objetivo da PCA é reduzir o espaço multidimensional para melhor visualização e compreensão do conjunto de dados. Isso é feito através da projeção das amostras em um novo sistema de eixos chamados de componentes principais, como pode ser observado na Figura 14, na qual é mostrada a projeção das duas primeiras componentes principais em um sistema multidimensional.

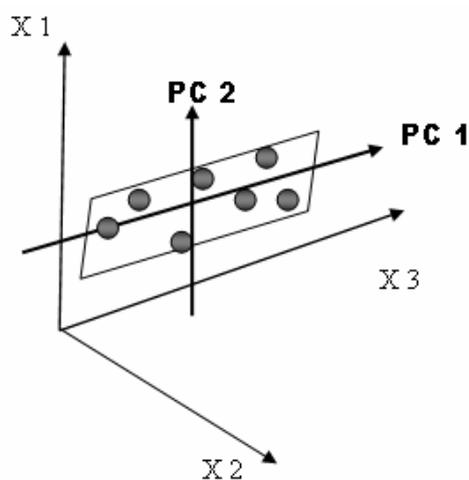


Figura 14 – Representação das duas primeiras componentes principais em um sistema multidimensional.

Nesse novo sistema de eixos cada uma das componentes principais é gerada a partir da combinação linear das m -variáveis originais, as quais são ortogonais entre si, ou seja, são completamente não correlacionadas. A decomposição dos dados originais \mathbf{X} é feita conforme a equação 7 e pode ser representada pelo esquema da Figura 15, sendo \mathbf{T} chamada de matriz de escores e \mathbf{L} conhecida como matriz de pesos (*loadings*).

$$\mathbf{X} = \mathbf{T} * \mathbf{L}^T \quad (7)$$

Matriz de Dados Originais Matriz dos Escores Matriz Transposta dos Pesos

Figura 15 – Representação da decomposição da matriz de dados \mathbf{X} em escores e pesos.

Os escores representam a projeção das amostras nesse novo sistema de coordenadas e os pesos são os coeficientes que medem a importância de cada variável original em cada componente principal. Os pesos são os cossenos dos ângulos entre o eixo da componente principal e os eixos das variáveis originais, e seus valores estarão entre 1 e -1 . Quanto mais próximo de ± 1 , maior será a influência de uma determinada variável na descrição desta componente principal; quanto mais próximo de zero, menor será sua influência. Dessa forma, através da análise dos pesos é possível identificar as variáveis que mais contribuem para a descrição de cada conjunto de dados. A representação da decomposição dos dados em componente principais é mostrada na Figura 16, sendo que a projeção das CP's ocorre na direção de maior variância dos dados (A) e em (B) e (C) é mostrado a projeção dos pesos (contendo as informações a respeito das variáveis) e dos escores (contendo as informações a respeito das amostras), respectivamente.

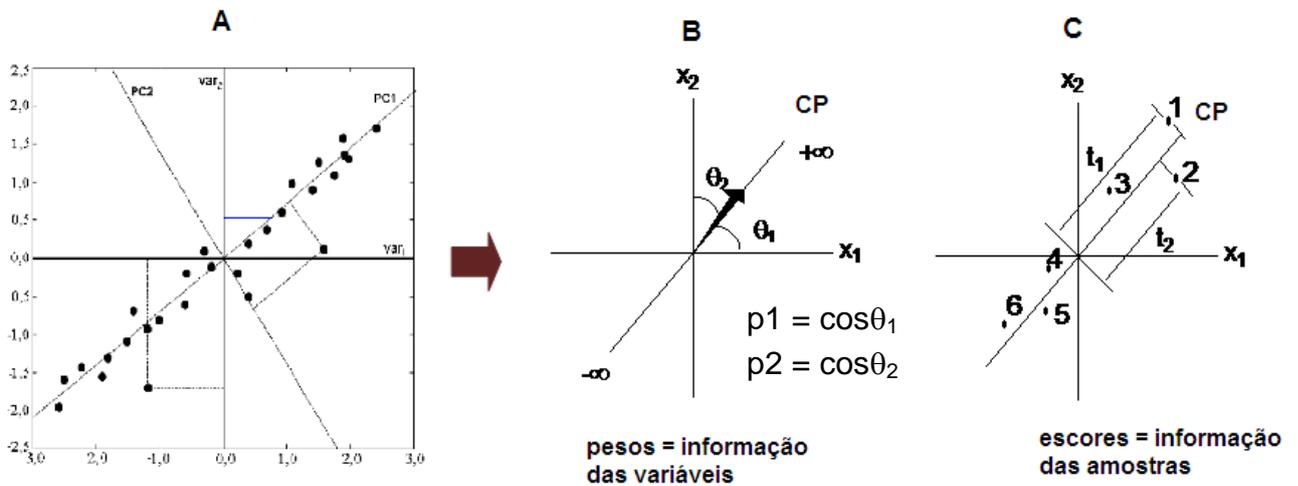


Figura 16 – Representação dos eixos das componentes principais (A); projeção dos pesos (B); projeção dos escores (C).

De forma geral, as combinações lineares das m -variáveis originais que geram cada componente principal podem ser representadas pela equação 8.

$$PC_i = x_{i1}p_1 + x_{i2}p_2 + \dots + x_{im}p_m \quad (8)$$

Na equação acima, x_{ij} (para $j = 1, 2, \dots, m$) são as variáveis originais e p_j (para $j = 1, 2, \dots, m$) são os coeficientes que medem a importância de cada variável na i -ésima componente principal (CP), ou seja, é o peso que cada variável tem na combinação linear. A escolha do número de componentes principais a serem utilizadas na construção do modelo é feita considerando o percentual de variância explicada pelas CP's e a variância residual. A Tabela 3 apresenta a variância explicada por cada componente principal.

Tabela 3 – Variância explicada pelos Componentes Principais.

Componente Principal	Variância explicada por cada componente	Variância total
1	β_1	β_1
2	β_2	$\beta_1 + \beta_2$
3	β_3	$\beta_1 + \beta_2 + \beta_3$
...
i	β_i	$\beta_1 + \beta_2 + \beta_3 + \dots + \beta_i = (100\%)$

3.4. CALIBRAÇÃO MULTIVARIADA

Calibração é um conjunto de operações que estabelecem uma relação matemática entre medidas instrumentais e propriedades físico-químicas de interesse [45]. Ela envolve duas fases, a primeira chamada de calibração, na qual ocorre a construção do modelo, e a segunda chamada de previsão, na qual é verificada a validade do modelo [46].

Os métodos de calibração podem ser divididos quanto à dimensionalidade e complexidade dos dados em calibração de ordem zero, primeira e superior (segunda ordem ou mais) [47]. Os modelos de calibração de ordem zero consistem de regressões univariadas, que são realizadas quando se utiliza apenas um número ou escalar como resposta instrumental para cada amostra. Para os modelos de calibração de primeira ordem é obtido um vetor como resposta instrumental para cada amostra, resultando em uma regressão multivariada [48]. Esses modelos constituem uma importante linha de pesquisa dentro da quimiometria. Dentre os modelos de regressão multivariada de primeira ordem, os mais utilizados em tratamentos de dados químicos são: regressão linear múltipla (MLR, *Multiple Linear Regression*), regressão por componentes principais (PCR, *Principal Components Regression*) e regressão por mínimos quadrados parciais (PLS, *Partial Least Squares*) [49].

Modelos de calibração de segunda ordem são construídos para métodos que geram uma matriz de dados por amostra. Esses dados podem ser gerados por técnicas como cromatografia líquida com detecção por arranjo de diodos (HPLC-DAD), cromatografia líquida de alta eficiência ou gasosa com detecção por espectrômetro de massas (HPLC-MS e CG-MS), ressonância nuclear magnética em duas dimensões (RMN-2D), espectralfluorimetria, análise por injeção em fluxo com gradiente de pH e detecção por arranjo de diodos, dados de microscopias e imagens em geral [50]. O principal modelo de calibração de segunda ordem é a análise de fatores paralelos (PARAFAC, *Parallel Factor Analysis*).

3.4.1. Regressão por mínimos quadrados parciais – PLS

O PLS [41,42,51-56] é um método, cujo o modelo matemático que tem por objetivo extrair informações relevantes de um conjunto de dados espectrais para correlacioná-las com uma ou mais variáveis de interesse, sendo considerado o método de regressão mais utilizado na construção de modelos de calibração para dados de primeira ordem.

O modelo PLS reduz o espaço multivariado através da criação de um novo conjunto de variáveis chamadas variáveis latentes, que podem ser decompostas em escores e pesos (*loadings*) e são combinações lineares das variáveis espectrais originais. É usado para a construção do modelo de calibração um número de variáveis latentes que proporcione o menor erro possível de previsão, ou seja, que minimize as diferenças entre os valores previstos e os valores experimentais.

As variáveis de interesse (por exemplo, a tensão interfacial de óleos isolantes) são chamadas de variáveis dependentes e são representadas pela letra Y. Para que o modelo matemático seja construído é necessário que se tenha alguma informação a respeito dessas variáveis dependentes. Essas informações são obtidas pela correlação com outras variáveis chamadas de variáveis independentes e representadas pela letra X. Exemplos dessas variáveis independentes são os sinais analíticos que podem ser relacionados com as propriedades físico-químicas [57].

Na construção do modelo PLS ocorre a decomposição simultânea das matrizes de dados \mathbf{X} (variáveis independentes) e \mathbf{Y} (variáveis dependentes). O processo de decomposição dessas matrizes pode ser representado pela equação 9 e 10.

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E}_x \quad (9)$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{E}_y \quad (10)$$

Onde \mathbf{T} e \mathbf{U} contêm os escores para as duas matrizes de dados e \mathbf{P} e \mathbf{Q} contêm os respectivos pesos. \mathbf{E}_x e \mathbf{E}_y são as matrizes de resíduos contendo a parte não modelada dos dados. As equações (9) e (10) podem ser representadas pelo esquema da Figura 17.

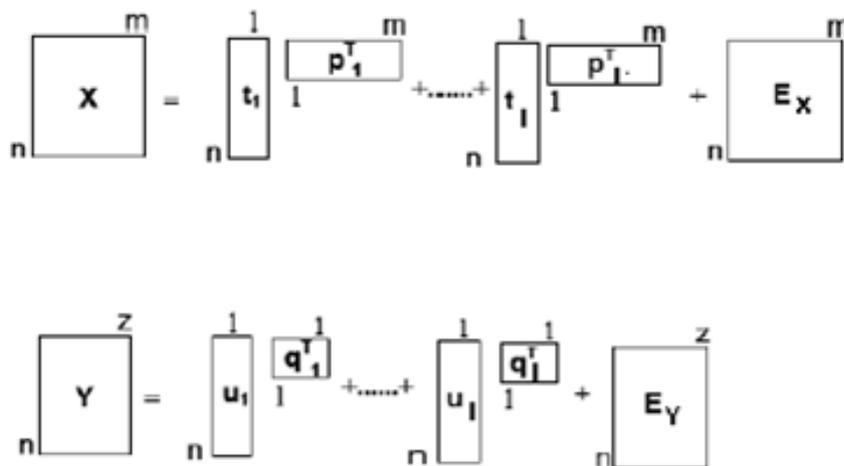


Figura 17 – Decomposição das matrizes \mathbf{X} e \mathbf{Y} em variáveis latentes para o modelo PLS.

Após a decomposição das matrizes \mathbf{X} e \mathbf{Y} , é estabelecida uma relação linear entre os escores de \mathbf{X} e os escores de \mathbf{Y} para cada variável latente, conforme a equação 11,

$$\mathbf{u}_h = \mathbf{b}_h \mathbf{t}_h \quad (11)$$

$$\mathbf{Y} = \mathbf{TBQ}^T + \mathbf{E} \quad (12)$$

onde \mathbf{b} é o vetor contendo os coeficientes de regressão linear da regressão do modelo para cada variável latente. O índice h da equação indica o número de variáveis latentes do modelo. Desta maneira, os escores de \mathbf{Y} (matriz \mathbf{U}) podem ser previstos a partir dos escores de \mathbf{X} (matriz \mathbf{T}) e, assim, a variável de interesse pode ser prevista por meio dos dados espectrais contidos em \mathbf{X} .

3.4.2. Escolha do Número de Variáveis Latentes

Na construção do modelo de regressão, um ponto muito importante é a escolha do número de variáveis latentes (LV, *Latent Variables*). Deve-se selecionar sempre um número de variáveis latentes com informações suficientes, evitando tanto sobreajuste quanto subajuste. Um modelo que contenha um número muito grande de variáveis latentes é sobreajustado, podendo incluir ruídos, enquanto um número pequeno de LV's pode deixar informação importante sem ser modelada. Dessa forma, um número ideal de variáveis latentes é aquele que permite a criação de um modelo com uma boa capacidade de previsão.

A validação cruzada (CV, *Cross Validation*) é o métodos mais usado para a escolha do número de variáveis latentes, podendo ser realizada em blocos, ou seja, um número determinado de amostras é deixado de fora no processo de construção do modelo e a seguir essas amostras são previstas pelo modelo construído, ou ainda, podendo ser feita individualmente (*leave one out*), ou seja, uma amostra de cada vez é deixada de fora no processo de construção do modelo e a seguir essa amostra é prevista pelo modelo construído. Em ambos os casos, o processo é repetido até que todas as amostras tenham sido previstas e determina-se o erro médio. A soma dos quadrados dos erros de previsão (PRESS, *Predicted Residual Error Sum of Squares*) é medida da mesma forma para cada número de variáveis latentes e pode ser definida conforme a equação 13,

$$\text{PRESS} = \sum (y_{\text{exp}} - y_{\text{prev}})^2 \quad (13)$$

onde y_{exp} e y_{prev} são os valores experimentais e previstos, respectivamente para a propriedade de interesse. Os resultados de PRESS são graficados em um gráfico em função do número de LV's e o perfil desse gráfico é analisado. De um modo geral, nesse gráfico é observado um mínimo ou um patamar, que é um indicativo da melhor dimensionalidade do modelo de regressão, ou seja, o menor número de LV's que produziu o menor erro de previsão sem perda significativa da variância dos dados.

3.4.3. Previsões do Modelo

Após a construção de modelo de calibração, a etapa seguinte consiste em verificar a sua validade. Para validar o modelo é necessário utilizar amostras que não foram utilizadas na etapa de calibração (validação externa). Nessa etapa, chamada de previsão, avalia-se se o modelo é capaz de prever corretamente ou com pequena margem de erro os valores das propriedades de interesse das novas amostras. Nesse trabalho foi utilizado o erro médio quadrático de previsão (RMSEP, *Root Mean Square Error Prediction*) para avaliar a capacidade de previsão do modelo. Tal valor expressa a exatidão do modelo, ou seja, a proximidade entre os valores calculados pelo modelo (y_{prev}) e os valores “verdadeiros” ou obtidos por um método de referência (y_{exp}). O RMSEP é definido através da equação 14, onde n é o número de amostras.

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_{prev} - y_{exp})^2}{n}} \quad (14)$$

3.5. MÉTODOS DE SEGUNDA ORDEM

3.5.1. PARAFAC

O PARAFAC [47,58-62] é um modelo matemático utilizado para tratar dados multidimensionais, ou seja, dados que são representados por um arranjo cúbico ou de ordem superior. Esse método foi proposto no início dos anos 70 por pesquisadores da área da psicologia (psicometria) e é uma generalização da análise de componentes principais (PCA) para dados de ordem superior.

No modelo PARAFAC o conjunto de dados multidimensionais representados por uma matriz cúbica \underline{X} de dimensões (I, J, K) é decomposto em três matrizes de pesos (*loadings*) \mathbf{A} (I, F), \mathbf{B} (J, F) e \mathbf{C} (K, F), em que F é o número de fatores do modelo, e as matrizes \mathbf{A} , \mathbf{B} e \mathbf{C} são as matrizes contendo os perfis de cada fator em cada dimensão.

A decomposição do modelo PARAFAC é expressa matematicamente pela equação 15, em que o arranjo $\underline{X}_{i,j,k}$ é a matriz cúbica, a, b e c são os elementos das matrizes \mathbf{A} , \mathbf{B} e \mathbf{C} , e $\underline{E}_{i,j,k}$ é a matriz de resíduos. Outra forma de representar o modelo PARAFAC é matricial, como mostrado na equação 16, na qual o símbolo \otimes representa o produto Khatri-Rao [63].

$$\underline{X}_{i,j,k} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + \underline{E}_{i,j,k} \quad (15)$$

$$\underline{X} = \mathbf{A}(\mathbf{C} \otimes \mathbf{B})^T + \underline{E} \quad (16)$$

Na Figura 18 é ilustrada a decomposição dos dados da matriz \underline{X} na forma de tríades, em que cada tríade representa um fator.

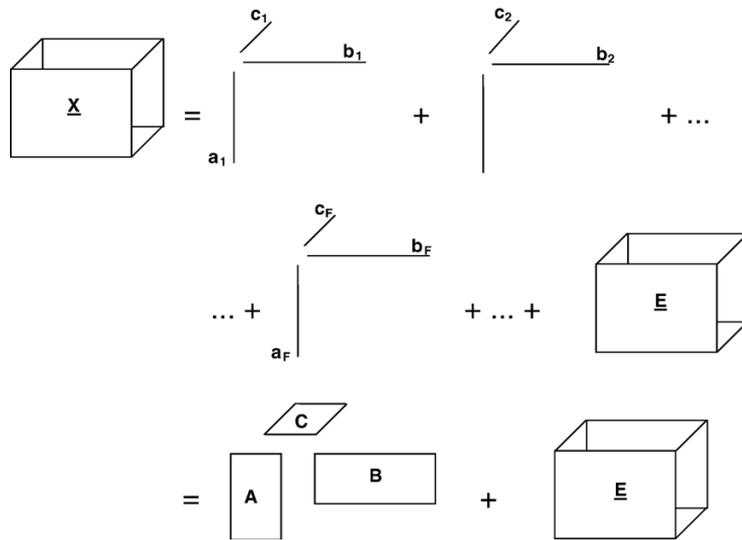


Figura 18 – Esquema da decomposição trilinear do modelo PARAFAC em tríades.

As aplicações do modelo PARAFAC envolvem alguns aspectos fundamentais, dentre eles: a consistência multilinear dos dados, o método de inicialização do algoritmo, o critério de convergência, as restrições impostas ao modelo e o número de fatores ideal. O algoritmo usado para resolver o modelo PARAFAC é o dos Mínimos Quadrados Alternados (ALS, *Alternating Least Squares*). O ALS sucessivamente assume os pesos em dois modos e então estima o conjunto desconhecido de parâmetros do último modo [62]. O ALS pode ser iniciado a partir de valores aleatórios ou usando algum tipo de decomposição dos dados, decomposição trilinear direta (DTD, *Direct Trilinear Decomposition*) ou decomposição em valores singulares (SVD, *Singular Value Decomposition*). Normalmente, o padrão (*default*) usado no PARAFAC é a DTD [62]. O algoritmo converge iterativamente até atingir um critério de convergência ou um número de iterações previamente definidos. O padrão usado costuma ser uma diferença de ajuste entre sucessivas iterações de 1.10^{-6} ou um número máximo de 2.500 iterações. As restrições impostas ao modelo, tais como não negatividade,

ortogonalidade e unimolalidade são alternativas utilizadas para melhorar as soluções do algoritmo e torná-las mais realistas.

A escolha do número de fatores é de grande importância na construção do modelo, porém dependendo do conjunto analisado essa escolha não é muito simples, não existindo um critério único. Em sistemas não muito complexos a escolha pode ser feita com base no conhecimento prévio do número de espécies. Para um conjunto de dados em que não haja esse conhecimento preliminar, essa escolha pode ser feita utilizando critérios como a variância explicada pelo modelo ou através do diagnóstico de consistência trilinear (CORCONDIA, *CORe CONSistency DIAgnostic*) [64].

O modelo de regressão é obtido a partir da decomposição dos dados pelo PARAFAC através de uma regressão por mínimos quadrados entre os pesos relativos à dimensão das amostras (**A**) e o vetor com as propriedades de referência (**y**) das amostras de calibração, de maneira análoga ao PCR.

3.5.2. NPLS

O NPLS [33,60,61,65] (*N-way Partial Least Squares*) é um método de calibração multivariada que pode ser usado para tratar um arranjo de dados multidimensionais, sendo uma extensão do método PLS para dados de ordem superior. O método NPLS relaciona um arranjo trilinear de variáveis independentes **X** (I, J, K) com uma matriz ou vetor de variáveis dependentes **Y**. Nesse modelo o arranjo trilinear **X** é decomposto de maneira trilinear, de forma a encontrar a máxima covariância entre os escores de **X** (matriz cúbica de dados independentes) e **Y** (matriz contendo a propriedade físico-química). A principal diferença em relação ao PARAFAC é a modelagem simultânea dos dados das variáveis dependentes e independentes.

Para a construção do modelo NPLS são necessárias duas etapas, as mesmas do modelo PLS, a etapa de calibração e a etapa de previsão. Na calibração o arranjo **X** é decomposto em escores (t_n) e pesos (*weights*, w_j e w_k), conforme representado na Figura 19.

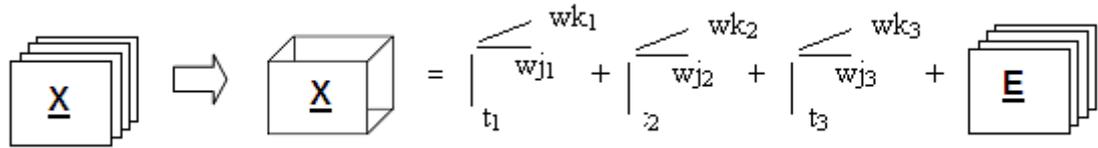


Figura 19 – Representação esquemática da decomposição de um arranjo multidimensional para o modelo NPLS.

Onde \underline{X} é o arranjo tridimensional, $\mathbf{T}(I, F)$ é a matriz relacionada ao modo das amostras (escores) e as matrizes $\mathbf{W}^J (J, F)$ e $\mathbf{W}^K (K, F)$ são relacionadas às outras duas dimensões, ou seja, são matrizes de pesos contendo as informações a respeito das variáveis. O arranjo \underline{E} contém os resíduos, que são a parte não explicada pelo modelo. Essa decomposição pode ser representada pela equação 17.

$$\underline{X}_{i,j,k} = \sum t_{ij} w_{j,f}^i w_{k,f}^k + \underline{E}_{i,j,k} \quad (17)$$

A matriz \mathbf{Y} também é decomposta em escores e pesos (*weights*), podendo esta ser representada pelo esquema da Figura 20 e pela equação 19,

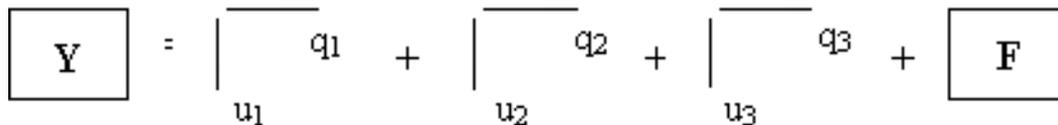


Figura 20 – Representação esquemática da decomposição da matriz \mathbf{Y} para o modelo NPLS.

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}' + \mathbf{F} \quad (18)$$

onde \mathbf{Y} é a matriz contendo a propriedade físico-química, \mathbf{U} é a matriz contendo os escores de \mathbf{Y} , \mathbf{Q} é a matriz contendo os pesos (*loadings*) de \mathbf{Y} e \mathbf{F} é a matriz de resíduos de \mathbf{Y} , que assim como no arranjo \underline{X} , não podem ser explicados pelo modelo.

Para relacionar a matriz de escores de \underline{X} e a matriz de escores de \mathbf{Y} é necessário encontrar uma relação linear entre a matriz de dados e a matriz da propriedade físico-química. Essa relação linear é representada pela equação 19,

$$\mathbf{U} = \mathbf{T}\mathbf{b} \quad (19)$$

onde \mathbf{b} é uma matriz contendo os coeficientes de regressão linear entre o arranjo de dados \mathbf{X} e \mathbf{Y} . A estimativa $\hat{\mathbf{Y}}$ para novas amostras pode ser feita a partir de novos escores \mathbf{T}^* , conforme a equação 20.

$$\hat{\mathbf{Y}} = \mathbf{T}^*\mathbf{b}^T \quad (20)$$

O NPLS apresenta vantagens em relação aos métodos de desdobramento, que são métodos de primeira ordem usados para tratar dados de segunda ordem (*unfold*-PLS, por exemplo). Dentre essas vantagens encontram-se a simplicidade do modelo e a facilidade na interpretação dos resultados, além da menor sensibilidade a ruído.

Capítulo 4

Parte Experimental



4. PARTE EXPERIMENTAL

4.1. AMOSTRAGEM

Foram obtidas 48 amostras de óleos isolantes usados em transformadores junto a Centrais Elétricas de Goiás S.A. (CELG). Essas amostras foram coletadas pela CELG no ano de 2008 entre os meses de março e setembro, pois esse é o período de baixa umidade. As coletas foram feitas em todas as regiões do estado de Goiás obedecendo a uma logística da própria empresa energética, com critérios baseados no histórico de análises físico-químicas e cromatográficas feitas nos óleos dos transformadores de todas as subestações do estado de Goiás. As amostras de óleos coletadas pela CELG apresentam uma diversidade muito grande de tempo de uso nos transformadores, podendo variar de 1 a 30 anos. Isso faz com que essas amostras apresentem uma variedade muito grande de compostos dissolvidos, dificultando muito a interpretação dos resultados.

Após a coleta das amostras, foram feitos ensaios físico-químicos no laboratório de análises de óleos da CELG. Esses ensaios são feitos para avaliar as propriedades dielétricas e refrigerantes do óleo isolante, ou até mesmo avaliar indiretamente o estado de degradação do sistema isolante (papel-óleo). Dentre esses ensaios, foram feitas análises de cor, medida de acidez por titulação, análise de umidade, fator de potência e tensão interfacial. O ensaio de tensão interfacial é feito usando um tensiômetro de torção Krüss modelo K8, mostrado na Figura 21, sendo que através desse ensaio é possível avaliar indiretamente o estado de degradação do sistema isolante, pois ele é uma medida indireta da quantidade de compostos polares existentes no óleo. Quanto maior a concentração desses compostos (água, por exemplo), menor será o valor da tensão interfacial e maior será o estado de degradação do sistema, pois tais compostos polares são oriundos da degradação do papel e do óleo.



Figura 21 – Tensiômetro de torção Krüss.

Dessa forma, a medida de tensão interfacial é de grande importância na avaliação das condições de operação dos transformadores elétricos, sendo escolhida para ser correlacionada com as medidas espectrais e com as imagens obtidas dos óleos. Embora a medida de tensão interfacial seja importante, o método padrão (NBR 6234) [22] apresenta as desvantagens de possuir baixa precisão e ser relativamente demorado.

4.2. OBTENÇÃO DAS MEDIDAS

4.2.1. Espectroscopia no Infravermelho Próximo de óleos isolantes

Foram obtidos espectros das 48 amostras de óleos minerais isolantes usando um espectrofotômetro Perkim Elmer Spectrum 100N FTIR e uma cubeta de quartzo de 1 mm de caminho óptico, na região de 7800 a 4000 cm^{-1} (1282 a 2500 nm), com uma resolução de 4 cm^{-1} e 32 varreduras. O espectrofotômetro utilizado no processo de obtenção dos espectros é mostrado na Figura 22.

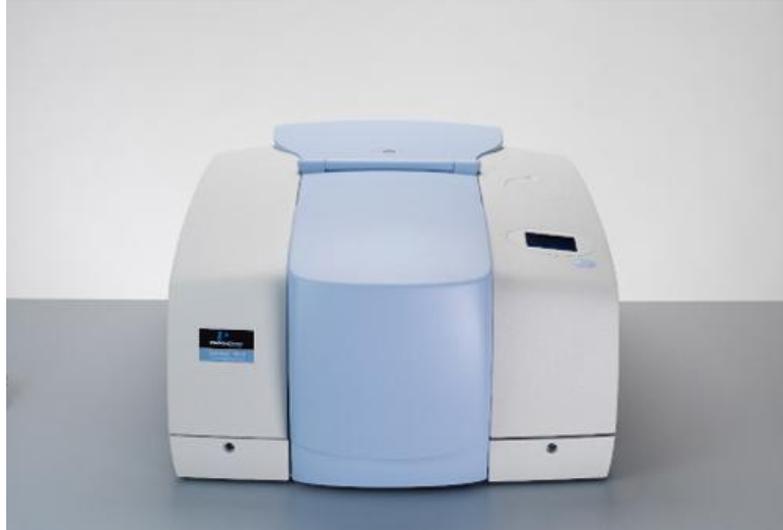


Figura 22 – Espectrofotômetro utilizado para a obtenção dos espectros das amostras de óleo.

4.2.2. Análise de imagens de óleos isolantes

Foram capturadas imagens das 48 amostras de óleos minerais (mesmo conjunto de amostras utilizadas na análise espectroscópica) usando um *scanner* de mesa modelo Genius. Na Figura 23 estão representadas imagens das 48 amostras de óleos usadas na construção do modelo de regressão.

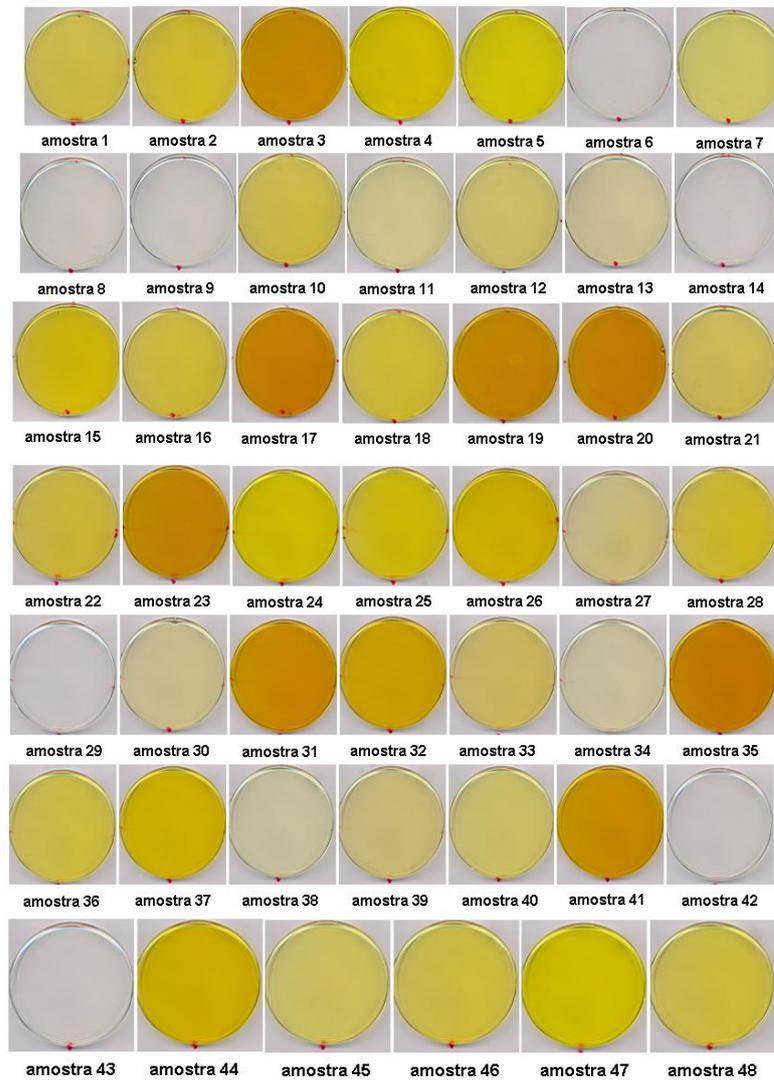


Figura 23 – Imagens das 48 amostras de óleos digitalizadas usadas na construção do modelo.

A obtenção das imagens foi realizada utilizando uma placa de petri, na qual foram adicionadas 15 mL de óleo. Esse procedimento foi realizado em triplicata para cada amostra e para cada amostragem (15 mL) foram digitalizadas três imagens; portanto, para cada amostra de óleo foram obtidas nove imagens. A placa foi posicionada sobre a região central do *scanner* e sobre a placa foi colocado um anteparo branco com o objetivo de impedir a entrada de luz de outras fontes de iluminação que não fossem provenientes do *scanner*. As imagens foram digitalizadas no sistema de cor RGB no formato bitmap (BMP) com resolução de 300

dpi e posteriormente convertidas para o sistema *gray-scale*. O *scanner*, a placa e o anteparo utilizados durante o processo de digitalização das imagens são mostrados na Figura 24.



Figura 24 – *Scanner* e a placa de petri utilizados na obtenção das imagens de óleos.

4.3. TRATAMENTO DOS DADOS

Todos os conjuntos de dados foram tratados no programa *MATLAB*TM, versão 6.5 (*The MathWorks*, Natick, EUA), usando o pacote *PLS Toolbox*, versão 2.1 (*Eigenvenctor Research*, Manson, EUA), para os dados de NIR e o pacote *The N-way Toolbox for MATLAB*, versão 2.10, para os dados de imagens [62].

Para cada amostra de óleo foram obtidas nove imagens, que foram importadas para a área de trabalho do Matlab. Essas imagens são coloridas, tridimensionais e de natureza multivariadas caracterizadas por arranjos do tipo objeto x objeto x variável, ou seja, **OOV** (pixel (**O**) x pixel (**O**) x canal de cor (**V**)) e sua manipulação é muito complicada, pois ao serem empilhadas formam arranjos de 4 dimensões. Em contrapartida, as imagens em escala de cinzas são bidimensionais

e univariadas caracterizadas por arranjos do tipo objeto x objeto, sendo chamadas de arranjos **OO** (pixel x pixel), o que torna sua manipulação mais vantajosa em relação às imagens coloridas, devido ao fato delas poderem ser empilhadas formando arranjos do tipo **OOO** (amostras x pixels x pixels). Embora seja vantajoso o uso de imagens em escala de cinzas, estas não atingem o critério de trilinearidade necessário à aplicação direta dos métodos de ordem superior, devido os arranjos formados serem do tipo (OOO). Neste caso, o uso de filtros é fundamental, pois através deles o arranjo **OOO** é convertido em **OVV** (amostra x canal de cor x canal de cor), podendo ser modelado por métodos multidimensionais como PARAFAC e NPLS. Assim, cada imagem colorida foi convertida para escala de cinzas e após essa conversão, foi efetuada a etapa de pré-processamento das imagens. Essa etapa consistiu em aplicar em cada imagem o filtro especial *unsharp* com o objetivo de realçar os contornos da imagem; em seguida, cada imagem foi submetida à transformada de Fourier bidimensional (2D-FFT), seguida de uma transformação logarítmica. As funções usadas nessa etapa encontram-se no *toolbox* de imagens do Matlab. Após essa etapa de pré-processamento, foi calculada uma média das imagens de cada amostra, obtendo-se dessa forma 48 imagens médias correspondentes as 48 amostras.

A Figura 25 mostra um esquema do procedimento realizado com cada imagem, desde a sua digitalização até a etapa de modelagem. A rotina para a realização desse procedimento foi desenvolvida pela própria autora no decorrer do trabalho e está descrita no apêndice.

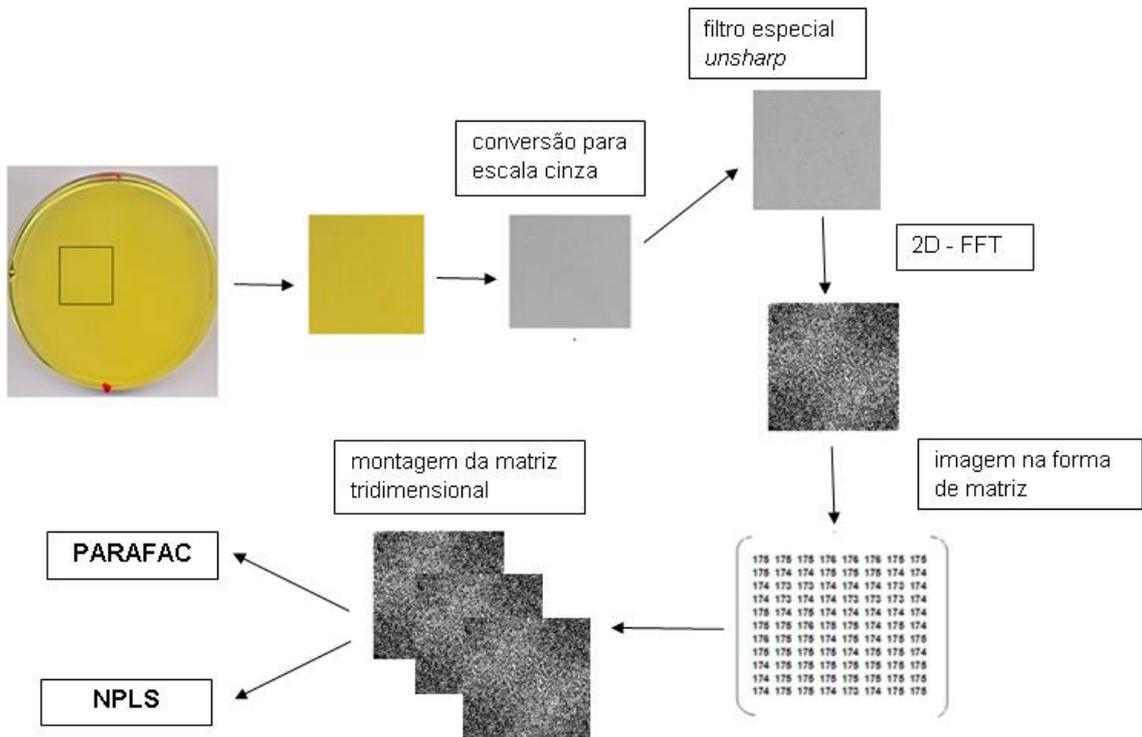


Figura 25 – Esquema do tratamento das imagens e montagem dos arranjos tridimensionais.

Capítulo 5

Resultados e Discussão



5. RESULTADOS E DISCUSSÃO

5.1. ESPECTROSCOPIA NIR

Foi construído um modelo de calibração para correlacionar os espectros das 48 amostras de óleos com a propriedade físico-química tensão interfacial. Esses valores são apresentados na Tabela 4, na qual pode se observar uma variação entre 15 a 44 dina/cm, sendo que valores de tensão interfacial baixos indicam a presença de compostos polares em consequência de uma alta degradação do sistema isolante. O conjunto de calibração foi construído com 38 amostras e o conjunto de validação com 10 amostras.

Tabela 4 – Tensão interfacial quantificada nas amostras de óleos e usada para construir os modelos de calibração para dados espectrais e de imagens.

amostras	tensão interfacial (dina/cm)	amostras	tensão interfacial (dina/cm)
1	33	25	29
2	26	26	25
3	16	27	29
4	32	28	32
5	37	29	45
6	43	30	29
7	39	31	15
8	39	32	18
9	40	33	20
10	27	34	31
11	20	35	18
12	19	36	34
13	30	37	21
14	44	38	44
15	24	39	33
16	29	40	33
17	16	41	19
18	34	42	43
19	20	43	42
20	18	44	25
21	33	45	36
22	20	46	36
23	23	47	40
24	31	48	44

As amostras de validação foram selecionadas a partir da observação dos escores de uma análise de componentes principais (PCA), de modo a serem homoganeamente representativas de toda a faixa de variação espectral. No gráfico da Figura 26 são mostrados os escores para a primeira e a segunda componente principal e as amostras de validação selecionadas.

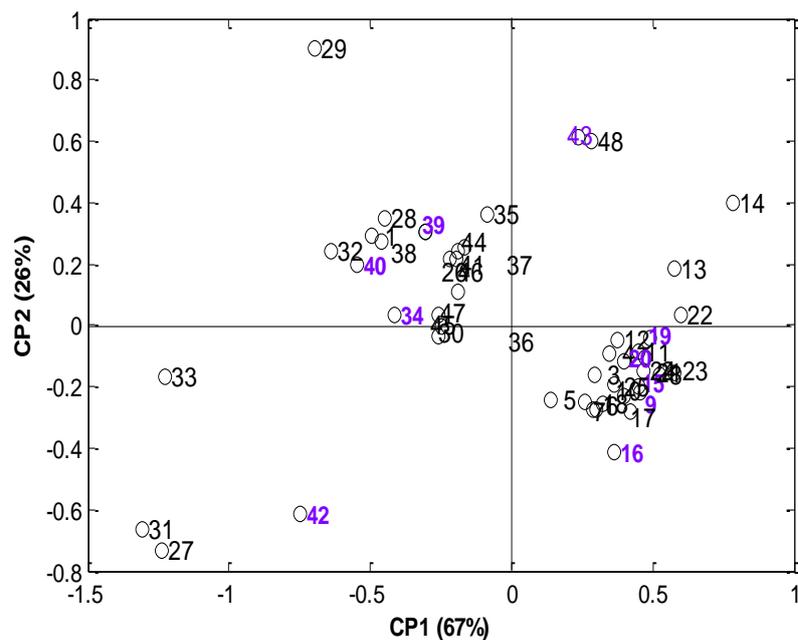


Figura 26 – Gráfico de escores para os dados espectrais.

Foram testados modelos em várias regiões espectrais e em várias condições e o melhor modelo PLS foi obtido na faixa de 1360 a 1430 nm, conforme a Figura 27. Essa região corresponde ao primeiro sobretom do estiramento de ligações OH isoladas, que podem ser associadas a moléculas de água presentes no óleo, pois a água é produto de degradação do sistema isolante (tanto papel, quanto óleo), sendo uma das variáveis de maior relevância no monitoramento desse sistema, pelo fato de acelerar o processo de degradação [66-68]. Essa vibração pode também ser associada à presença de moléculas de 5-hidroximetil-2-furfuraldeído, que é o principal composto furânico a contaminar o óleo em função da degradação do papel [69]. Os compostos furânicos são os principais produtos de degradação da celulose do papel Kraft [11].

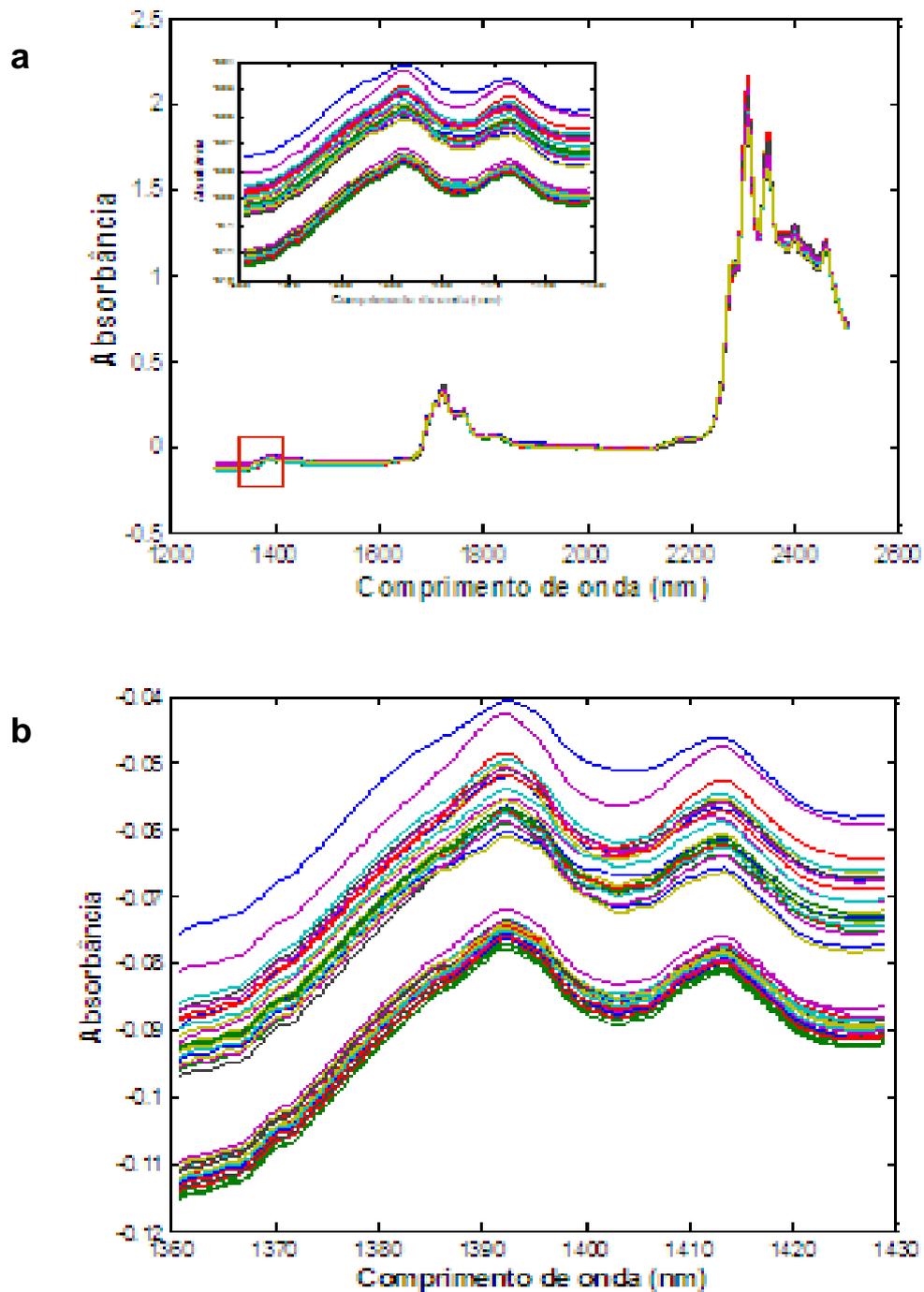


Figura 27 – Espectros das 48 amostras de óleos na região NIR (a). Faixa espectral selecionada de 1360 a 1430 nm (b).

Os espectros de NIR foram pré-tratados usando alisamento *Savitsky-Golay* com janela de 15 pontos e ajuste de função quadrática, com o objetivo de diminuir o ruído experimental, aumentando a razão sinal-ruído. Após o pré-tratamento, foram testadas várias faixas espectrais até se encontrar o melhor modelo de calibração. Outros pré-tratamentos foram testados, dentre eles a primeira derivada, segunda derivada e a correção de espalhamento multiplicativa (MSC, *Multiple Scatter Correction*), porém os erros absolutos foram muito altos. O modelo de calibração PLS foi construído com 8 variáveis latentes, que explicaram 100% da variância do bloco X e 77,75% do bloco Y. O erro médio quadrático de previsão (RMSEP) foi de 2,1 dina/cm. Na Tabela 5 são mostrados os valores da propriedade físico-química prevista pelo modelo e os erros percentuais. Como o método de referência [22] tem baixa precisão, os valores de referência da tensão interfacial foram obtidos somente com dois algarismos significativos. Em todo este trabalho, os valores previstos serão expressos usando a estimativa de um terceiro algarismo subscrita (o primeiro algarismo não significativo), de acordo com a notação usada em vários livros didáticos de química analítica [70].

Tabela 5 – Valores de tensão interfacial previstos pelo modelo PLS e erros percentuais

Nº da amostra	tensão interfacial experimental (dina/cm)	tensão interfacial previstos (dina/cm)	erro relativo(%)
9	40	43,8	7,1
15	24	24,6	2,6
16	29	30,5	5,2
19	20	18,4	-7,9
20	18	18,2	1,1
34	31	32,8	5,7
39	33	35,3	7,0
40	33	32,1	-2,8
42	43	47,4	10,2
43	42	41,0	-2,5

No gráfico da Figura 28 é mostrada a correlação entre os valores de tensão interfacial experimental e os valores previstos pelo modelo PLS. O valor do coeficiente de correlação obtido foi de $r = 0,985$.

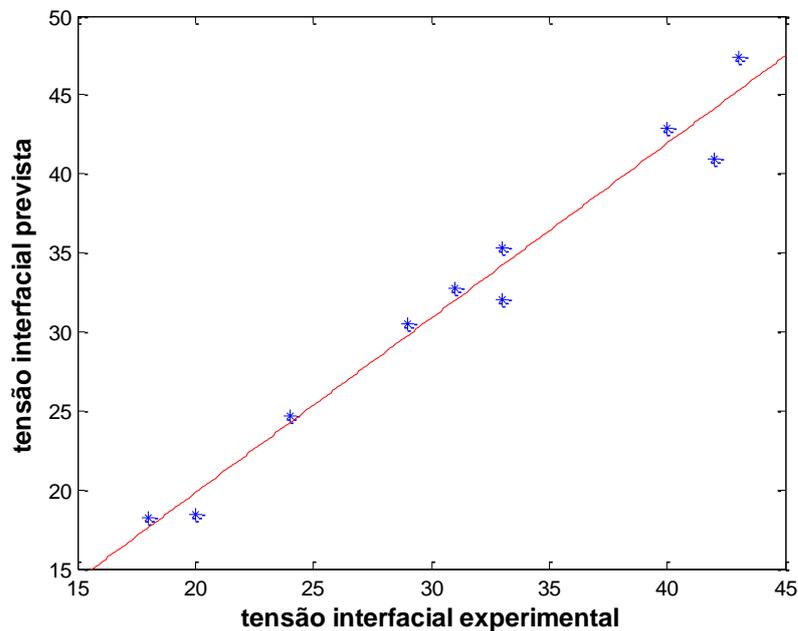


Figura 28 – Valores experimentais versus valores previstos pelo modelo PLS.

Através dos valores de tensão interfacial previstos mostrados na Tabela 2, conclui-se que foi possível modelar o conjunto de dados com erros de até aproximadamente 10%, o que faz com que a técnica de espectroscopia NIR possa ser utilizada no monitoramento dos sistemas isolantes, uma vez que esta é não destrutiva, rápida e utiliza pequena quantidade de amostra.

5.2. ANÁLISE DE IMAGENS

Foi montado um arranjo tridimensional empilhando as matrizes médias das imagens para formar um arranjo \underline{X} de dimensões (48, 200, 200). Esse arranjo de dados foi dividido em um conjunto de calibração (38 amostras) e um de previsão (10 amostras). No modelo PARAFAC, ao contrário do modelo NPLS, as amostras de calibração e previsão são decompostas conjuntamente. Portanto nesse caso, a matriz \underline{X} , contendo todas as 48 amostras, foi decomposta e uma matriz de escores contendo somente as 38 amostras de calibração foi correlacionado com um vetor coluna \underline{y} (38,1), o qual contém os valores da propriedade físico-química tensão

interfacial para essas amostras. A seguir, são mostrados os resultados para os dois modelos.

5.2.1. Modelo PARAFAC

O modelo PARAFAC foi construído usando 38 amostras para o conjunto de calibração e 10 amostras para o conjunto de previsão, com a mesma seleção do modelo usando NIR. Na construção do modelo os dados não foram autoescalados e nem centrados na média. Testou-se o número de fatores de 1 a 8, sendo que o melhor modelo foi obtido com 6 fatores e explicou 99.86% da variância dos dados.

O arranjo de dados de imagens tratado, além de muito grande, não é explicado por nenhuma lei físico-química, tornando a modelagem um pouco complexa. Devido a isso, não foi usada nenhuma restrição na construção do modelo. O erro médio quadrático de previsão (RMSEP) foi de 3,9 dinas/cm. Na tabela 3 são mostrados os valores da propriedade físico-química previstos pelo modelo e os erros percentuais.

Tabela 6 – Valores de tensão interfacial previstos pelo modelo PARAFAC e erros percentuais.

Nº da amostra	tensão interfacial experimental (dina/cm)	tensão interfacial previstos (dina/cm)	erro relativo(%)
9	40	42,0	5,0
15	24	32,4	35,0
16	29	26,4	-9,0
19	20	19,4	-3,0
20	18	19,8	10,0
34	31	38,4	23,9
39	33	28,8	-12,7
40	33	32,5	-1,5
42	43	43,2	0,5
43	42	42,0	0,0

Os valores de tensão interfaciais previstos mostrados na Tabela 6 apresentam erros relativamente baixos, se considerarmos que para o conjunto de dados analisado tanto as matrizes de óleos quanto a propriedade físico-química são sistemas reais, ou seja, trata-se de uma calibração natural, e o objetivo da empresa

energética no monitoramento do sistema isolante não é avaliar a precisão da medida, pois a degradação do sistema isolante é lenta. O objetivo da empresa ao fazer as medidas físico-químicas é, em alguns aspectos, qualitativo, mesmo porque o laboratório de análises mantém um histórico de todos os transformadores do estado de Goiás, contendo as análises feitas desde que o transformador iniciou seu tempo de vida. Além disso, as matrizes de imagens são complexas e de difícil interpretação, como pode ser observado no gráfico da Figura 29, no qual se encontra os pesos (*loadings*) para o modelo PARAFAC.

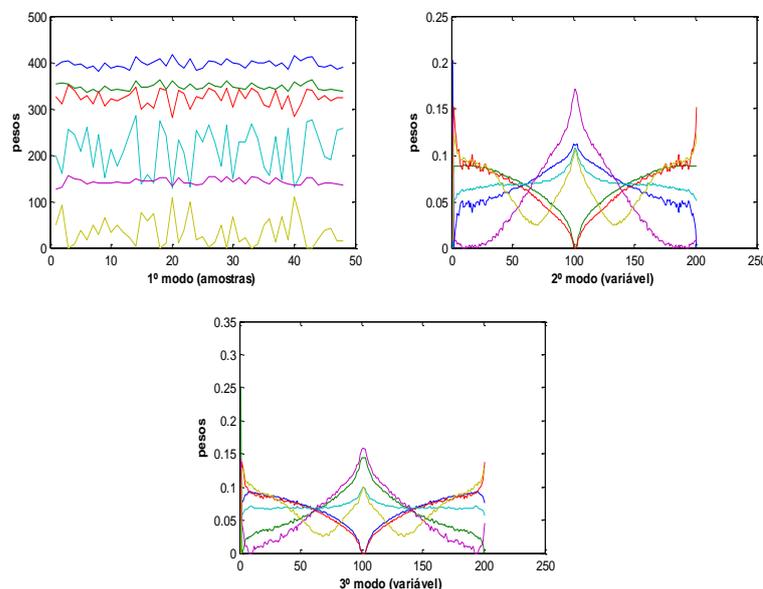


Figura 29 – Gráfico de pesos para o modelo PARAFAC.

Diante desses aspectos, a quantificação da tensão interfacial realizada pela empresa é feita sem avaliar a precisão, a partir de uma única medida para cada amostra. Isso se deve à rotina de trabalho dos técnicos da empresa ser muito tumultuada, pois existe apenas um laboratório para monitorar as análises dos sistemas de todos os transformadores do estado, mas principalmente pelo fato da medida dessa propriedade ser muito demorada, além dos aspectos mencionados anteriormente a respeito do critério de avaliação feita pela CELG.

Dessa forma, é possível afirmar que através da aplicação do PARAFAC foi possível modelar bem o conjunto de dados de imagens, mesmo havendo 3 amostras de previsão com erros acima de 10%. É possível inferir que esse erro

relativamente maior foi devido ao conjunto de amostras de calibração ser pequeno sendo composto por apenas 38 amostras de óleo, e os fatores responsáveis pela degradação do óleo serem muito diversificados. No gráfico da Figura 30 é mostrada a correlação entre os valores de tensão interfacial experimental e os valores previstos pelo modelo PARAFAC. O coeficiente de correlação obtido foi de $r = 0,898$.

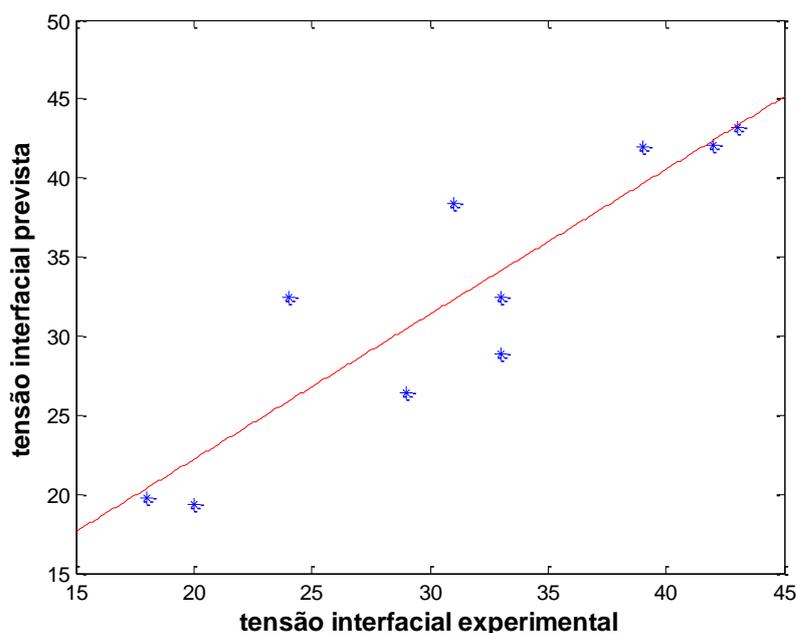


Figura 30 – Valores experimentais versus valores previstos pelo modelo PARAFAC.

5.2.2. Modelo NPLS

Foi construído um modelo NPLS para esse conjunto de dados de imagens, buscando prever o valor da propriedade físico-química com o menor erro percentual. O modelo NPLS foi construído usando validação cruzada para selecionar o melhor modelo. Variou-se o número de fatores de 1 até 6, e o melhor modelo foi obtido com 3 fatores. A Tabela 7 mostra o percentual de variância explicada por cada fator para as matrizes \underline{X} e \underline{Y} .

Tabela 7 – Variância explicada pelas matrizes **X** e **Y** para o modelo NPLS para 3 fatores.

Número de fatores	Bloco X	Bloco Y
1	43,64	18,06
2	44,42	38,37
3	46,08	57,67

O modelo foi construído com a mesma seleção de amostras usadas pelo modelo PARAFAC, ou seja, 38 amostras para o conjunto de calibração e 10 amostras para o conjunto de validação.

Na Tabela 8 são mostrados os valores da tensão interfacial previstos pelo modelo e os erros percentuais.

Tabela 8 – Valores de tensão interfacial previstos pelo modelo NPLS e erros percentuais.

Nº da amostra	tensão interfacial experimental (dina/cm)	tensão interfacial previstos (dina/cm)	erro relativo (%)
9	40	37,5	-6,3
15	24	25,2	5,0
16	29	28,9	-0,3
19	20	18,6	-7,0
20	18	18,0	0,0
34	31	35,5	14,5
39	33	32,3	-2,1
40	33	31,2	-5,5
42	43	38,1	-11,4
43	42	36,4	-13,3

Através dos valores de tensão interfacial previstos pelo modelo NPLS, mostrados na tabela 8, é possível concluir que esse método modelou o conjunto de dados de imagens com um menor erro relativo do que o modelo PARAFAC, com um erro médio quadrático de previsão (RMSEP) de 3,0 dina/cm. Esse método apresenta

a vantagem de decompor os dados das matrizes \mathbf{X} e \mathbf{Y} simultaneamente, além de ser um método mais simples de ser interpretado do que o modelo PARAFAC.

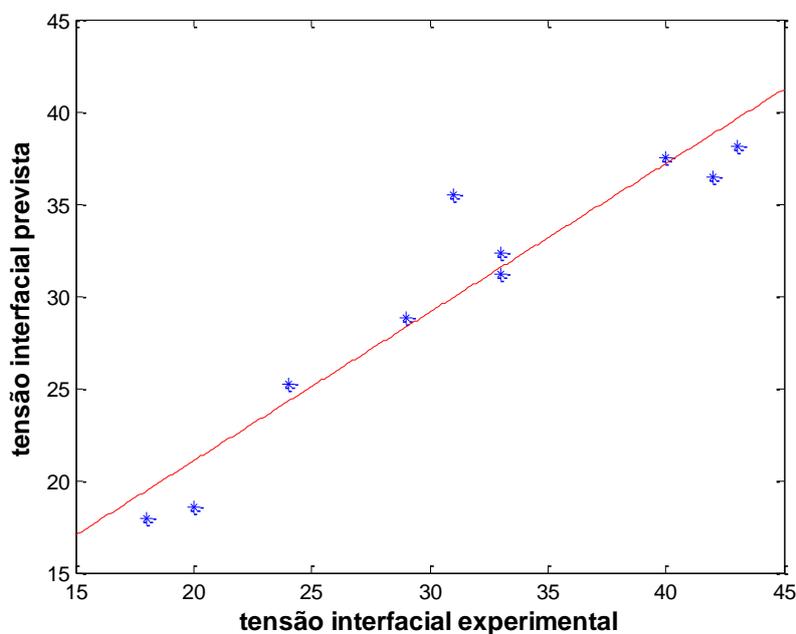


Figura 31 – Valores experimentais versus valores previstos pelo modelo NPLS.

No gráfico da Figura 31 é mostrada a correlação entre os valores de tensão interfacial experimental e os valores previstos pelo modelo NPLS. O coeficiente de correlação obtido foi de $r = 0,9497$. Huang e outros [27] utilizaram o método NPLS para modelar dados de imagens de queijos, os quais foram correlacionados com algumas propriedades específicas. O coeficiente de correlação encontrado por eles foi de $r = 0,91$. Assim, é possível afirmar que embora os conjuntos de calibração e validação para os dados de óleos sejam pequenos, o método NPLS conseguiu fazer boas previsões, mesmo não havendo estimativa da precisão na medida da propriedade físico-química.

5.3. OUTROS MÉTODOS DE TRATAR IMAGENS

5.3.1. Histogramas em escalas de cinza

Outra metodologia utilizada no tratamento desse conjunto de dados de imagens foi baseada na extração dos histogramas de frequência para imagens em tons de cinzas. Para cada amostra foi obtido um histograma médio de frequência, variando de 0 a 256 tons de cor entre o branco e o preto. Dessa forma, foi construído um modelo de PLS com 4 variáveis latentes, que explicaram 38,57% da variância do bloco X e 69,14% do bloco Y. O erro médio quadrático de previsão (RMSEP) foi de 3,1 dina/cm. Na Tabela 9 são mostrados os valores da propriedade físico-química prevista pelo modelo e os erros percentuais.

Tabela 9 – Valores de tensão interfacial previstos pelo modelo PLS dos histogramas de frequências para as imagens em escala de cinza e erros percentuais.

Nº da amostra	tensão interfacial experimental (dina/cm)	tensão interfacial previstos (dina/cm)	erro relativo (%)
9	40	43,6	9,0
15	24	28,0	16,5
16	29	29,5	1,7
19	20	20,2	0,9
20	18	16,6	-7,9
34	31	34,4	11,1
39	33	28,0	-15,1
40	33	29,3	-11,1
42	43	39,9	-7,2
43	42	40,8	-2,9

No gráfico da Figura 32 é mostrada a correlação entre os valores de tensão interfacial experimental e os valores previstos pelo modelo PLS para os histogramas de frequência das imagens em escala de cinza. O coeficiente de correlação obtido foi de $r = 0.9343$.

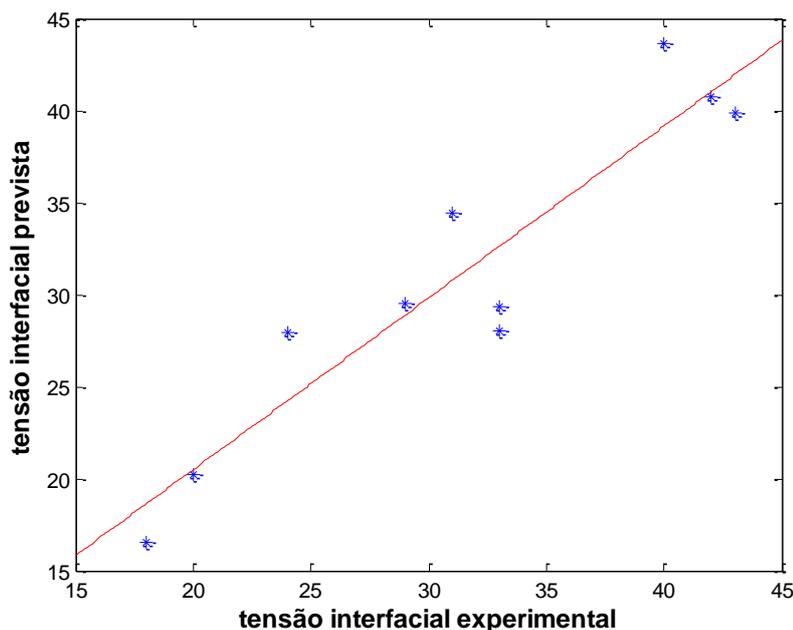


Figura 32 – Valores experimentais versus valores previstos pelo modelo PLS para os histogramas de frequência das imagens em escala de cinzas.

5.3.2. Histogramas RGB

Foram extraídos histogramas de frequência para os canais RGB para as imagens das 48 amostras de óleos, sendo obtida uma matriz contendo 48 linhas e 768 colunas referentes aos três canais R (0 a 255 tons de cor), G (256 a 512 tons de cor) e B (513 a 768 tons de cor). Foi construído um modelo PLS para correlacionar os histogramas de frequências dos canais de cor com a propriedade físico-química tensão interfacial. O modelo foi construído com 5 variáveis latentes explicando 68,76% da variância do bloco X e 91,85% da variância do bloco Y. O RMSEP foi de 3,6 dina/cm e os valores da propriedade físico-química prevista pelo modelo e os erros percentuais são mostrados na Tabela 10.

Tabela 10 – Valores de tensão interfacial previstos pelo modelo PLS dos histogramas de freqüências para os canais RGB e erros percentuais.

Nº da amostra	tensão interfacial experimental (dina/cm)	tensão interfacial previstos (dina/cm)	erro relativo (%)
9	40	44,1	10,2
15	24	24,1	0,5
16	29	30,3	4,6
19	20	16,8	-22,2
20	18	15,6	-6,7
34	31	37,3	20,5
39	33	33,2	0,7
40	33	26,7	-19,2
42	43	40,1	-6,7
43	42	40,5	-3,5

Na Figura 33 é mostrado o gráfico de correlação entre os valores de tensão interfacial experimental e os valores previstos pelo modelo PLS para os histogramas de freqüência dos canais RGB. O coeficiente de correlação obtido foi de $r = 0.9266$.

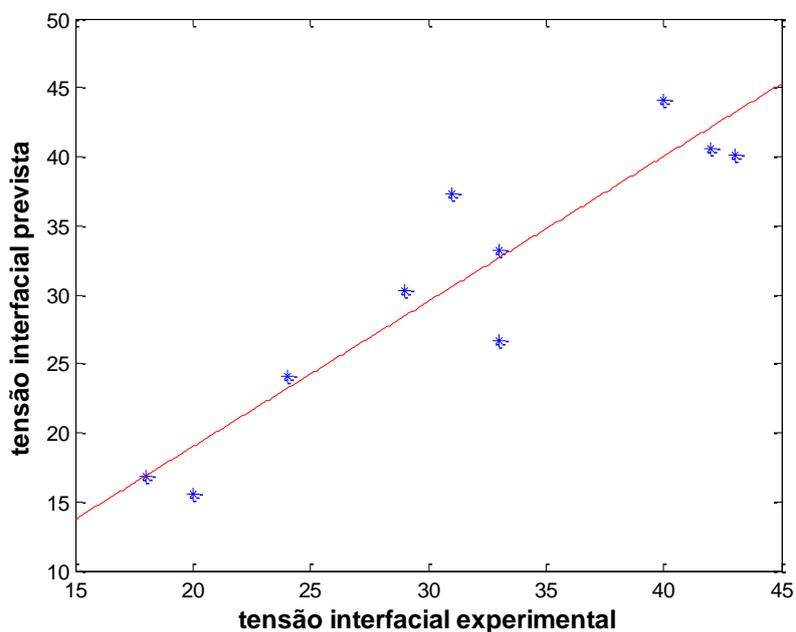


Figura 33 – Valores experimentais versus valores previstos pelo modelo PLS para os histogramas de freqüência dos canais RGB.

Tratar imagens utilizando histogramas de frequência tem como vantagem a obtenção de modelos mais simples que podem ser analisados por métodos de primeira ordem, não sendo necessária a utilização de etapas de pré-processamentos das imagens. Dessa forma, o objetivo foi mostrar que é possível prever a propriedade tensão interfacial utilizando métodos de primeira e segunda ordem para os dados de imagens.

Portanto, a maior vantagem de usar técnicas como a espectroscopia NIR e a análise de imagens é a possibilidade de monitorar a degradação do óleo sem a necessidade de fazer medidas físico-químicas, utilizando uma quantidade muito pequena de amostra e gastando pouco tempo. Além disso, ressalta-se o fato de essas técnicas não serem destrutivas e o método de referência [22] ser obsoleto, tendo sido publicado em 1965 sem reformulação nos últimos anos.

5.3.3. PLS desdobrado

Para esse mesmo conjunto de dados foi construído um modelo PLS desdobrado (*unfold-PLS*), que se baseia no desdobramento de uma matriz tridimensional em uma matriz bidimensional. Esse método é uma alternativa para tratar dados de segunda ordem usando um modelo de primeira ordem, no qual a decomposição dos dados é bilinear. Os resultados encontrados na construção desse modelo apresentaram erros acima de 40% e, portanto não serão apresentados.

Capítulo 6

Conclusões e Perspectivas

6.1. CONCLUSÃO

Os resultados obtidos neste trabalho mostraram ser possíveis determinar a propriedade físico-química tensão interfacial em amostras de óleos usando as técnicas de espectroscopia NIR e de análise de imagens, juntamente com métodos de calibração de primeira e segunda ordem, respectivamente. Com isso, foi possível propor novos métodos para monitorar o sistema isolante de transformadores elétricos, considerando que a avaliação do óleo é uma forma indireta de monitorar o estado de degradação do papel, pois os métodos tradicionais de avaliação desse sistema, tanto do óleo quanto do papel, demandam muito tempo, além de gerar uma grande quantidade de resíduos. O modelo NIR-PLS forneceu os melhores resultados, com erros de previsão entre -7,9 e 10,2%, considerados aceitáveis pelas empresas de energia elétrica. Os métodos propostos são rápidos e não destrutivos, mostrando grandes vantagens sobre os métodos tradicionais, os quais demandam muito tempo, além de gerar uma grande quantidade de resíduos. Este trabalho abre perspectivas para um monitoramento direto (*on-line*) do sistema isolante.

6.2. PERSPECTIVAS

Outra técnica que pode ser utilizada para avaliar as condições de degradação do sistema isolante é a espectroscopia de ressonância magnética nuclear (RMN). Foram realizadas análises preliminares em 20 amostras de óleos isolantes fornecidos pela CELG e utilizando PCA foi possível classificar as amostras em muito degradadas, de degradação média e pouco degradadas, pois os perfis dos espectros são muito característicos. Assim, pode ser possível através de espectros de RMN prever a concentração dos compostos 2-furfuraldeído e 5-hidroximetil-2-furfuraldeído, que são os principais produtos da degradação do papel e se encontram dissolvidos no óleo [69].

A vantagem de utilizar RMN na quantificação de compostos furânicos em relação ao método já estabelecido, a cromatografia líquida, é a menor quantidade de reagente, um menor tempo na realização da análise e, principalmente, uma quantidade muito pequena de amostra (aproximadamente 2 mL), não gerando resíduos para serem tratados e descartados.

Embora o espectrômetro de RMN seja muito caro em relação ao cromatógrafo, a CELG tem mostrado grande interesse em trabalhar em parceria e até com a possibilidade de terceirizar essa análise, o que seria de grande vantagem tanto para a universidade quanto para a empresa.

Referências Bibliográficas

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] [http:// www.dee.ufrn.br/~joao/apostila/cap09.htm](http://www.dee.ufrn.br/~joao/apostila/cap09.htm) - acessada em 30 de abril de 2009.
- [2] Sá, A. L. C. R. “Avaliação da degradação do papel tipo Kraft por espectroscopia na região do infravermelho e calibração multivariada”. Dissertação de Mestrado, Universidade Federal de Goiás, Goiânia, 2003.
- [3] Paixão, L. A. “Avaliação da qualidade do óleo isolante em transformadores com o emprego da função discriminante quadrática”. Dissertação de Mestrado, Universidade Federal do Paraná, Curitiba, 2006.
- [4] Laurentino, A. C. “Regeneração de óleo mineral isolante usando como adsorvente o compósito quitosana/fosfato de cálcio amorfo carbonatado”, Dissertação de Mestrado, Universidade Federal de Santa Catarina, Florianópolis, 2005.
- [5] Muthanna, K. T.; Sarkar, A.; Das, K.; Waldner, K. “Transformer Insulation Life Assessment”. *IEEE Transactions on Power Delivery*, **2006**, 21, 150.
- [6] Quadros, A. M. “Papel Nomex utilizado em transformadores: avaliação das características e degradação térmica”. Dissertação de Mestrado, Universidade Federal de Santa Catarina, Florianópolis, 2006.
- [7] Lundgaard, L. E.; Hansen, W.; Linhjell, D.; Painter, T. J. “Aging of oil-impregnated paper in power transformers”. *IEEE Transactions on Power Delivery*, **2004**, 19, 230.
- [8] Saha, T. K.; Purkait, P. “Understanding the impacts of moisture and thermal ageing on transformer’s insulation by dielectric response and molecular weight measurements”. *IEEE, Transactions on Dielectrics and Electrical Insulation*, **2008**, 15, 569.
- [9] Shroff, D. H.; Stannett, A. W. “A review of paper aging in power transformers”. *IEE Proceedings*, **1985**, 132, 312.
- [10] Martins, M. A. G. “Furfuraldeído - um indicador prático da degradação térmica do papel Kraft de transformadores”. *Ciência e Tecnologia dos Materiais*, **2007**, 19, 1.
- [11] Emsley, A. M.; Stevens, G. C. “Review of chemical indicators of degradation of cellulosic electrical paper insulation in oil-filled transformers”. *IEE proceedings. Science, Measurement and Technology*, **1994**, 141, 324.
- [12] Emsley, A. M.; Xiao, X.; Heywood, R. J.; Ali, M. “Degradation of cellulosic insulation in power transformers. Part 2: Formation of furan products in insulating oil”. *IEE Proceedings - Science, Measurement and Technology*, **2000**, 147, 110.

- [13] Morais, R. M.; Mannheimer, W. A.; Carballeira, M.; Noualhaguet, J. C. "Furfural analysis for assessing degradation of thermally upgraded papers in transformer insulation". *IEEE, Transactions on Dielectrics and Electrical Insulation*, **1999**, 6, 159.
- [14] Duval, M. "Dissolved gas analysis: It can save your transformer". *IEEE Electrical Insulation Magazine*, **1989**, 5, 22.
- [15] Saha, T. K. "Review of modern diagnostic techniques for assessing insulation condition in aged transformers". *IEEE, Transactions on Dielectrics and Electrical Insulation*, **2003**, 10, 903.
- [16] Bolhuis, J. V. V.; Gulski, E.; Smit, J. J. "Monitoring and diagnostic of transformer solid insulation". *IEEE Transactions on Power Delivery*, **2002**, 17, 528.
- [17] Associação Brasileira de Normas Técnicas. NBR 10576: Óleo mineral isolante de equipamentos elétricos – Diretrizes para supervisão e manutenção. ABNT, 2006.
- [18] Associação Brasileira de Normas Técnicas NBR 10710: Determinação do teor de água - Ensaio, ABNT, p. 1-9, 1988.
- [19] Associação Brasileira de Normas Técnicas NBR 10859: Determinação da rigidez dielétrica - Ensaio, ABNT, p. 1-5, 1989.
- [20] Associação Brasileira de Normas Técnicas NBR 12133: Determinação do fator de perdas dielétricas - Ensaio, ABNT, p. 1-7, 1991
- [21] Associação Brasileira de Normas Técnicas NBR 14248: Determinação do índice de neutralização - Método do indicador - Ensaio, ABNT, p. 1-9, 1998.
- [22] Associação Brasileira de Normas Técnicas NBR 6234: Determinação da tensão interfacial de óleo água, ABNT, 1965.
- [23] Kim, H. C.; Burgess, D. J. "Prediction of interfacial tension between oil mixtures and water". *Journal of Colloid and Interface Science*, **2001**, 241, 509.
- [24] Geladi, P.; Grahn, H. "Multivariate image analysis". Chichester: John Wiley, 1996.
- [25] Ribeiro, F. A. L. "Análise de imagens gray-scale utilizando métodos quimiométricos". Tese de Doutorado, Universidade Estadual de Campinas, 2007.
- [26] Gonzalez, R. C.; Woods, R. E. "Processamento de imagens digitais". São Paulo: Edgard Blucher, 2000.

- [27] Pataca, L. C. M. “Análises de mel e própolis utilizando métodos Quimiométricos de Classificação e Calibração”. Tese de Doutorado, Universidade Estadual de Campinas, 2006.
- [28] Gurden, S. P.; Lage, E. M.; Faria, C. G.; Joekes, I.; Ferreira, M. M. C. “Analysis of video images from a gas-liquid transfer experiment: a comparison of PCA and PARAFAC for multivariate images analysis”. *Journal of Chemometrics*, **2003**, *17*, 400.
- [29] Godinho, M. S.; Pereira, R. O.; Ribeiro, K. O.; Schimidt, F.; Oliveira, A. E.; Oliveira, S. B. “Classificação de refrigerantes através de Análise de Imagens e Análise de Componentes Principais (PCA)”. *Química Nova*, **2008**, *31*, 1485
- [30] Gonzalez, R. C.; Woods, R. E. “Digital Image Processing”. New Jersey: Prentice Hall, 1992.
- [31] Pratt, W. K. “Digital Image Processing”. New York: John Wiley, 2001.
- [32] Geladi, P.; Bengtsson, E.; Esbensen, K.; Grahn, H. “Image analysis in chemistry I. Properties of images, greylevel operations, the multivariate image”. *Trends in Analytical Chemistry*, **1992**, *11*, 41.
- [33] Huang, J.; Wium, H.; Qvist, K. B.; Esbensen, K. H. “Multi-way methods in image analysis – relationships and applications”. *Chemometrics and Intelligent Laboratory Systems*, **2003**, *66*, 141.
- [34] Reed, T. R. “Digital Image Sequence Processing, Compression, and Analysis”. New York: CRC Press LLC, 2005.
- [35] Pasquini, C. “Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications”. *Journal of the Brazilian Chemical Society*, **2003**, *14*, 198.
- [36] Barton, F. E. “Theory and Principles of near infrared spectroscopy”. *Spectroscopy Europe*, **2002**, *14*, 12.
- [37] Blanco, M.; Villarroya, I. “NIR spectroscopy: a rapid-response analytical tool”. *Trends in Analytical Chemistry*, **2002**, *21*, 240.
- [38] Borges, W. N. “Parâmetros de qualidade de lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada e seleção de variáveis”. Tese de Doutorado, Universidade Estadual de Campinas, 2005.
- [39] Nicolai, B. M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K. I.; Lammertyn, J. “Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review”. *Postharvest Biology and Technology*, **2007**, *46*, 99.

- [40] Miller, C. E. "Chemometrics for on-line spectroscopy applications - Theory and Practice". *Journal of Chemometrics*, **2000**, *14*, 513.
- [41] Brereton, R. G. "Chemometrics: Data Analysis for the Laboratory and Chemical Plant". Chichester: John Wiley, 2003.
- [42] Otto, M. "Chemometrics". Weinheim: Wiley-VCH, 1999.
- [43] Geladi, P.; Sethson, B.; Nystrom, J.; Lillhonga, T.; Lestander, T.; Burger, J. "Chemometrics in spectroscopy Part. 2. Examples". *Spectrochimica Acta B*, **2004**, *59*, 1347.
- [44] Sena, M. M.; Collins, C.H.; Collins, K. E. "Aplicação de métodos quimiométricos na especiação de Cr(VI) em solução aquosa". *Química Nova*, **2001**, *24*, 331.
- [45] Martens, H.; Naes, T. "Multivariate Calibration". New York: John Wiley, 1989.
- [46] Wold, S.; Sjostrom, M.; Eriksson, L. "PLS-regression: a basic tool of chemometrics". *Chemometrics and Intelligent Laboratory Systems*, **2001**, *58*, 109.
- [47] 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*, **2005**, *28*, 910.
- [48] Trygg, J. "Prediction and spectral profile estimation in multivariate calibration". *Journal of Chemometrics*, **2004**, *18*, 166.
- [49] Haswell, S. J.; Walmsley, A. D. "Chemometrics: the issues of measurement and modelling". *Analytica Chimica Acta*, **1999**, *400*, 399.
- [50] Braga, J. W. B. "Aplicação e validação de modelos de calibração de segunda ordem em Química Analítica". Tese de Doutorado, Universidade Estadual de Campinas, 2008.
- [51] Kruger, U.; Zhou, Y.; Wang, X.; Rooney, D. "Robust partial least square regression: Part I, algorithmic developments". *Journal of Chemometrics*, **2007**, *22*, 1.
- [52] Nadler, B.; Coifman, R. R. "Partial least squares, Beer's law and the net analyte signal: statistical modeling and analysis". *Journal of Chemometrics*, **2005**, *19*, 45.
- [53] Indah, U. "A twist to partial least squares regression". *Journal of Chemometrics*, **2005**, *19*, 32.
- [54] Xu, Q.; Liang, Y.; Shen, H. "Generalized PLS regression". *Journal of Chemometrics*, **2001**, *15*, 135.

- [55] Bereton, R. G. "Introduction to multivariate calibration in analytical". *The Analyst*, **2000**, 125, 2125
- [56] Ferreira, M. M. C.; Antunes, A. M.; Melgo, M. S.; Volpe, P. L. O. "Quimiometria I: Calibração Multivariada, um tutorial". *Química Nova*, **1999**, 22, 724.
- [57] Janne, K.; Pettersen, J.; Lindberg, N. O.; Lundstedt, T. "Hierarchical principal component analysis (PCA) and projection to latent structure (PLS) technique on spectroscopic data as a data pretreatment for calibration". *Journal of Chemometrics*, **2001**, 15, 203.
- [58] Bro, R. "PARAFAC. Tutorial and application". *Chemometrics and Intelligent Laboratory Systems*, **1997**, 38, 149.
- [59] Acar, E.; Yener, B. "Unsupervised Multiway Data Analysis: A Literature Survey". *IEEE Transactions on Knowledge and Data Engineering*, **2009**, 21, 7.
- [60] Sena, M. M. "Aplicação de métodos quimiométricos de primeira e segunda ordem na determinação direta de fármacos por espectroscopia molecular". Tese de Doutorado, Universidade Estadual de Campinas, 2004.
- [61] Bro, R. "Multiway Analysis in the Food Industry: Models, Algorithms and Applications", Tese de Doutorado, Universidade de Amsterdam, 1998.
- [62] Anderson, C.A.; Bro, R. "The N-way Toolbox for MATLAB". *Chemometrics and Intelligent Laboratory Systems*, **2000**, 52, 1.
- [63] Khatri, C. G.; Rao, C. R.; *Generalized Inverse of Matrices and its Applications*, John Wiley: New York, 1968.
- [64] Bro, R.; Kiers, H. A. L.; "A new efficient method for determining the number of components in PARAFAC models". *Journal of Chemometrics*, **2003**, 17, 274.
- [65] Bro, R. "Multiway calibration Multilinear PLS". *Journal of Chemometrics*. **1996**, 10, 47.
- [66] Percherancier, J. P.; Vuarchex, P. J. "Fourier Transform Infrared (FT-IR) Spectrometry to Detect Additives and Contaminants in Insulating Oils". *IEEE Electrical Insulation Magazine*, **1998**, 14, 23.
- [67] Kóréh, O.; Torkos, K.; Mahara, M. B.; Borossay, J. "Study of Water Clusters in Insulating Oils by Fourier Transform Infrared Spectroscopy". *IEEE Transactions on Dielectrics and Electrical Insulation*, **1998**, 5, 896.
- [68] Ali, M.; Emsley, A. M.; Herman, H.; Heywood, R. J. "Spectroscopic Studies of the Ageing of Cellulosic Paper". *Polymer*, **2001**, 42, 2893.

[69] Madhavan, K. S.; Murthy, T. S. R.; Sethuraman, R. "Estimation of Degree of Polymerization and Residual Age of Transformers Based on Furfural Levels in Insulating Oil through Generalized Regression Neural Networks". Em: Reusch, B. (Ed.), "Computational Intelligence. Theory and Applications", Berlin: Springer, 2006, pp. 751-756.

[70] Harris, D. C. "Análise Química Quantitativa", 6^a ed., Rio de Janeiro: LTC, 2005, pp. 66.

Apêndice

APÊNDICE – Rotina utilizada para processamento das imagens obtidas das amostras de óleos

Para cada amostra (obtenção de 9 imagens)

Imagem 1

```
I=imread('amostra1.1.bmp'); leitura da imagem
img=l(250:450,250:450,:); seleção de uma parte da imagem
img1=rgb2gray(img); conversão da imagem para escala de cinza
h=fspecial('unsharp'); filtro especial
i=imfilter(img1,h);
F1=fft2(i);
T1=log(F1); transformada de Fourier bidimensional
t1=log(abs(F1)); retirada da parte real
clear img I i img1 h
```

Imagem 2

```
I=imread('amostra1.2.bmp');
img=l(250:450,250:450,:);
img2=rgb2gray(img);
h=fspecial('unsharp');
i=imfilter(img2,h);
F2=fft2(i);
T2=log(F2);
t2=log(abs(F2));
clear img I i img2 h
```

Imagem 3

```
I=imread('amostra1.3.bmp');
img=l(250:450,250:450,:);
img3=rgb2gray(img);
h=fspecial('unsharp');
i=imfilter(img3,h);
F3=fft2(i);
T3=log(F3);
t3=log(abs(F3));
clear img I i img3 h
```

Imagem 4

```
I=imread('amostra2.1.bmp');
```

```
img=l(250:450,250:450,:);  
img4=rgb2gray(img);  
h=fspecial('unsharp');  
i=imfilter(img4,h);  
F4=fft2(i);  
T4=log(F4);  
t4=log(abs(F4));  
clear img I i img4 h
```

Imagem 5

```
I=imread('amostra2.2.bmp');  
img=l(250:450,250:450,:);  
img5=rgb2gray(img);  
h=fspecial('unsharp');  
i=imfilter(img5,h);  
F5=fft2(i);  
T5=log(F5);  
t5=log(abs(F5));  
clear img I i img5 h
```

Imagem 6

```
I=imread('amostra2.3.bmp');  
img=l(250:450,250:450,:);  
img6=rgb2gray(img);  
h=fspecial('unsharp');  
i=imfilter(img6,h);  
F6=fft2(i);  
T6=log(F6);  
t6=log(abs(F6));  
clear img I i img6 h
```

Imagem 7

```
I=imread('amostra3.1.bmp');  
img=l(250:450,250:450,:);  
img7=rgb2gray(img);  
h=fspecial('unsharp');  
i=imfilter(img7,h);  
F7=fft2(i);  
T7=log(F7);  
t7=log(abs(F7));  
clear img I i img1 h
```

Imagem 8

```

I=imread('amostra3.2.bmp');
img=I(250:450,250:450,:);
img8=rgb2gray(img);
h=fspecial('unsharp');
i=imfilter(img8,h);
F8=fft2(i);
T8=log(F8);
t8=log(abs(F8));
clear img I i img8 h

```

Imagem 9

```

I=imread('amostra3.3.bmp');
img=I(250:450,250:450,:);
img9=rgb2gray(img);
h=fspecial('unsharp');
i=imfilter(img9,h);
F9=fft2(i);
T9=log(F9);
t9=log(abs(F9));
clear img I i img9 h

```

```

for i=1:9
    arq=['F' num2str(i)];
    fft(:,i)=eval([arq]);
end

```

```

for i=1:9
    arq=['T' num2str(i)];
    FF(:,i)=eval([arq]);
end

```

```

for i=1:9
    arq=['t' num2str(i)];
    tt(:,i)=eval([arq]);
end

```

```

FF1=mean(FF,3);
FFT1=mean(fft,3);
tt1=mean(tt,3);

```

cálculo das médias