



**UNIVERSIDADE FEDERAL DA PARAÍBA**  
**CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA**  
**DEPARTAMENTO DE QUÍMICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**DISSERTAÇÃO DE MESTRADO**

**DETERMINAÇÃO SIMULTÂNEA DE ACIDEZ, ÍNDICE DE  
REFRAÇÃO E VISCOSIDADE EM ÓLEOS VEGETAIS  
USANDO ESPECTROMETRIA NIR, CALIBRAÇÃO  
MULTIVARIADA E SELEÇÃO DE VARIÁVEIS**

**Alessandra Félix da Costa Pereira**

**1º Orientador: Prof. Dr. Mário César Ugulino de Araújo**

**2º Orientador: Prof. Dr. Sérgio Ricardo Bezerra dos Santos**

*João Pessoa – PB – Brasil*

*Março/2007*



**UNIVERSIDADE FEDERAL DA PARAÍBA**  
**CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA**  
**DEPARTAMENTO DE QUÍMICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**DISSERTAÇÃO DE MESTRADO**

**DETERMINAÇÃO SIMULTÂNEA DE ACIDEZ, ÍNDICE DE REFRAÇÃO E  
VISCOSIDADE EM ÓLEOS VEGETAIS USANDO ESPECTROMETRIA  
NIR, CALIBRAÇÃO MULTIVARIADA E SELEÇÃO DE VARIÁVEIS**

**Alessandra Félix da Costa Pereira\***

Dissertação apresentada como  
requisito para obtenção do título de  
Mestre em Química pela Universidade  
Federal da Paraíba.

**1º Orientador: Prof. Dr. Mário César Ugulino de Araújo**

**2º Orientador: Prof. Dr. Sérgio Ricardo Bezerra dos Santos**

**\*Bolsista CNPQ.**

*João Pessoa – PB - Brasil*

*Março/2007*

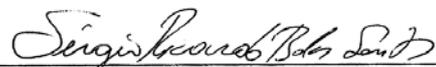
**Determinação simultânea de acidez, índice de refração e viscosidade em óleos vegetais usando espectrometria NIR, calibração multivariada e seleção de variáveis.**

Aprovada pela banca examinadora:



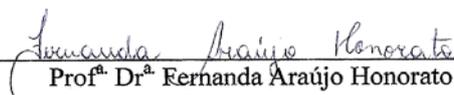
---

Prof. Dr. Mário César Ugulino de Araújo  
Orientador/Presidente



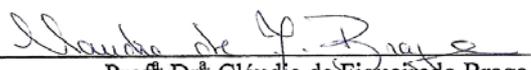
---

Prof. Dr. Sérgio Ricardo Bezerra dos Santos  
2º. Orientador



---

Profª Drª Fernanda Araújo Honorato  
Examinadora



---

Profª Drª Cláudia de Figueiredo Braga  
Examinadora

*A Deus, por sua infinita bondade e misericórdia. Por ter posto em minha vida uma família abençoada e amigos maravilhosos. Por tudo que tenho e que sou!*

*À minha amada e querida mãe Valdevina, por todo seu carinho, compreensão e amor, por ter sempre acreditado em mim e ser o meu maior exemplo de vida!*

*Ao meu amado e querido pai Paulo Roberto, por seu amor, apoio e carinho.*

*Amo-os profundamente!*

*Com carinho, dedico.*

## AGRADECIMENTOS

---

- Em especial, aos meus queridos irmãos Ana Paula, Leonardo e Ludmilla, pela convivência maravilhosa, amizade e apoio. Amo vocês!
- A toda minha família, pelo incentivo e por ter sempre acreditado nas minhas potencialidades. Em especial, a minha querida avó Adalzira Fabião (Dadá) (*in memoriam*).
- Ao meu querido sobrinho Matheus, que iluminou muitos dos meus dias com os seus sorrisos.
- Ao Prof. Dr. Mário Ugulino, pela orientação, apoio e amizade.
- Ao Prof. Dr. Sérgio Santos pela orientação, paciência, apoio científico-emocional e, principalmente, pela valiosa amizade.
- Ao amigo Francisco Gambarra, pela grande contribuição, apoio e compromisso.
- Ao amigo Wellington Lyra, pela contribuição no início deste trabalho.
- Em especial, ao grande amigo Márcio Coelho, pela incansável ajuda desde o início deste trabalho, pelas horas de dedicação, paciência, orientação, companheirismo, compromisso e, acima de tudo, pela nossa amizade. Muito obrigada por tudo!
- A todos os amigos da família LAQA, que me proporcionaram momentos maravilhosos e inesquecíveis (extraíndo de mim muitos sorrisos) e, acima de tudo, pela convivência harmoniosa e amizades construídas e cultivadas: Amália Gama, Elaine Cristina, Heronides Dantas, Ilanna Campelo (txi), Josenita Alves, Luciano Farias (Lulu), Osmundo Dantas, Pedro Germano (Big D), Simone Simões, Sueny Freitas e Valdomiro Lacerda. Em especial, aos meus queridos amigos que representam muito em minha vida: Edilene Oliveira (Leninha), Glauciene Paula (Gal), Márcio Coelho, Pablo Nogueira (Pablito), Ricardo Alexandre (Xande4) e Sérgio Ricardo. Amo vocês!

## AGRADECIMENTOS

---

- Aos queridos amigos que, apesar da distância, sempre estiveram presentes me apoiando: Daniele Muniz, Herbert Luckwi, Janainy Gomes, Kelly Cristinne e Patrícia Ferreira.
- À coordenação de Pós-Graduação em Química e aos professores do Departamento de Química pela minha formação acadêmica.
- Ao CNPq, pela bolsa concedida.
- Finalmente, a todos aqueles que de alguma forma contribuíram para a realização deste gratificante trabalho.

**Muito obrigada!**

*"Ninguém acende uma candeia para cobri-la com um vaso ou para escondê-la debaixo da cama, mas a coloca no candelabro para que aqueles que entram possam ver a luz".*

**Jesus Cristo**

*(Lucas 8:16)*

---

|  |      |
|--|------|
| ÍNDICE DE FIGURAS.....   | iii  |
| ÍNDICE DE TABELAS.....   | v    |
| LISTA DE ABREVIATURAS E SIGLAS.....                            | vi   |
| RESUMO.....  | viii |
| ABSTRACT.....  | ix   |
| <br>   |      |
| CAPÍTULO I - INTRODUÇÃO.....                                   | 1    |
| 1.1. Óleos Vegetais.....                                       | 2    |
| 1.2. A Importância dos Óleos Vegetais para a Saúde Humana..... | 3    |
| 1.3. Produção de Óleos Vegetais.....                           | 4    |
| 1.3.1. Processo de Refinação dos Óleos Vegetais.....           | 6    |
| 1.4. Principais Óleos Vegetais Comercializados.....            | 7    |
| 1.4.1. Óleo de Soja.....                                       | 8    |
| 1.4.2. Óleo de Milho.....                                      | 10   |
| 1.4.3. Óleo de Girassol.....                                   | 12   |
| 1.4.4. Óleo de Canola.....                                     | 13   |
| 1.5. Análise de Óleos Vegetais.....                            | 15   |
| 1.5.1. Métodos de Análise de Óleos Vegetais.....               | 16   |
| 1.6. Espectroscopia no Infravermelho.....                      | 21   |
| 1.6.1. Espectroscopia no Infravermelho Próximo (NIR).....      | 23   |
| 1.7. Quimiometria.....   | 24   |
| 1.7.1. Métodos de Pré-Processamentos de Dados.....             | 25   |
| 1.7.2. Métodos de Calibração Multivariada.....                 | 25   |
| 1.7.2.1. Seleção de Variáveis.....                             | 29   |
| 1.7.2.1.1. Algoritmo <i>i</i> PLS.....                         | 29   |
| 1.7.2.1.2. Algoritmo das Projeções Sucessivas.....             | 30   |
| 1.8. Objetivo do trabalho.....                                 | 33   |
| <br>   |      |
| CAPÍTULO II - EXPERIMENTAL.....                                | 34   |
| 2.1. Amostras.....   | 35   |
| 2.2. Divisão do Conjunto de Amostras.....                      | 35   |
| 2.3. Análise dos Óleos Vegetais por Métodos de Referência..... | 35   |

---

|  |    |
|--|----|
| 2.4. Registro dos Espectros NIR.....                         | 37 |
| 2.5. Softwares.....  | 38 |
| <br>   |    |
| CAPÍTULO III – RESULTADOS E DISCUSSÕES.....                  | 39 |
| 3.1. Seleção da Região Espectral de Trabalho.....            | 40 |
| 3.2. Atribuição das Bandas dos Espectros NIR.....            | 40 |
| 3.3. Pré-Processamento dos Dados Espectrais.....             | 41 |
| 3.4. Elaboração dos Modelos PLS, <i>i</i> PLS e SPA-MLR..... | 42 |
| <br>   |    |
| CAPÍTULO IV – CONCLUSÕES.....                                | 50 |
| 4.1. Conclusões.....   | 51 |
| 4.2. Propostas Futuras.....                                  | 51 |
| <br>   |    |
| REFERÊNCIAS.....   | 52 |

---

|   |    |
|---|----|
| <b>Figura 1.1.</b> Esquema da reação de formação do óleo vegetal.....   | 2  |
| <b>Figura 1.2.</b> Estrutura do ácido oléico.....   | 4  |
| <b>Figura 1.3.</b> Participação das grandes regiões na produção de oleaginosas.....   | 5  |
| <b>Figura 1.4.</b> Informações de oleaginosas por região.....   | 8  |
| <b>Figura 1.5.</b> Diagrama de energia potencial para os osciladores harmônico (A) e anarmônico (B).....  | 23 |
| <b>Figura 1.6.</b> Representação do procedimento de uma calibração multivariada.....  | 25 |
| <b>Figura 1.7.</b> Representação de matrizes de dados (X e Y).....  | 26 |
| <b>Figura 1.8.</b> Representação do funcionamento do algoritmo SPA: projeção do plano ortogonal a $X_3$ (a) e visão espacial dos vetores no plano horizontal (b).....                                   | 31 |
| <b>Figura 2.1.</b> Viscosímetro Viscotest 6 L.....  | 36 |
| <b>Figura 2.2.</b> Refratômetro de Abbé.....  | 37 |
| <b>Figura 2.3.</b> Sistema para registro dos espectros NIR dos óleos vegetais. A: espectrômetro NIR; B: amostra de óleo vegetal; C: cubeta de fluxo de quartzo; D: bomba peristáltica; E: descarte..... | 38 |
| <b>Figura 2.4.</b> Cubeta de fluxo de quartzo de 1 mm de caminho óptico.....  | 38 |
| <b>Figura 3.1.</b> Espectros NIR brutos das 70 amostras de óleos vegetais (6184 a 3068 $\text{cm}^{-1}$ ).....  | 40 |
| <b>Figura 3.2.</b> Espectros derivativos das 70 amostras de óleos vegetais analisados.....  | 42 |
| <b>Figura 3.3.</b> Intervalos de variáveis selecionados pelo <i>i</i> PLS para acidez..   | 43 |
| <b>Figura 3.4.</b> Intervalos de variáveis selecionados pelo <i>i</i> PLS para índice de refração.....  | 44 |
| <b>Figura 3.5.</b> Intervalos de variáveis selecionados pelo <i>i</i> PLS para viscosidade.....   | 44 |
| <b>Figura 3.6.</b> Variáveis escolhidas pelo SPA-MLR para acidez no intervalo do <i>i</i> PLS.....  | 45 |

---

|  |    |
|--|----|
| <b>Figura 3.7.</b> Variáveis escolhidas pelo SPA-MLR para índice de refração no intervalo do <i>i</i> PLS.....   | 46 |
| <b>Figura 3.8.</b> Variáveis escolhidas pelo SPA-MLR para viscosidade no intervalo do <i>i</i> PLS.....          | 46 |
| <b>Figura 3.9.</b> Valor previsto pelo SPA-MLR <i>versus</i> o valor de referência para acidez.....              | 48 |
| <b>Figura 3.10.</b> Valor previsto pelo SPA-MLR <i>versus</i> o valor de referência para índice de refração..... | 48 |
| <b>Figura 3.11.</b> Valor previsto pelo SPA-MLR <i>versus</i> o valor de referência para viscosidade.....        | 49 |

|   |    |
|---|----|
| <b>Tabela 1.1.</b> Características físico-químicas do óleo de soja.....                           | 9  |
| <b>Tabela 1.2.</b> Composição de ácidos graxos do óleo de soja.....                               | 10 |
| <b>Tabela 1.3.</b> Características físico-químicas do óleo de milho.....                          | 11 |
| <b>Tabela 1.4.</b> Composição de ácidos graxos do óleo de milho.....                              | 11 |
| <b>Tabela 1.5.</b> Características físico-químicas do óleo de girassol.....                       | 12 |
| <b>Tabela 1.6.</b> Composição de ácidos graxos do óleo de girassol.....                           | 13 |
| <b>Tabela 1.7.</b> Características físico-químicas do óleo de canola.....                         | 14 |
| <b>Tabela 1.8.</b> Composição de ácidos graxos do óleo de canola.....                             | 15 |
| <b>Tabela 1.9.</b> Regiões espectrais do infravermelho.....                                       | 21 |
| <b>Tabela 3.1.</b> Valores de RMSEP e correlação (r) dos modelos PLS, <i>i</i> PLS e SPA-MLR..... | 47 |

---

|                   |  |
|-------------------|--|
| $\nu$             | Freqüência Vibracional Clássica                          |
| $h$               | Constante de Plank                                       |
| $y$               | Fator de Anarmonicidade                                  |
| $n$               | Número Quântico Vibracional                              |
| $y^{verd}$        | Valores Verdadeiros do parâmetro Y avaliado              |
| $\hat{y}_{cal,n}$ | Valores de Predição para n Amostras de calibração        |
| $y_{cal,n}$       | Valores de Referência para n Amostras de calibração      |
| $y^{prev}$        | Valores previstos do parâmetro Y avaliado                |
| AGL               | Ácidos Graxos Livres                                     |
| ANVISA            | Agência Nacional de Vigilância Sanitária                 |
| AOCS              | American Oil Chemists' Society                           |
| ATR               | Attenuated Total Reflectance                             |
| DA                | Discriminant Analysis                                    |
| $E_{vib}$         | Energia vibracional                                      |
| FIR               | Far Infrared   |
| GA                | Genetic Algorithm  |
| HCA               | Hierarchical Cluster Analysis                            |
| IBGE              | Instituto Brasileiro de Geografia e Estatística          |
| $i$ PLS           | Interval Partial Least Square regression                 |
| LDA               | Linear Discriminant Analysis                             |
| MIR               | Medium Infrared  |
| MSC               | Multiplicative Scatter Correction                        |
| NIR               | Near Infrared  |
| PCA               | Principal Component Analysis                             |
| PCR               | Principal Component Regression                           |
| PC                | Principal Component                                      |
| PLS               | Partial Least Square                                     |
| $r$               | Valores de Correlação                                    |
| RMSECV            | Root Mean Square Error of Prediction of Cross-Validation |
| RMSEP             | Root Mean Square Error of Prediction                     |

---

|       |  |
|-------|--|
| SG    | Savitzky-Golay                                       |
| SIMCA | Soft Independent Modeling of Class Analogy           |
| SPA   | Successive Projection Algorithm                      |
| SPXY  | Sample Set Partitioning Based on Joint x-y Distances |
| VL    | Variável Latente                                     |

---

Óleos vegetais são muito importantes para a saúde porque desempenham funções fisiológicas que são essenciais para os seres vivos. A qualidade dos óleos depende de muitos fatores, desde a escolha das sementes até a etapa de produção. Desta forma, muitos parâmetros físicos e químicos são monitorados e os métodos da AOCS (American Oil Chemists' Society) são comumente utilizados para esse propósito devido à simplicidade de operação. Entretanto, estes métodos são lentos, trabalhosos e, em alguns casos, por exemplo, em medidas de acidez, carecem de precisão. A espectrometria no infravermelho próximo (*Near Infrared* - NIR) é uma técnica que tem superado tais inconvenientes fornecendo procedimentos analíticos rápidos, diretos, precisos e não-invasivos. Desta forma, é proposta deste trabalho desenvolver uma nova metodologia para a determinação simultânea de acidez, índice de refração e viscosidade em quatro tipos de óleos vegetais (soja, milho, girassol e canola). A metodologia é baseada na modelagem simples por regressão linear múltipla dos parâmetros estudados. As variáveis são selecionadas pelo algoritmo das projeções sucessivas empregado na região espectral selecionada previamente pelo algoritmo dos mínimos quadrados parciais por intervalos. Os resultados mostraram que os modelos obtidos apresentaram baixos valores de erros de predição (0,009 % m/m; 0,0003 e 0,8 mPa.s) e bons valores de correlação (0,94, 0,98 e 0,96) para acidez, índice de refração e viscosidade, respectivamente. Com efeito, uma calibração simples pode ser realizada com sucesso para cada parâmetro, sem a necessidade de desenvolver um modelo separado para cada tipo de óleo vegetal.

**Palavras-chave:** óleos vegetais; espectrometria NIR; calibração multivariada; seleção de variáveis.

---

Vegetable oils are very important for the human health because they fulfill physiological functions that are essential for the living beings. Oils quality depends on a lot of factors, from the seeds choice until the production step. Therefore, many physic-chemical parameters are monitored and AOCS methods (American Oil Chemists' Society) are usually applied for this purpose due to their operation simplicity. However, the AOCS methods are time consuming, laborious and, for some cases, for example, in acidity measurements, lack precision. Near infrared (NIR) spectrometry is a technique that had surpassed such drawbacks yielding fast, direct, precise, reproducible and non-invasive analytical procedures. Therefore, it is a purpose of this work to develop a new methodology for the simultaneous determination of acidity, refraction index and viscosity in four types of vegetable oil (soya, corn, sunflower and canola). The methodology is based on the simple Multiple Linear Regression modeling of the studied parameters. The variables selected by the Successive Projection Algorithm applied on the spectral region chosen beforehand by the Interval Partial Least Square. The results showed that the obtained modeling presented low error of prediction (0,009 % w/w; 0,0003 and 0,8 mPa.s) and good correlations values (0,94, 0,98 and 0,96) for acidity, refraction index and viscosity, respectively. In fact, a simple calibration can be carried out successfully for each parameter without needing to develop independent models for each type of vegetable oil.

**Keywords:** Edible oils; NIR spectrometry; multivariate calibration variable selection.

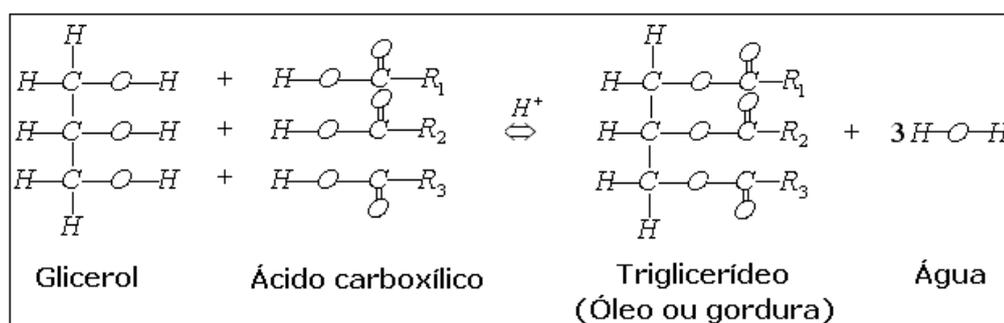
# **CAPÍTULO I**

## **INTRODUÇÃO**

## 1.1 - Óleos Vegetais

As diversidades sociais, econômicas e ambientais geram distintas motivações regionais para a produção e consumo de óleos vegetais (brutos e refinados) devido à grande quantidade de espécies oleaginosas das quais se podem extrair óleos e gorduras através de processos adequados.

Entende-se por óleos vegetais, um conjunto de compostos que compreendem o glicerol, ácidos graxos (principais componentes) e compostos resultantes da condensação destes, os ésteres (**Figura 1.1**). Estes ésteres, denominados triglicerídeos ou triacilgliceróis<sup>[1]</sup>, possuem cadeias de átomos de carbono com diferentes graus de insaturação. Além de possuírem ácidos graxos e glicerídeos em sua composição, os óleos vegetais também são constituídos de algumas substâncias como esteróis, tocoferóis, fenóis, flavonóides, compostos voláteis, vitaminas, pigmentos, etc<sup>[2]</sup>. São hidrofóbicos, ou seja, insolúveis em água, pois suas moléculas apolares são incapazes de formar ligações intermoleculares fortes com as moléculas de água. Possuem estabilidade oxidativa em diversas aplicações, como na fritura, por exemplo, pois contêm altos níveis de antioxidantes naturais (como os tocoferóis, constituintes da vitamina E), tornando-os mais resistentes a rancidez<sup>[1]</sup>.



**Figura 1.1.** Esquema da reação de formação do óleo vegetal.

Os lipídeos e derivados de vegetais possuem grande importância nutricional devido às vitaminas e ácidos graxos dos respectivos óleos e gorduras. A maior parte desses ácidos são monocarboxílicos com número

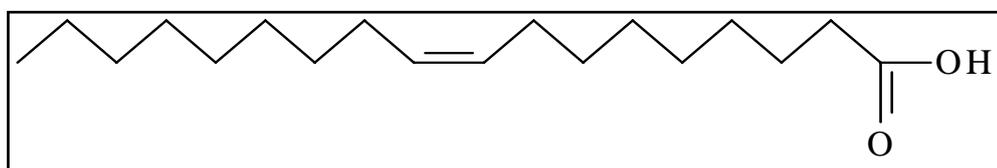
par de átomos de carbono dispostos numa cadeia linear provenientes da bioprodução da unidade acetato ( $\text{CH}_3\text{COSCoA}$  = acetil coenzima A combinada com  $\text{CO}_2$  no malonil coenzima A). Diferem um do outro por diferentes graus de insaturação, os quais caracterizam os ácidos como mono, di, tri e poliinsaturados. No grupo dos ácidos graxos saturados encontram-se os ácidos esteárico (o mais importante), butírico, cáprico, láurico, mirístico, palmítico, araquídico e behênico. Entre os insaturados, que ocorrem com mais freqüência na natureza, encontra-se o grupo do oléico, linoléico,  $\alpha$ -linolênico, que são os predominantes, seguidos dos ácidos araquidônico, elaídico e erúico<sup>[1,3]</sup>.

## 1.2 - A Importância dos Óleos Vegetais para a Saúde Humana

O consumo de gorduras é muito importante para a saúde, uma vez que exercem um efeito protetor de doenças em nosso organismo. No entanto, só são indicadas aquelas que não propiciam o surgimento de doenças coronárias e que não comprometem as vias circulatórias. As gorduras saturadas, por exemplo, que estão presentes no coco, na manteiga e em óleos industrialmente endurecidos, são gorduras que comprometem a saúde, pois aumentam o nível de colesterol do sangue e também promovem a arteriosclerose dos vasos sanguíneos. Todavia, as gorduras insaturadas exercem uma influência benéfica nas funções vasculares e ajudam a reduzir o nível de colesterol. São encontradas nos óleos comestíveis, principalmente aqueles que possuem grandes quantidades de ácidos graxos mono e poliinsaturados, como os de soja, arroz, girassol, milho e muitos outros, assim como em algumas margarinas especiais. Eles podem ser utilizados como temperos em alimentos, mas em frituras devem ser evitados devido ao excesso de calorias<sup>[1]</sup>.

Os ácidos graxos  $\alpha$ -linolênico e linoléico pertencentes às famílias ômega 3 e ômega 6 ( $\omega$ -3 e  $\omega$ -6), respectivamente, fazem parte de um grupo especial de ácidos graxos poliinsaturados, denominados de

essenciais. A ingestão desses ácidos na alimentação humana é muito importante, pois se sabe que eles ajudam a reduzir os níveis de colesterol do sangue e desempenham funções fisiológicas imprescindíveis nos organismos vivos, sendo, portanto, essenciais à vida. Já que não podem ser sintetizados pelas células, estes ácidos devem ser administrados pelos alimentos que os contêm. O ácido oléico (**Figura 1.2**) é um ácido graxo que pode ser produzido em pequenas quantidades no organismo e funciona como precursor dos ácidos graxos essenciais citados acima. É um dos ácidos graxos mais distribuídos na natureza, pois constitui mais da metade do total de muitos óleos e gorduras comestíveis. O óleo de oliva, por exemplo, pode conter até 80% de ácido oléico, enquanto que nas gorduras animais esses níveis chegam a pouco mais de 40%. Estes ácidos ajudam a diminuir o risco de ataques cardíacos e também a formação de arteriosclerose (espessamento e endurecimento da parede arterial)<sup>[4]</sup>.

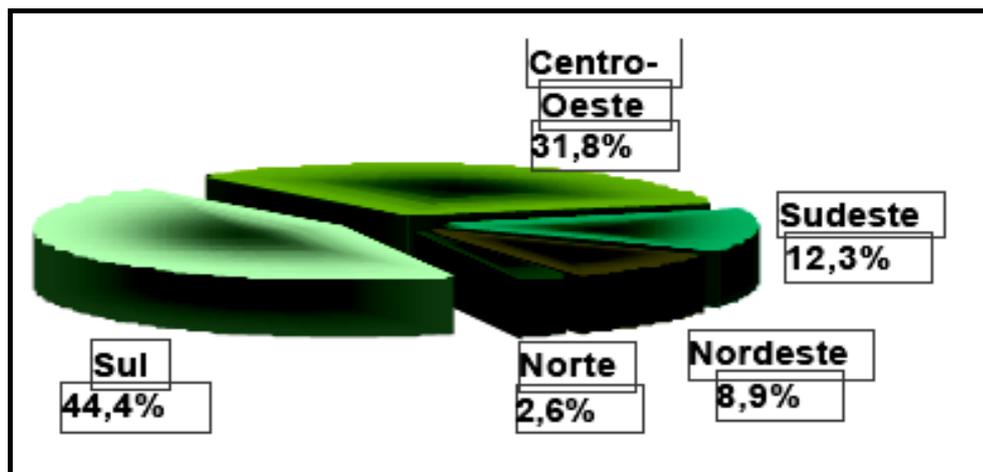


**Figura 1.2.** Estrutura do ácido oléico.

### 1.3 - Produção de Óleos Vegetais

No Brasil, a maior produção de oleaginosas encontra-se na região Sul, cuja produção em 2007 foi de 58,1 milhões de toneladas. Em seguida, encontram-se as regiões Centro-Oeste, Sudeste, Nordeste e Norte com 41,5; 16,0; 11,7 e 3,4 milhões de toneladas, respectivamente.

A **Figura 1.3** mostra a participação das grandes regiões na produção de oleaginosas cultivadas no país<sup>[5]</sup>.



**Figura 1.3.** Participação das grandes regiões na produção de oleaginosas(Fonte:IBGE)<sup>[5]</sup>

No Brasil, a maior parte da produção de óleos vegetais está voltada principalmente para o setor de alimentos, ou seja, para o consumo humano (óleo de cozinha), no entanto, eles podem ter inúmeras outras aplicações, tais como: pintura, lubrificante, cosméticos, farmacêutica, iluminação, combustível (Biodiesel ou Puro) e para usos industriais.

Segundo o Ministério da Ciência e Tecnologia<sup>[6]</sup>, com o incentivo do Governo Federal, os óleos vegetais estão sendo aproveitados como fonte energética destinada a diversos segmentos, abrindo oportunidades nos setores social, econômico e, sobretudo, ambiental. A produção está sendo destinada para a estruturação do programa de produção e uso do biodiesel em muitos países, principalmente no Brasil, já que diversas espécies oleaginosas, tais como a soja, a mamona, o girassol, etc são cultivadas e posteriormente utilizadas como matéria-prima na produção deste biocombustível.

No que se refere ao setor de alimentos, a qualidade dos óleos depende de diversos fatores associados a todos os estágios de produção, como a escolha do material bruto, fases de processamento, refinamento, armazenamento, etc<sup>[1]</sup>. Portanto, torna-se indispensável um rigoroso controle para suprir a demanda dos consumidores por produtos de alta qualidade. Para esta finalidade, muitos parâmetros físicos e/ou químicos são monitorados como acidez, densidade, cor, índice de refração,

viscosidade, umidade, volatilidade, valores de peróxido, índice de iodo, entre outros<sup>[7]</sup>.

### 1.3.1 - Processo de Refinação dos Óleos Vegetais

Para se obter óleos de alta qualidade é necessário que as condições de cultivo, preparo e armazenamento das sementes sejam as mais adequadas possíveis, pois além de obter produtos de qualidade, o rendimento também é maior. Caso contrário, pode ocorrer auto-aquecimento das sementes, aumentando assim, a acidez dos óleos produzidos bem como o escurecimento, as modificações organolépticas e as alterações estruturais dos mesmos.

Após os cuidados com o armazenamento e preparação da matéria-prima, os óleos brutos, que contém 95% de triglicerídeos e 5% de impurezas, são extraídos (mecanicamente ou com solventes) das sementes tratadas e aptos a passarem por uma tecnologia de processamentos que visa torná-los comestíveis. Essa tecnologia é conhecida como refinação e tem como objetivo melhorar a aparência, o odor e o sabor dos óleos brutos produzidos, removendo alguns componentes indesejáveis<sup>[1]</sup>.

As principais etapas do processo de refinação dos óleos vegetais são:

- ✚ **Degomagem:** Consiste em remover substâncias que causam escurecimento do óleo na etapa de desodorização, tais como as gomas (fosfatídeos hidratáveis), ceras, corantes (clorofila, carotenóides) e substâncias coloidais presentes nos óleos brutos;
- ✚ **Neutralização:** Tem a finalidade de remover os ácidos graxos livres (com hidróxido de sódio (NaOH)) e outros componentes indesejáveis como produtos de decomposição de glicerídeos, proteínas, ácidos oxidados, etc. A neutralização ocorre na interface do óleo e da solução alcalina e como essas fases não são intersolúveis, esse processo exige uma dispersão de solução alcalina no óleo;

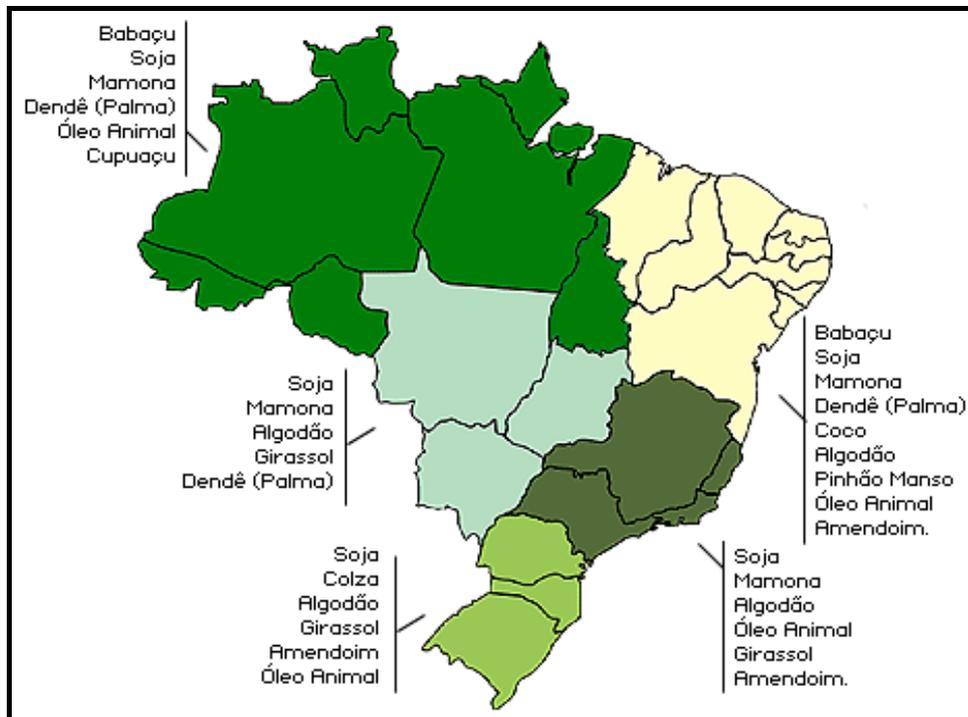
- ✚ **Branqueamento:** Apesar do processo de degomagem remover certa quantidade de corantes presentes no óleo e o de neutralização promover o seu branqueamento, eles não deixam o óleo completamente límpido, quase incolor, como exigem os consumidores. Por isso, as indústrias utilizam terras clarificantes (naturais ou ativadas) e algumas vezes misturadas com carvão ativado para adsorver os corantes presentes nos óleos;
- ✚ **Desodorização:** Objetiva remover os odores e sabores indesejáveis causados pelos peróxidos e ácidos graxos livres, bem como alguns compostos que se desenvolveram durante armazenagem e processamento das sementes e óleos, tais como: aldeídos, cetonas, ácidos graxos oxidados, produtos de decomposição de proteínas, carotenóides, esteróis, fosfatídeos, etc<sup>[1]</sup>.

#### 1.4 - Principais Óleos Vegetais Comercializados

De acordo com a Agência Nacional de Vigilância Sanitária (ANVISA)<sup>[8]</sup>, os principais óleos e gorduras vegetais comercializados são:

- ✚ **Azeite de oliva;**
- ✚ **Óleo de algodão;**
- ✚ **Óleo de amendoim;**
- ✚ **Óleo de arroz;**
- ✚ **Óleo de canola;**
- ✚ **Óleo de gergelim;**
- ✚ **Óleo de girassol;**
- ✚ **Óleo de milho;**
- ✚ **Óleo de soja;**
- ✚ **Óleo de uva;**
- ✚ **Óleo e gordura vegetal modificados;**
- ✚ **Óleo misto ou composto;**
- ✚ **Óleo ou gordura de coco de babaçu;**
- ✚ **Óleo ou gordura de coco;**
- ✚ **Óleo ou gordura de palma;**
- ✚ **Óleo ou gordura de palmiste;**
- ✚ **Óleo vegetal saborizado e azeite saborizado.**

No Brasil, os diversos tipos de oleaginosas produtoras de óleo estão distribuídos conforme está mostrado na **Figura 1.4**.



**Figura 1.4.** Informações de oleaginosas por região (Fonte: PETROBRÁS)<sup>[9]</sup>.

### 1.4.1 - Óleo de Soja

Segundo o IBGE<sup>[5]</sup>, a soja (*glycine max* L.) é a oleaginosa mais cultivada no Brasil, pois pode ser plantada em qualquer região. O Brasil ocupa a 2ª posição do *ranking* da produção mundial, perdendo apenas para os Estados Unidos<sup>[10]</sup>. Em 2007, a cultura dessa semente chegou a ocupar cerca de 20,7 milhões de hectares plantados no país.

Em média, a soja possui 40% de proteínas, 20% de lipídios (óleo), 5% de minerais e 34% de carboidratos (açúcares como glicose, frutose e sacarose, fibras e os oligossacarídeos como rafinose e estaquiose)<sup>[11]</sup>. Devido a essas características, a produção do óleo de soja predomina no mercado brasileiro tornando-o o mais consumido não só no Brasil, mas em todo o mundo.

No Brasil, além de possuir preços mais baixos para consumo humano, o óleo de soja vem se destacando, em curto prazo, na produção

nacional de biodiesel, abrindo também as portas para os demais óleos vegetais, uma vez que eles passaram a representar um dos principais insumos para a produção deste biocombustível.

O óleo de soja pode ser produzido como bruto, refinado comestível, refinado industrial e lecitina. Dentre os óleos, é o que apresenta o maior teor de gorduras saturadas, possui bom percentual de gorduras monoinsaturadas e poliinsaturadas, sendo por isso, indicado para frituras.

Uma das principais características do óleo de soja é que o mesmo possui originalmente um odor pronunciado. Isto exige que ele passe sempre por um processo de desodorização, última etapa do processo de refinação. Quando refinado, ele apresenta um aspecto límpido e uma cor levemente amarelada, bem como odor e sabor característicos<sup>[12]</sup>.

As **Tabelas 1.1** e **1.2** abaixo apresentam as características físico-químicas e a composição de ácidos graxos, respectivamente, presentes no óleo de soja<sup>[8]</sup>.

**Tabela 1.1.** Características físico-químicas do óleo de soja (Fonte: ANVISA)<sup>[8]</sup>.

| <b>Propriedades Físico-Químicas</b> | <b>Unidades</b>         | <b>Faixa de aceitação para consumo</b>                 |
|-------------------------------------|-------------------------|--|
| Densidade Relativa                  | -                       | 0,919 - 0,925 (20°C/20°C)<br>0,916 - 0,922 (25°C/25°C) |
| Índice de Refração                  | -                       | 1,466 - 1,470  |
| Índice de Iodo                      | g I <sub>2</sub> / 100g | 120 - 143  |
| Índice de Saponificação             | -                       | 189 - 195  |
| Matéria Insaponificável             | g/100 g                 | < 1,5  |
| Acidez, óleo refinado               | g ácido oléico/100g     | < 0,3  |
| Acidez, óleo bruto                  | g ácido oléico/100g     | < 2,0  |
| Acidez, óleo semi-refinado          | g ácido oléico/100g     | < 0,5  |
| Acidez, óleo degomado               | g ácido oléico/100g     | < 1,0  |
| Índice de Peróxido                  | meq/kg                  | < 10,0   |
| Fósforo (óleo degomado)             | g/100g                  | < 0,02   |

**Tabela 1.2.** Composição de ácidos graxos do óleo de soja (Fonte: ANVISA)<sup>[8]</sup>.

| Ácidos graxos                       | Estrutura | (g / 100 g) |
|-------------------------------------|-----------|-------------|
| -                                   | C<14      | < 0,1       |
| Mirístico                           | C 14:0 *  | < 0,5       |
| Palmítico                           | C 16:0    | 7,0 - 14,0  |
| Palmitoléico                        | C 16:1    | < 0,5       |
| Esteárico                           | C 18:0    | 1,4 - 5,5   |
| Oléico ( $\omega$ -9)               | C 18:1    | 19,0 - 30,0 |
| Linoléico ( $\omega$ -6)            | C 18:2    | 44,0 - 62,0 |
| $\alpha$ -Linolênico ( $\omega$ -3) | C 18:3    | 4,0 - 11,0  |
| Araquídico                          | C 20:0    | < 1,0       |
| Eicosenóico                         | C 20:1    | < 1,0       |
| Behênico                            | C 22:0    | < 0,5       |

\*O primeiro valor corresponde ao número de átomos de carbono e o segundo, o número de insaturações.

### 1.4.2 - Óleo de Milho

O milho (*zea mays*) é o segundo produto mais plantado no Brasil. É muito saudável, pois é rico em vitamina E, amido e proteína e contém ácidos graxos essenciais. O Brasil ocupa a terceira posição no *ranking* mundial de produção deste cereal, ficando atrás apenas dos Estados Unidos e China. Os maiores produtores são os estados do Paraná e de São Paulo<sup>[10,13]</sup>.

O óleo de milho reduz o nível de colesterol e ajuda a regular a pressão arterial. O ácido graxo mais abundante no óleo de milho é o linoléico ( $\omega$ -6), que pode chegar a 50% ou mais do total. O ácido  $\alpha$ -linolênico triinsaturado é altamente sujeito à oxidação e está presente em pequena percentagem neste óleo, se comparado com o de soja<sup>[14,15]</sup>. É obtido por processo de extração e refino e possui um aspecto límpido, de cor amarelo-claro e com odor e sabor característicos.

Nas **Tabelas 1.3** e **1.4** são apresentadas as características físico-químicas e a composição de ácidos graxos presentes no óleo de milho<sup>[8]</sup>.

**Tabela 1.3.** Características físico-químicas do óleo de milho (Fonte: ANVISA)<sup>[8]</sup>.

| <b>Propriedades Físico-Químicas</b> | <b>Unidades</b>         | <b>Faixa de aceitação para consumo</b>                 |
|-------------------------------------|-------------------------|--|
| Densidade Relativa                  | -                       | 0,917 - 0,925 (20°C/20°C)<br>0,914 - 0,922 (25°C/25°C) |
| Índice de Refração                  | -                       | 1,465 - 1,468  |
| Índice de Iodo                      | g I <sub>2</sub> / 100g | 103 - 128  |
| Índice de Saponificação             | -                       | 187 - 195  |
| Matéria Insaponificável             | g/100g                  | < 2,8  |
| Acidez, óleo refinado               | g ácido oléico/100g     | < 0,3  |
| Acidez, óleo semi- refinado         | g ácido oléico/100g     | < 0,5  |
| Acidez, óleo bruto                  | g ácido oléico/100g     | < 6,0  |
| Índice de Peróxido                  | meq/kg                  | < 10,0   |

**Tabela 1.4.** Composição de ácidos graxos do óleo de milho (Fonte: ANVISA)<sup>[8]</sup>.

| <b>Ácidos graxos</b>                | <b>Estrutura</b> | <b>(g / 100 g)</b> |
|-------------------------------------|------------------|--------------------|
| -                                   | C<14             | < 0,3              |
| Mirístico                           | C 14:0           | < 0,1              |
| Palmítico                           | C 16:0           | 9,0 - 14,0         |
| Palmitoléico                        | C 16:1           | < 0,5              |
| Estearico                           | C 18:0           | 0,5 - 4,0          |
| Oléico ( $\omega$ -9)               | C 18:1           | 24,0 - 42,0        |
| Linoléico ( $\omega$ -6)            | C 18:2           | 34,0 - 62,0        |
| $\alpha$ -Linolênico ( $\omega$ -3) | C 18:3           | < 2,0              |
| Araquídico                          | C 20:0           | < 1,0              |
| Eicosenóico                         | C 20:1           | < 0,5              |
| Behênico                            | C 22:0           | < 0,5              |
| Lignocérico                         | C 24:0           | < 0,5              |

### 1.4.3 - Óleo de Girassol

Os maiores produtores do grão girassol (*helianthus annuus*) são Argentina, Rússia, Ucrânia, USA, China, França e Índia. A produção brasileira ainda é pouco expressiva, mas atualmente o produto vem se consolidando e conquistando o seu espaço no mercado, de modo que esta oleaginosa já ocupa o quarto lugar em volume de produção, perdendo apenas para a soja, palma e colza.

Das sementes do girassol se extrai um óleo de excelente qualidade nutricional e com um ótimo rendimento em torno de 48 a 52%. Este óleo possui um alto teor de ácido linoléico e tocoferóis, que são responsáveis pelo tratamento da esclerose múltipla e pela prevenção de doenças cardiovasculares produzidas pelo excesso de colesterol nos vasos sanguíneos. Além disso, é altamente digestivo e possui o maior teor de gorduras poliinsaturadas. Apresenta aspecto límpido, cor amarelo-dourado claro e odor e sabor suave característicos<sup>[16,17]</sup>.

As **Tabelas 1.5** e **1.6** apresentam as características físico-químicas e a composição de ácidos graxos presentes no óleo de girassol<sup>[8]</sup>.

**Tabela 1.5.** Características físico-químicas do óleo de girassol (Fonte: ANVISA)<sup>[8]</sup>.

| Propriedades Físico-Químicas | Unidades                | Faixa de aceitação para consumo                        |
|------------------------------|-------------------------|--|
| Densidade Relativa           | -                       | 0,918 - 0,923 (20°C/20°C)<br>0,915 - 0,920 (25°C/25°C) |
| Índice de Refração           | -                       | 1,467 - 1,469  |
| Índice de Iodo               | g I <sub>2</sub> / 100g | 110 - 143  |
| Índice de Saponificação      | mg KOH/g                | 188 - 194  |
| Matéria Insaponificável      | g/100g                  | < 1,5  |
| Acidez, óleo refinado        | g ácido oléico/100g     | < 0,3  |
| Acidez, óleo semi- refinado  | g ácido oléico/100g     | < 0,5  |
| Acidez, óleo bruto           | g ácido oléico/100g     | < 2,0  |
| Acidez, óleo virgem          | g ácido oléico/100g     | < 2,0  |
| Índice de Peróxido           | meq/kg                  | < 10,0   |

**Tabela 1.6.** Composição de ácidos graxos do óleo de girassol (Fonte: ANVISA)<sup>[8]</sup>.

| Ácidos graxos               | Estrutura | (g / 100 g) |
|-----------------------------|-----------|-------------|
| -                           | C<14      | < 0,4       |
| Mirístico                   | C 14:0    | < 0,5       |
| Palmítico                   | C 16:0    | 3,0 - 10,0  |
| Palmitoléico                | C 16:1    | < 1,0       |
| Estearico                   | C 18:0    | 1,0 -10,0   |
| Oléico ( $\omega$ -9)       | C 18:1    | 14,0 - 35,0 |
| Linoléico ( $\omega$ -6)    | C 18:2    | 55,0 - 75,0 |
| A-Linolênico ( $\omega$ -3) | C 18:3    | < 0,3       |
| Araquídico                  | C 20:0    | < 1,5       |
| Eicosenóico                 | C 20:1    | < 0,5       |
| Behênico                    | C 22:0    | < 1,0       |
| Lignocérico                 | C 24:0    | < 0,5       |
| Erúcido                     | C 22:1    | < 0,5       |
| Tetracosenóico              | C 24:1    | < 0,5       |

#### 1.4.4 - Óleo de Canola

O óleo de colza é considerado inviável para consumo humano uma vez que possui alto teor de ácido erúcido e glucolisanato, compostos tóxicos que são considerados venenosos. Por este motivo, laboratórios canadenses conseguiram, por meio de melhoramento genético da colza, obter uma planta transgênica com baixo teor de tais compostos. Esta planta (*brassica nabus e brassica campestri*) é, hoje em dia, mais conhecida com o nome de canola, do inglês *Canadian Oil Low Acid*, da qual se pôde extrair um óleo próprio para o consumo humano<sup>[11]</sup>.

Os maiores produtores da canola são China, Canadá, Índia, Alemanha, França, Austrália e Polônia<sup>[10]</sup>. No Brasil, as pesquisas e o cultivo da colza tiveram início em 1974 no Rio Grande do Sul (RS), em meados dos anos 80 no Paraná (PR) e em 2003 no estado de Goiás (GO).

Os grãos de canola produzidos no país possuem cerca de 24 a 27% de proteína e de 34 a 40% de óleo.

O óleo de canola é considerado de ótima qualidade por apresentar diversas vantagens, tais como:

- ✚ Possui menor índice de gorduras saturadas (6%, contra 15% do óleo de soja e 11% do óleo de milho);
- ✚ É rico em vitamina E, em ômega 3 e em gorduras monoinsaturadas.

As gorduras monoinsaturadas presentes no óleo de canola são gorduras redutoras de uma lipoproteína de baixa densidade, denominada LDL (*low density lipoprotein*). Esta lipoproteína é responsável pelo depósito do colesterol na parede das artérias (arteriosclerose) e é popularmente conhecida como "mau colesterol"<sup>[1,18,19]</sup>.

São mostrados nas **Tabelas 1.7** e **1.8** as características físico-químicas e a composição de ácidos graxos presentes no óleo de canola<sup>[8]</sup>.

**Tabela 1.7.** Características físico-químicas do óleo de canola (Fonte: ANVISA)<sup>[8]</sup>.

| Propriedades Físico-Químicas | Unidades                | Faixa de aceitação para consumo                        |
|------------------------------|-------------------------|--|
| Densidade Relativa           | -                       | 0,914 - 0,920 (20°C/20°C)<br>0,911 - 0,917 (25°C/25°C) |
| Índice de Refração           | -                       | 1,465 - 1,467  |
| Índice de Iodo               | g I <sub>2</sub> / 100g | 110 - 126  |
| Índice de Saponificação      | -                       | 182 - 193  |
| Matéria Insaponificável      | g/100g                  | < 2,0  |
| Acidez, óleo refinado        | g ácido oléico/100g     | < 0,3  |
| Acidez, óleo semi-refinado   | g ácido oléico/100g     | < 0,5  |
| Acidez, óleo degomado        | g ácido oléico/100g     | < 1,0  |
| Acidez, óleo bruto           | g ácido oléico/100g     | < 2,0  |
| Fósforo, óleo degomado       | g/100 g                 | < 0,02   |
| Índice de Peróxido           | meq/kg                  | < 10,0   |
| Brassicasterol               | g/100 g                 | > 5,0  |
| Ácido erúcido                | g/100 g                 | > 2,0  |

**Tabela 1.8.** Composição de ácidos graxos no óleo de canola (Fonte: ANVISA)<sup>[8]</sup>.

| Ácidos graxos                       | Estrutura | (g / 100 g) |
|-------------------------------------|-----------|-------------|
| Mirístico                           | C 14:0    | < 0,2       |
| Palmítico                           | C 16:0    | 2,5 - 6,5   |
| Palmitoléico                        | C 16:1    | < 0,6       |
| Esteárico                           | C 18:0    | 0,8 - 3,0   |
| Oléico ( $\omega$ -9)               | C 18:1    | 53,0 - 70,0 |
| Linoléico ( $\omega$ -6)            | C 18:2    | 15,0 - 30,0 |
| $\alpha$ -Linolênico ( $\omega$ -3) | C 18:3    | 5,0 - 13,0  |
| Araquídico                          | C 20:0    | 0,1 - 1,2   |
| Eicosenóico                         | C 20:1    | 0,1 - 4,3   |
| Behênico                            | C 22:0    | < 0,6       |
| Lignocérico                         | C 24:0    | < 0,2       |
| Erúcico                             | C 22:1    | < 2,0       |
| Tetracosenóico                      | C 24:1    | < 0,2       |

### 1.5 - Análise de Óleos Vegetais

Algumas características são indicativas da qualidade de um produto. No que se refere aos óleos vegetais, parâmetros físicos e químicos são analisados e utilizados para garantir a qualidade dos mesmos. Dentre os mais comumente analisados, encontram-se densidade, índice de iodo, índice de peróxido, índice de saponificação, medição de cor<sup>[20]</sup> e os que serão abordados neste trabalho, tais como: índice de acidez, índice de refração e medidas de viscosidade. Estes parâmetros de qualidade podem ser determinados por métodos de referência como aqueles estabelecidos pela AOCS (*American Oil Chemists' Society*)<sup>[21]</sup>, os quais têm sido bastante empregados por muitos laboratórios e indústrias.

A AOCS tem também proposto um conjunto de normas com limites específicos dos parâmetros de qualidades de óleos vegetais, de modo a garantir a conformidade dos mesmos dentro de um certo padrão de qualidade, levando em conta se eles estarão aptos ou não para consumo.

O teor de acidez é uma das principais características que confere qualidade aos óleos vegetais. Vários fatores podem influenciar a acidez de um óleo vegetal, mas o principal é o tratamento dado ainda às sementes durante a colheita e armazenamento. Assim, a acidez está diretamente relacionada com a qualidade da matéria-prima, com o processamento e, principalmente, com as condições de conservação dos óleos vegetais. Segundo a AOCS<sup>[21]</sup>, a acidez em óleos vegetais pode ser determinada através da titulação ácido-base, utilizando hidróxido de sódio e fenolftaleína como titulante e indicador, respectivamente. A AOCS recomenda que o índice de acidez permitido seja no máximo 0,3 g/100 g.

O índice de refração e a viscosidade são parâmetros físicos não menos importantes que os químicos para garantir a qualidade dos óleos. O índice de refração é a relação existente entre a velocidade da luz no ar e no meio (substância em exame). Ele varia na razão inversa da temperatura e tende a aumentar com o grau de insaturação dos ácidos graxos constituintes dos triglicerídeos. A viscosidade aumenta com o comprimento das cadeias dos ácidos graxos constituintes dos triglicerídeos e diminui com o aumento das insaturações<sup>[22]</sup>.

### **1.5.1 - Métodos de Análise de Óleos Vegetais**

Muitas metodologias analíticas têm sido propostas para a determinação, caracterização e autenticação de óleos comestíveis, dentre elas, a Cromatografia Gasosa<sup>[23]</sup>, Espectrometria de Ressonância Magnética Nuclear de  $^1\text{H}$  e  $^{13}\text{C}$ <sup>[24]</sup>, Espectrometria de Massa<sup>[25]</sup> e Espectrometria no Infravermelho Próximo (*Near Infrared* - NIR)<sup>[20]</sup>. No que se refere à determinação dos parâmetros de qualidade dos óleos, os métodos da AOCS<sup>[21]</sup> são muito utilizados devido à simplicidade de operação. Entretanto, estes métodos são lentos, trabalhosos e, em alguns casos, por exemplo, em medidas de acidez, carecem de precisão. Na determinação deste parâmetro, grandes quantidades de amostras e reagentes são consumidas durante a operação e há muita dificuldade na

visualização do ponto de viragem do indicador<sup>[20]</sup>. Um método analítico que tem sido capaz de superar tais inconvenientes é a Espectrometria NIR. Essa técnica fornece procedimentos analíticos rápidos, diretos (sem pré-tratamento), reproduzíveis, não-invasivos e, além disso, o custo por análise é relativamente baixo.

A literatura dispõe de alguns estudos<sup>[26-36]</sup> que objetivam classificar e/ou autenticar óleos vegetais pela análise de alguns parâmetros como acidez, índice de refração, viscosidade, índice de iodo, etc ou determinaram alguns constituintes presentes nos óleos, tais como: triglicerídeos, esteróis, tocoferóis e ácidos graxos. Devido à complexidade das matrizes dos óleos vegetais, ferramentas quimiométricas<sup>[37-39]</sup> têm sido muito utilizadas. Dentre elas, podemos citar: a análise hierárquica de agrupamentos (HCA), a análise de componentes principais (PCA), a modelagem independente flexível por analogia de classe (SIMCA), a análise discriminante linear (LDA) e os mínimos quadrados parciais (PLS).

Dahlberg *et al*<sup>[26]</sup> determinaram índice de refração e viscosidade em 27 marcas de diferentes tipos de óleos vegetais e margarinas, empregando a espectrometria MIR e calibração multivariada (PLS). Com os modelos PLS construídos foi possível prever os dois parâmetros físicos analisados com um erro padrão estimado de 2,2% e 2,8% para índice de refração e viscosidade, respectivamente.

Técnicas quimiométricas como PCA e LDA, combinadas à cromatografia gasosa, foram empregadas por Lee *et al*<sup>[27]</sup> para classificar oito tipos de óleos vegetais de acordo com a composição de seus ácidos graxos. A PCA selecionou essas amostras dividindo-as em grupos distintos e pôde distinguí-las claramente enquanto a LDA identificou amostras desconhecidas em um desses grupos. Os resultados cromatográficos obtidos mostraram adulteração em um dos tipos de óleo analisado.

Martín *et al*<sup>[28]</sup> empregaram LDA para classificar óleos vegetais usando um arranjo de sensores de gás denominado "nariz eletrônico" (do inglês: *Electronic Nose*). Os resultados mostraram que a combinação deste método com a LDA permitiu classificar corretamente todos os óleos

vegetais. Eles conseguiram resultados satisfatórios, apesar de concluírem que melhores resultados poderiam ser obtidos se usassem mais sensores, assim como outras técnicas de reconhecimento de padrões.

Barthus e Poppi<sup>[29]</sup> determinaram o índice de iodo em 24 amostras de óleos vegetais usando a espectrometria Raman e PLS e compararam seus resultados com os métodos de referência adotados pela AOCS. O modelo foi construído e 3 variáveis latentes foram suficientes para validá-lo empregando, como técnica de validação, a validação cruzada. Eles concluíram que o método proposto foi eficaz, visto que seus resultados foram concordantes com os do método de referência, fornecendo um coeficiente de correlação de 0,996.

Iñón *et al*<sup>[30]</sup> utilizaram a espectrometria MIR com transformada de Fourier e medidas de reflectância total atenuada (ATR) e PLS para a determinação da concentração de ácidos graxos livres (AGL) em óleos de oliva. Antes da construção dos modelos PLS, diferentes aspectos foram estudados, tais como: métodos para seleção do conjunto de calibração, faixa espectral a ser considerada e as técnicas de pré-processamentos a serem empregadas. Usando um conjunto de calibração com 16 amostras, as propriedades de 28 amostras foram preditas com uma precisão relativa de 0,017 % (m/m).

Christy *et al*<sup>[31]</sup> classificaram e quantificaram a adulteração em óleos de oliva por cinco tipos de óleos vegetais (soja, girassol, milho, noz e avelã) empregando a espectrometria NIR e técnicas quimiométricas (PLS e PCA). Os espectros produzidos foram pré-processados usando uma correção de espalhamento multiplicativo (MSC) e suavizados pelo método de Savitzky-Golay (SG). Os modelos PLS foram construídos para prever os óleos vegetais adulterantes presentes nas amostras de óleos de oliva. Os resultados mostraram que os modelos PLS construídos forneceram erros relativos percentuais de predição de  $\pm 0,57$ ,  $\pm 1,32$ ,  $\pm 0,96$ ,  $\pm 0,56$  e  $\pm 0,57$  % m/m para soja, girassol, milho, noz e avelã, respectivamente. Os modelos PCA, com três componentes principais, explicaram 99% da

variância dos dados, distribuindo as amostras adulteradas nos 5 grupos distintos esperados.

Um estudo para classificar 80 amostras de óleos de oliva de diferentes origens empregando PCA e SIMCA foi desenvolvido por Diaz *et al*.<sup>[32]</sup> Eles analisaram conteúdos de triglicerídeos e esteróis presentes nos óleos e também alguns parâmetros, tais como: acidez, índice de peróxido, cor, ácidos graxos e estabilidade. A análise de triglicerídeos foi realizada usando a cromatografia líquida e a de esteróis, a cromatografia gasosa. Os demais parâmetros foram analisados usando o método oficial da Comunidade Européia. Os resultados dos modelos PCA obtidos para os valores de triglicerídeos e esteróis mostraram que cinco componentes principais foram suficientes para explicar 97,3% da variância dos dados. O modelo SIMCA foi construído para classificar um dos tipos de óleo de oliva definido por MC pelos autores. O modelo classificou dois grupos: um que continha amostras de óleos MC e outro que não continha, denominando os óleos desse segundo grupo de NMC. Como parte do estudo, eles empregaram a PCA novamente para visualizar os dois grupos e puderam verificar que seis componentes principais foram suficientes para explicar 95,8% e 98,1% da variância dos dados para as classes MC e NMC, respectivamente. Eles concluíram que os resultados obtidos com a composição dos triglicerídeos combinados ao modelo SIMCA para classificar a classe de amostras MC foram melhores do que aqueles obtidos com a composição dos esteróis e dos demais parâmetros analisados. Assim, a análise de triglicerídeos foi determinante para classificar corretamente esse tipo de óleo.

Osawa *et al*.<sup>[33]</sup> fizeram um estudo comparativo do uso da titulação potenciométrica e com indicador na determinação de ácidos graxos livres (AGL), expressos em % de ácido oléico, em óleos vegetais refinados, brutos e degomados e em gordura extraída de rações para animais. Os autores verificaram que, em geral, os resultados obtidos com ambas as titulações se equiparam. Entretanto, foi observado que o coeficiente de variação obtido na titulação potenciométrica de óleos brutos (0,6 a 2,6%)

não foi concordante com aqueles obtidos pela titulação com o indicador (0,3 a 105,8%). Uma explicação para essa diferença está na dificuldade da visualização do ponto final da titulação com indicador devido a presença de pigmentos nestas amostras.

Christy *et al*<sup>[34]</sup> usaram a espectrometria NIR e PLS para determinação de ácidos graxos saturados e insaturados em óleos vegetais e compararam os resultados obtidos com os valores determinados empregando a cromatografia gasosa. O modelo PLS pôde prever o conteúdo dos ácidos graxos saturados e insaturados com uma margem de erro de predição de 2,5%. Eles também determinaram, especificamente, os teores do ácido oléico e do linoléico e obtiveram, através dos modelos PLS construídos, erros de predição de 3,5 e 4,1%, respectivamente.

Um estudo para distinguir óleos vegetais crus e refinados usando métodos cromatográficos e técnicas quimiométricas foi desenvolvido por Giacomelli *et al*<sup>[35]</sup>. O estudo se baseou na análise de componentes principais (PCA) relacionando a composição de ácidos graxos com níveis de tocoferóis e medidas de cor (pela análise de pigmentos). Os resultados mostraram que 3 componentes principais foram suficientes para distinguir os óleos crus dos óleos refinados, explicando mais de 80% do total da variância dos dados.

Um método para determinação de acidez e índice de peróxido em diferentes tipos de óleos vegetais usando a Espectroscopia NIR combinada a técnicas quimiométricas de classificação (HCA) e calibração multivariada (PLS) foi proposto por Armenta *et al*<sup>[36]</sup>. Um primeiro grupo de amostras de óleo foi usado para a construção dos modelos de calibração, validação e predição e algumas amostras desse grupo foram separadas para testar a robustez dos modelos em função da variabilidade. Um segundo grupo foi usado para aplicar a metodologia desenvolvida. O HCA distribuiu 11 conjuntos correspondentes aos dois grupos. Os resultados mostraram que as amostras do segundo grupo não foram claramente diferenciadas. Os modelos PLS construídos forneceram um erro médio quadrático de predição (RMSEP) de 0,034% (grupo I) e 0,037% (m/m) (grupo II) para a

acidez e 1,87 e 0,79 (meq.O<sub>2</sub>.kg<sup>-1</sup>) para o índice de peróxido, respectivamente.

Como se pode observar nos parágrafos acima, a espectrometria NIR combinada a técnicas quimiométricas é uma ferramenta muito valiosa para a determinação dos parâmetros de qualidade dos óleos vegetais. Na realidade, nas duas últimas décadas o interesse pela espectrometria NIR tem crescido consideravelmente por sua capacidade de obter inúmeras informações sobre uma espécie de interesse, sendo então, empregada em diversos tipos de análises para determinações qualitativas e/ou quantitativas. As grandes vantagens desta técnica, comparada às técnicas clássicas de referência, são que as análises fornecem resultados mais rápidos, não são destrutivas e invasivas e não geram subprodutos químicos tóxicos. Por isso, a sua aplicação tem sido indispensável nos laboratórios e nas indústrias, dentre elas as: de alimentos, químicas, farmacêuticas, de tintas, de lãs, petroquímicas, têxteis, de polímeros, de cosméticos, de fibras, de borrachas, etc. A sua aplicação tem também auxiliado a agricultura na análise da composição química dos solos, assim como também tem sido empregada em restaurações de imagens decorativas de pinturas históricas<sup>[38,40,41]</sup>.

## 1.6 - Espectroscopia no Infravermelho

A espectroscopia no infravermelho engloba os fenômenos associados à absorção e à emissão de energia radiante que ocorre em três regiões do espectro eletromagnético (**Tabela 1.9**): NIR, MIR e FIR.

**Tabela 1.9.** Regiões espectrais do Infravermelho (Fonte: SKOOG)<sup>[42]</sup>.

| Região            | Região em Comprimentos de Onda (μm) | Região em Números de Onda (cm <sup>-1</sup> ) | Região em Freqüência (Hz)                   |
|-------------------|-------------------------------------|---|---|
| <b>Próximo</b>    | 0,78 a 2,5                          | 12.800 a 4.000                                | 3,8x10 <sup>14</sup> a 1,2x10 <sup>14</sup> |
| <b>Médio</b>      | 2,5 a 50                            | 4.000 a 200                                   | 1,2x10 <sup>14</sup> a 6,0x10 <sup>12</sup> |
| <b>Distante</b>   | 50 a 1000                           | 200 a 10                                      | 6,0x10 <sup>12</sup> a 3,0x10 <sup>11</sup> |
| <b>Mais usada</b> | 2,5 a 15                            | 4.000 a 670                                   | 1,2x10 <sup>14</sup> a 2,0x10 <sup>13</sup> |

A condição para que haja absorção da radiação infravermelha é que uma molécula apresente uma variação no momento de dipolo como consequência do movimento vibracional ou rotacional. Isto ocorre, por exemplo, em moléculas heteronucleares, como HCl. Todavia, não ocorre em moléculas homonucleares, como H<sub>2</sub>, que não absorvem radiações no infravermelho<sup>[42]</sup>.

Quando uma molécula absorve radiação infravermelha, seus átomos vibram com maior amplitude. Esse comportamento é similar ao de um oscilador harmônico (**Figura 1.5**) que, pela mecânica quântica, apresenta os níveis de energia vibracional entre dois átomos de uma molécula quantizados, conforme a **Equação 1.1**.

$$E_{vib} = h\nu \left(n + \frac{1}{2}\right) \quad (1.1)$$

onde  $E_{vib}$  é a energia vibracional,  $h$  é a constante de Plank,  $\nu$  é a frequência vibracional clássica e  $n$  é o número quântico vibracional.

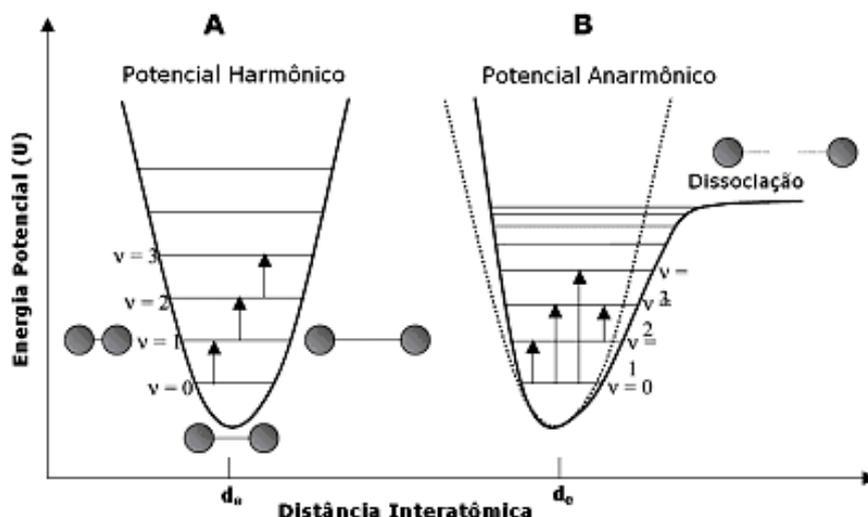
Na prática, os níveis de energia quantizados não são igualmente espaçados porque o comportamento das moléculas se desvia da idealidade, ou seja, o modelo do oscilador harmônico não pode explicar o comportamento de moléculas reais, pois ele não considera a interação coulômbica entre átomos ou o rompimento das ligações. Desta forma, o comportamento das moléculas reais assemelha-se mais ao modelo de um oscilador anarmônico, onde a diferença dos níveis de energia diminui com o aumento de  $n$ , como mostrado na **Equação 1.2**.

$$E_{vib} = h\nu [1 - (2n + \Delta n + 1)y] \quad (1.2)$$

onde  $y$  é o fator de anarmonicidade.

A anarmonicidade pode resultar em transições entre estados de energia vibracional, onde  $\Delta n = \pm 2, \pm 3, \dots$ . Essas transições produzem bandas de absorção conhecidas como sobretons e são, assim, muito pouco prováveis de ocorrer se comparadas às transições vibracionais

fundamentais. Portanto, os sobretons são caracterizados por transições vibracionais mais fracas<sup>[38,42]</sup>.



**Figura 1.5.** Diagrama de energia potencial para os osciladores harmônico (A) e anarmônico (B) (Fonte: PASQUINI, C)<sup>[38]</sup>.

### 1.6.1 - Espectroscopia no Infravermelho Próximo (NIR)

A absorção de radiação no infravermelho próximo (12.800 a 4.000  $\text{cm}^{-1}$ ) está associada a grupos funcionais que apresentam ligações muito polarizadas como C-H, N-H e O-H. Os espectros de absorção na região NIR são caracterizados pela presença de sobretons de transições vibracionais fundamentais e por sobretons de combinações dessas transições fundamentais que aparecem normalmente na região do MIR. Os espectros NIR não fornecem apenas informações químicas das amostras, mas também físicas, como: densidade, viscosidade, índice de refração, etc. As bandas de absorção na região NIR são de intensidades fracas (cerca de 10 a 1000 vezes menos intensas) e bem mais sobrepostas do que aquelas que ocorrem nas regiões MIR e FIR. Essa característica não constitui uma desvantagem da técnica NIR, pois além de permitir o seu uso em medidas de amostras que apresentam um alto teor de umidade, garante alta penetração da radiação em medidas de

refletância difusa de amostras sólidas. Além disso, quase todos os compostos orgânicos absorvem na região NIR<sup>[38,42-45]</sup>.

Pelo exposto, pode-se afirmar que a espectrometria NIR apresenta-se como uma ferramenta poderosa para determinações analíticas, mostrando-se capaz de realizar medidas com rapidez e precisão<sup>[38]</sup>. Entretanto, as bandas espectrais de baixa intensidade e muito sobrepostas, geradas por matrizes complexas, dificultam a obtenção de uma boa correlação entre grupos de átomos presentes numa molécula e a sua composição nas amostras. Por isso, as metodologias analíticas desenvolvidas usando a espectrometria NIR, geralmente, necessitam da utilização de métodos quimiométricos de análise para auxiliar na elucidação dos perfis espectrais.

## 1.7 - Quimiometria

A inserção de microcomputadores em laboratórios químicos, o desenvolvimento de técnicas instrumentais, a disponibilidade de softwares e a necessidade de ferramentas mais sofisticadas capazes de extrair maiores informações de conjuntos de dados muito complexos contribuíram para o surgimento da área da química conhecida como Quimiometria. Pode-se definir a Quimiometria como a ciência que utiliza métodos matemáticos e estatísticos para extrair de dados químicos complexos de natureza multivariada, o máximo de informações relevantes<sup>[46,47]</sup>.

Na Quimiometria existem várias técnicas extremamente úteis para os químicos modernos. Dentre elas, podemos citar: técnicas de pré-processamento de dados<sup>[48,49]</sup>, de planejamento e otimização experimental<sup>[50]</sup>, de reconhecimento de padrões e classificação<sup>[25,32]</sup>, de calibração multivariada<sup>[51-55]</sup>, de transferência de calibração<sup>[56,57]</sup>, etc.

Neste trabalho foram utilizadas algumas técnicas de pré-processamento de dados e de calibração multivariada que serão discutidas nas seções a seguir.

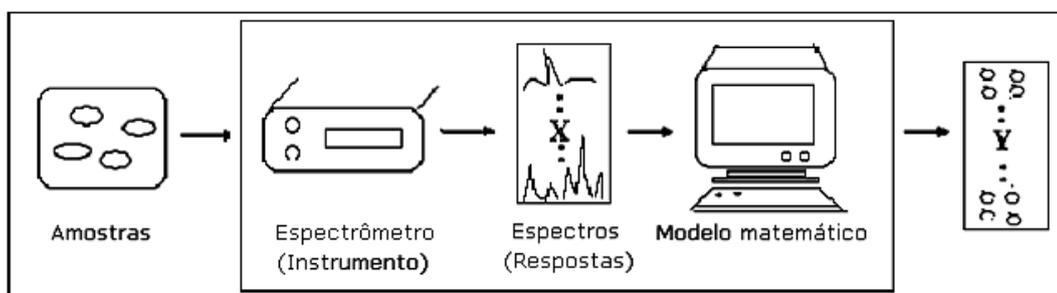
### 1.7.1 - Métodos de Pré-Processamento de Dados

Em geral, as matrizes de dados de natureza multivariada necessitam de manipulações matemáticas que possibilitem contornar ou reduzir fontes de variação não informativas para futuras modelagens. Estas manipulações são conhecidas como pré-processamento ou pré-tratamento de dados.

Em medidas baseadas em espectros NIR, geralmente é possível observar ruído aleatório e variação da linha de base. A técnica de pré-processamento mais simples para superar essas variações não informativas é a técnica de derivação de Savitzky-Golay, como muito bem discutido por Beeb *et al*<sup>[37]</sup>.

### 1.7.2 - Métodos de Calibração Multivariada

O processo de calibração multivariada baseado em dados espectrais consiste em construir um modelo matemático que relacione as medidas de absorvância, ao longo dos espectros, com os parâmetros de interesse (concentrações de analitos, acidez, viscosidade, índice de refração, etc) nas amostras. O diagrama abaixo (**Figura 1.6**) representa o procedimento de uma calibração multivariada usando o método espectroscópico.



**Figura 1.6.** Representação do procedimento de uma calibração multivariada (Fonte: FERREIRA, M.)<sup>[46]</sup>.

Para uma melhor compreensão das equações que serão apresentadas neste trabalho, vale a pena indicar a notação que será usada, como mostrada a seguir.

### NOTAÇÃO:

Matrizes e operadores lineares são representados por letras maiúsculas em negrito, os vetores coluna por letras minúsculas em negrito e os escalares por caracteres em itálico. Os elementos de uma seqüência são denotados por caracteres em itálico com um índice entre parênteses. O sobrescrito T significa transposta e o símbolo ( $\wedge$ ) indica os valores previstos.

Na construção de modelos de calibração multivariada, os dados espectrais de natureza multivariada (**Figura 1.7**) são representados numa matriz **X**, de dimensão ( $n \times m$ ), onde  $n$  representa o número de amostras cujos espectros foram registrados (ou número de linhas da matriz **X**) e  $m$  representa o número de variáveis (comprimentos de onda) em que as absorbâncias foram medidas (ou número de colunas da matriz **X**). Os parâmetros de interesse são representados pela matriz **Y** de dimensão ( $n \times p$ ), onde  $p$  são os parâmetros a serem estimados pelos modelos construídos<sup>[46]</sup>.

$$\begin{array}{c}
 \mathbf{X} = \begin{vmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1m} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & X_{n3} & \cdots & X_{nm} \end{vmatrix} \\
 \text{Variáveis independentes}
 \end{array}
 \qquad
 \begin{array}{c}
 \mathbf{Y} = \begin{vmatrix} Y_{11} & Y_{12} & Y_{13} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & Y_{23} & \cdots & Y_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ Y_{n1} & Y_{n2} & Y_{n3} & \cdots & Y_{np} \end{vmatrix} \\
 \text{Variáveis dependentes}
 \end{array}$$

**Figura 1.7.** Representação de matrizes de dados (X e Y)

Dentre os métodos mais empregados na calibração multivariada podemos citar a regressão linear múltipla (MLR) e a regressão por mínimos quadrados parciais (PLS), que foram utilizados nesse trabalho.

MLR é um método simples que usa todas as informações disponíveis contidas nas variáveis independentes (matriz  $\mathbf{X}$ ), incluindo os ruídos instrumentais. Apesar de sua simplicidade, os modelos obtidos por este método apresentam algumas restrições que muitas vezes limitam a sua aplicação, como por exemplo, o número de variáveis não pode ser maior do que o número de amostras de calibração. Outra limitação do método é que ele pode produzir modelos com problemas de colinearidade (instabilidade numérica de regressão). Todavia, esse problema pode ser reduzido empregando um processo de seleção de variáveis.

O método PLS é mais adequado para lidar com os problemas apontados no MLR, pois reduz a colinearidade, a quantidade de ruídos inseridos nos modelos e utiliza toda a informação espectral. Além disso, a calibração pode ser realizada mesmo na presença de interferentes. O princípio do método PLS consiste em modelar as matrizes  $\mathbf{X}$  e  $\mathbf{Y}$  em fatores ou variáveis latentes (VL), não necessitando assim, o uso de métodos de seleção de variáveis. As matrizes modeladas formam dois grupos de matrizes de escores ( $\mathbf{P}$  e  $\mathbf{Q}$ ) e pesos ( $\mathbf{R}$  e  $\mathbf{S}$ ) e dois blocos de equações, como mostradas nas **Equações 1.3 e 1.4**.

$$\mathbf{X} = \mathbf{PR}^T + \mathbf{E} \quad (1.3)$$

$$\mathbf{Y} = \mathbf{QS}^T + \mathbf{F} \quad (1.4)$$

onde  $\mathbf{E}$  e  $\mathbf{F}$  representam as matrizes residuais.

O modelo final relaciona linearmente os escores do bloco X (matriz  $\mathbf{P}$ ) com os escores do bloco Y (matriz  $\mathbf{Q}$ ), conforme mostram as **Equações 1.5 e 1.6**.

$$\mathbf{Q} = \mathbf{BP} + \mathbf{G} \quad (1.5)$$

$$\mathbf{Y} = \mathbf{BPS}^T + \mathbf{H} \quad (1.6)$$

onde **B** é a matriz dos coeficientes de regressão linear e **G** e **H** são as matrizes residuais.

A construção de cada variável latente é o resultado da combinação linear das variáveis originais e os pesos são os coeficientes dessa combinação. Eles indicam quanto cada variável original contribui para uma dada variável latente, enquanto os escores apontam diferenças e similaridades entre as amostras.

Para determinar o número ótimo de VLs, duas técnicas de validação são comumente utilizadas: série de teste e validação cruzada. O método de validação cruzada, utilizado neste trabalho, consiste em usar uma matriz contendo as amostras de calibração, cujo procedimento se faz com a retirada de uma ou mais amostras para a predição e as amostras remanescentes são utilizadas para modelagem. Repete-se este procedimento até que se tenha feito a predição com todas as amostras. O número ótimo de VLs será aquele capaz de produzir o menor erro médio quadrático de predição (RMSEP), que pode ser calculado de acordo com a **Equação 1.7**.

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (y_i^{prev} - y_i^{verd})^2}{n}} \quad (1.7)$$

onde  $y^{prev}$  corresponde ao valor previsto pelo modelo,  $y^{verd}$  é o valor verdadeiro para a amostra  $i$  e  $n$ , o número total de amostras.

De acordo com o exposto, pode-se dizer que o modelo PLS é geralmente empregado quando não há necessidade de se utilizar um processo de seleção de variáveis, que é o caminho mais viável para reduzir problemas de colinearidade. No entanto, quando há necessidade de utilizar um processo de seleção de variáveis, o MLR deve ser preferido porque trabalha com variáveis originais não modificadas e permite uma interpretação física e química do processo de forma mais fácil e direta<sup>[37,58,59]</sup>.

### 1.7.2.1 - Seleção de Variáveis

A escolha da região espectral é muito importante para a eficiência de um modelo multivariado. Quando escolhida adequadamente, permite minimizar os erros de predição de forma a obter modelos mais robustos e de simples interpretação.

Técnicas de seleção de variáveis têm sido muito usadas para este propósito e têm obtido muito sucesso em diversos trabalhos. São algoritmos matemáticos que selecionam as variáveis mais informativas para a construção de um modelo de calibração, eliminando as variáveis não relevantes<sup>[37,60]</sup>. Alguns exemplos são o algoritmo genético (GA)<sup>[61,54]</sup>, o algoritmo dos mínimos quadrados parciais por intervalos (*i*PLS)<sup>[62,63]</sup> e o algoritmo das projeções sucessivas (SPA)<sup>[52,55]</sup>. Nas duas sessões a seguir, serão discutidas as técnicas de seleção de variáveis utilizadas nesse trabalho.

#### 1.7.2.1.1 - Algoritmo *i*PLS

O algoritmo *i*PLS, desenvolvido por Nørgaard e colaboradores<sup>[62]</sup>, é uma extensão interativa desenvolvida para o PLS, onde é feita uma regressão por mínimos quadrados parciais em cada subintervalo equidistante, previamente definido, ao longo de toda a faixa espectral. Deste modo, é avaliada a relevância da informação nas diferentes subdivisões espectrais e é possível identificar e selecionar faixa de variáveis que apresenta informações mais relevantes e com o menor valor do erro médio quadrático de validação cruzada (RMSECV), cuja equação será apresentada na sessão seguinte. O modelo é então construído na faixa selecionada e comparado com aquele construído utilizando toda a faixa espectral<sup>[62,63]</sup>.

### 1.7.2.1.2 - Algoritmo das Projeções Sucessivas

Araújo e colaboradores<sup>[52]</sup> desenvolveram um método de seleção de variáveis denominado de Algoritmo das Projeções Sucessivas (SPA), cujo objetivo é reduzir a colinearidade de um conjunto de dados de modelos MLR, melhorando assim, a capacidade preditiva desses modelos. O princípio do algoritmo é baseado numa busca sistemática de variáveis informativas de modo que a operação seja iniciada partindo-se de um comprimento de onda ou vetor inicial escolhido, e em seguida, outros comprimentos de onda (vetores), não-colineares ao vetor escolhido, são incorporados um a um a cada iteração, em projeções sucessivas, até que um número específico  $N$  de comprimentos de onda seja alcançado.

Considere:

- ✚  $k(n)$  como o comprimento de onda na  $n$ -ésima iteração do SPA;
- ✚  $X_{cal}$  como a matriz ( $M_{cal} \times J$ ) de dados das respostas instrumentais (variáveis independentes) para  $M_{cal}$  misturas de calibração e  $J$  comprimentos de onda;
- ✚  $Y_{cal}$  como sendo a matriz ( $M_{cal} \times A$ ) de concentrações de calibração (variáveis dependentes) para  $A$  analitos;
- ✚  $Y_{teste}$  com sendo a matriz ( $M_{teste} \times A$ ) de concentração para  $M_{teste}$  misturas de teste.

O procedimento realizado pelo SPA pode ser exemplificado em algumas etapas que serão mostradas abaixo, onde o vetor  $k(0)$  é escolhido como o vetor de partida e  $N$ , que é conhecido, é o número de comprimentos de onda que serão extraídos.

#### ETAPAS:

1. Antes da primeira iteração ( $n=1$ ), faz-se  $x_j = j$ -ésima coluna de  $X_{cal}$ ;  $j = 1, \dots, J$ ;

2. Um conjunto  $S$  de comprimentos de onda que ainda não foram selecionados é criado, de forma que  $S = \{j \text{ tal que } 1 \leq j \leq J \text{ e } j \notin \{k(0), \dots, k(N-1)\}\}$ ;

3. Calcula-se a projeção de  $\mathbf{x}_j$  no subespaço ortogonal para  $\mathbf{x}_{k(n-1)}$ :

$$P\mathbf{x}_j = \mathbf{x}_j - \left(\mathbf{x}_j^T \mathbf{x}_{k(n-1)}\right) \mathbf{x}_{k(n-1)} \left(\mathbf{x}_{k(n-1)}^T \mathbf{x}_{k(n-1)}\right)^{-1} \quad (1.8)$$

onde  $P$  é o operador de projeção;

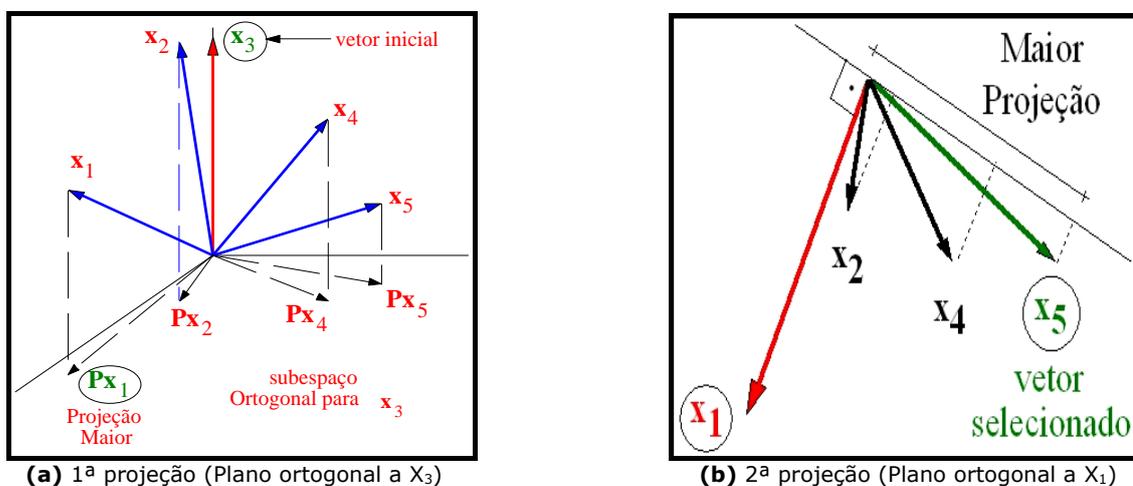
4. Um comprimento de onda é selecionado, onde  $k(n) = \arg(\max_{j \in S} \|\mathbf{P}\mathbf{x}_j\|)$ ;

5. Faz-se  $\mathbf{x}_j = \mathbf{P}\mathbf{x}_j, \forall j \in S$ ;

6. O processo é repetido até que  $n = n + 1$ . Se  $n < N$ , retorna-se para a etapa 1;

7. Nesta última etapa, obtêm-se os resultados dos comprimentos de onda que são  $\{k(n); n = 0, \dots, N-1\}$ .

O número de operações de projeções executadas no processo de seleção é dado por  $(N-1)(J-N/2)$ . A **Figura 1.8** mostra uma visão espacial do funcionamento do SPA para uma melhor compreensão das etapas descritas acima.



**Figura 1.8.** Representação do funcionamento do algoritmo SPA: projeção do plano ortogonal a  $X_3$  (a) e visão espacial dos vetores no plano horizontal (b).

Caso  $k(0)$  e  $N$  não sejam conhecidos, alguns critérios são estabelecidos para determinar os valores ótimos  $k^*(0)$  e  $N^*$  da seguinte forma:

- ✚ Especifica-se um conjunto de validação;
- ✚ Especifica-se um número mínimo ( $N_{min}$ ) e um máximo ( $N_{max}$ ) de comprimentos de onda para posterior seleção de forma que se possa encontrar o número ótimo de comprimentos de onda ( $N^*$ ), onde  $N$  vai de  $N_{min}$  a  $N_{max}$  e o *inicial* vai de  $j$  a  $J$ .

Deste modo, usando-se as etapas descritas anteriormente:

- ✚ Seleciona-se  $N$  comprimentos de onda, começando por  $k(0) = inicial$ ;
- ✚ Constrói-se um modelo MLR usando os comprimentos de onda selecionados;
- ✚ Usa-se o modelo para prever as amostras do conjunto de validação, calculando o erro médio quadrático de validação cruzada RMSECV, definido por:

$$RMSECV = \sqrt{\frac{1}{N_{cal}} \sum_{n=1}^{N_{cal}} (y_{cal,n} - \hat{y}_{cal,n})^2} \quad (1.9)$$

onde  $y_{cal,n}$  e  $\hat{y}_{cal,n}$  correspondem aos valores de referência e previsto, respectivamente, para  $n$  amostras e  $N_{cal}$  corresponde ao número de amostras do conjunto de calibração<sup>[52,55,64]</sup>.

### **1.8 - Objetivo do Trabalho**

Este trabalho tem como objetivo desenvolver um novo método que utilize a espectrometria NIR, a calibração multivariada, a seleção de faixas de variáveis (*i*PLS) e a seleção de variáveis (SPA) para a determinação simultânea de acidez, índice de refração e viscosidade em quatro tipos de óleos vegetais (óleos de soja, milho, girassol e canola).

# **CAPÍTULO II**

## **EXPERIMENTAL**

## 2.1 - Amostras

Setenta amostras de óleos vegetais (soja, milho, girassol e canola), de diferentes lotes e fabricantes, foram adquiridas em diversos supermercados da cidade de João Pessoa, Paraíba, no período de agosto a setembro de 2006.

## 2.2 - Divisão do Conjunto de Amostras

O algoritmo SPXY (do inglês: *Sample set partitioning based on joint x-y distances*), recentemente proposto por Galvão e colaboradores<sup>[65]</sup>, foi empregado para dividir as setenta amostras em dois conjuntos: 45 para calibração e 25 para a predição.

## 2.3 – Análises dos Óleos Vegetais por Métodos de Referência

Acidez, índice de refração e viscosidade dos quatro tipos de óleos vegetais (óleo de soja, milho, girassol e canola) foram analisados por métodos de referência conforme recomendado pela AOCS<sup>[21]</sup>.

As medidas de acidez foram realizadas em duplicata utilizando o seguinte procedimento:

- ✚ Pesou-se  $5,0 \pm 0,2$  g de cada amostra em um erlemeyer de 125 ml. Em seguida, foi adicionado 25 ml de uma solução de álcool etílico a 95%, sob constante agitação e, posteriormente, duas gotas de solução alcoólica de fenolftaleína 1%;
- ✚ Finalmente, cada solução foi titulada com uma solução de hidróxido de sódio (NaOH)  $0,1000 \text{ mol L}^{-1}$ , sob agitação, até o aparecimento de uma cor levemente rósea, que indicou o ponto final da titulação.

Segundo a AOCS, a acidez livre de óleos vegetais comestíveis é expressa em percentagem de ácido oléico, como mostra a **Equação 2.1**.

$$\% \text{ de ácido oléico} = \frac{V \times M \times 28,2}{\text{massa da amostra (g)}} \quad (2.1)$$

onde:

V = volume de hidróxido de sódio gasto na titulação;

M = Concentração de hidróxido de sódio em mol L<sup>-1</sup>.

Para a determinação da viscosidade dos óleos, as medidas foram realizadas em quintuplicata utilizando o viscosímetro Viscotest 6L da marca ThermoHaake (**Figura 2.1**). Para processar a leitura de uma medida para outra, o acessório do viscosímetro, conhecido como *Stringer*, foi adequadamente limpo com uma solução éter-álcool (1:1).



**Figura 2.1.** Viscosímetro Viscotest 6L.

As medidas para determinação do índice de refração foram realizadas em triplicata utilizando-se um refratômetro de Abbé da marca BAUSH & LOMB (**Figura 2.2**). Nos intervalos das medidas, os prismas do refratômetro foram cuidadosamente limpos usando-se, também, uma solução éter-álcool (1:1).



**Figura 2.2.** Refratômetro de Abbé.

## 2.4 - Registro dos Espectros NIR

Os espectros NIR das amostras de óleos vegetais foram adquiridos usando um espectrômetro de infravermelho com Transformada de Fourier, da marca Perkin Elmer, modelo Spectrum, série GX e uma cubeta de fluxo de quartzo de 1 mm de caminho óptico. Uma bomba peristáltica da marca Gilson e modelo Miniplus 3 foi usada para aspiração das amostras até o enchimento da cubeta.

Cada espectro NIR registrado foi sempre um espectro médio resultante de 16 varreduras, empregando uma resolução de  $4\text{ cm}^{-1}$ . O espectro do branco foi registrado com o ar ambiente. A umidade relativa do ar e a temperatura ambiente encontravam-se em torno de 45% e  $24^{\circ}\text{C}$ , respectivamente.

A limpeza da cubeta de fluxo foi feita com a amostra a ser analisada aspirando-se bolhas de ar em intervalos. Finalmente, o registro de todas as amostras foi arquivado num microcomputador para posterior tratamento quimiométrico.

A **Figura 2.3** mostra os detalhes do instrumento utilizado para o registro dos espectros.



**Figura 2.3.** Sistema para registro dos espectros NIR dos óleos vegetais. **A:** espectrômetro NIR; **B:** amostra de óleo vegetal; **C:** cubeta de fluxo de quartzo; **D:** bomba peristáltica; **E:** descarte.



**Figura 2.4.** Cubeta de fluxo de quartzo de 1 mm de caminho óptico.

## 2.5 – Softwares

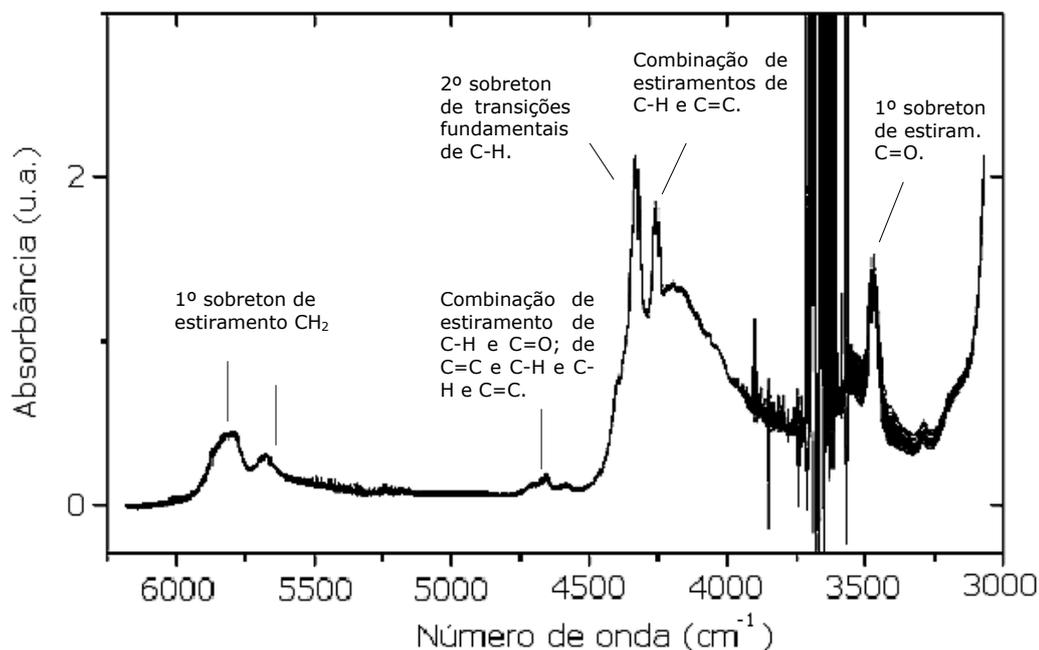
O pré-processamento dos dados e a elaboração dos modelos PLS foram realizados utilizando o programa *Unscrambler*® 9.6 da Camo S/A.

Os algoritmos de seleção de amostras (SPXY)<sup>[65]</sup> e variáveis (*i*PLS<sup>[62,63]</sup> e SPA-MLR<sup>[64]</sup>) foram construídos e executados em ambiente Matlab 6.5.

**CAPÍTULO III**  
**RESULTADOS E**  
**DISCUSSÕES**

### 3.1 - Seleção da Região Espectral de Trabalho

Os espectros das amostras de óleos vegetais foram registrados na região de 6184 a 3068  $\text{cm}^{-1}$ , conforme ilustra a **Figura 3.1**.



**Figura 3.1.** Espectros NIR brutos das 70 amostras de óleos vegetais (6184-3068  $\text{cm}^{-1}$ ).

### 3.2 - Atribuição das Bandas dos Espectros NIR

Analisando os espectros da **Figura 3.1** é claramente perceptível que a região compreendida entre 3752 e 3520  $\text{cm}^{-1}$  encontra-se saturada, mesmo assim, ela foi mantida a fim de comprovar a habilidade do algoritmo *i*PLS de descartar os intervalos de números de onda não informativos. Portanto, toda a região espectral foi utilizada como a região inicial de trabalho.

Pode-se verificar na **Figura 3.1** que as bandas que possuem dois picos máximos em aproximadamente 5791 e 5677  $\text{cm}^{-1}$  estão associadas ao 1º sobreton de estiramento  $\text{CH}_2$ . Os três pequenos picos em aproximadamente 4707, 4662 e 4595  $\text{cm}^{-1}$  estão associados à combinação de estiramento de C-H e C=O de ácidos graxos insaturados, à combinação

de estiramento de C=C e C-H e à combinação de estiramento assimétrico de C-H e C=C, respectivamente.

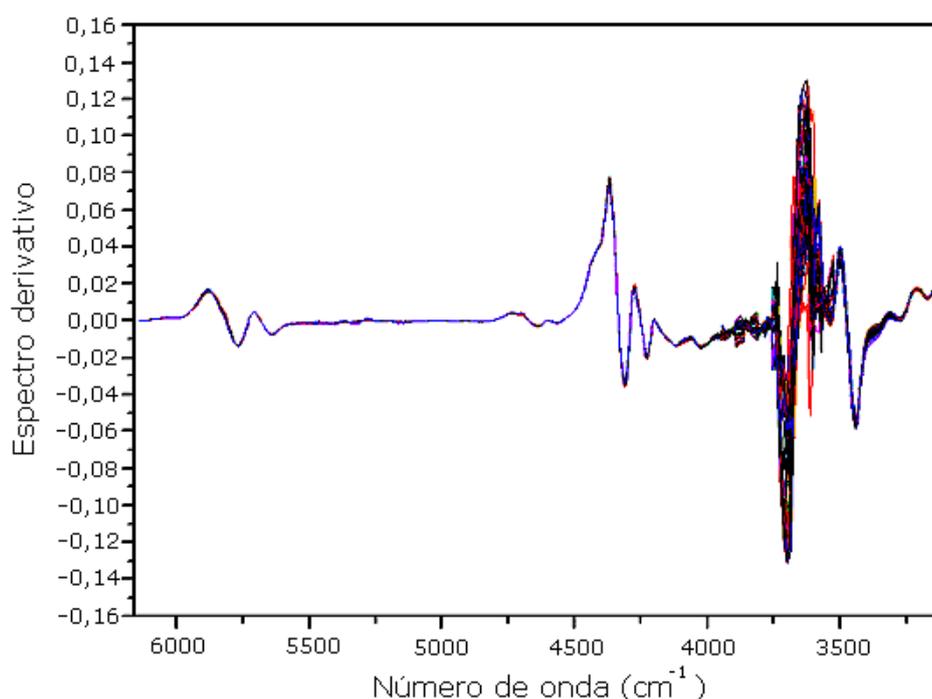
Os dois picos largos em aproximadamente 4329 e 4258  $\text{cm}^{-1}$ , estão associados ao segundo sobreton de transições fundamentais relativas às ligações C-H (de grupos metileno) e a combinação de estiramentos de C-H e C=C, respectivamente. O pico que se encontra em 3468  $\text{cm}^{-1}$  na região MIR apresenta uma absorção relativa ao 1º sobreton de estiramento C=O. É importante ressaltar que em 3500  $\text{cm}^{-1}$  é esperado um pico característico relacionado ao estiramento de O-H que deve estar sobreposto pela banda da carbonila.

Não há nenhuma informação relativa ao pico 3288  $\text{cm}^{-1}$  associado a análises de óleos vegetais, provavelmente devido às intensidades extremamente fracas que não são associadas aos parâmetros de interesse se comparados com os sinais intensos encontrados na região MIR. O aparecimento desta banda se deve, provavelmente, à formação de dímeros de ácidos carboxílicos e está associada, portanto, ao 1º sobreton de estiramento C=O, que se encontra entre 3383 e 3136  $\text{cm}^{-1}$ , aproximadamente<sup>[45]</sup>.

### 3.3 - Pré-Processamento dos Dados Espectrais

Os espectros apresentados na **Figura 3.1** apresentam problemas de variação sistemática de linha de base, bem como a presença de ruídos. Para contornar estes problemas, utilizou-se a 1ª derivada pelo método de Savitzky-Golay<sup>[37]</sup>, empregando um polinômio de segunda ordem e uma janela de 21 pontos. Desta forma, os espectros derivativos resultantes apresentaram 760 variáveis.

A **Figura 3.2** mostra os espectros derivativos das 70 amostras de óleos vegetais analisados.



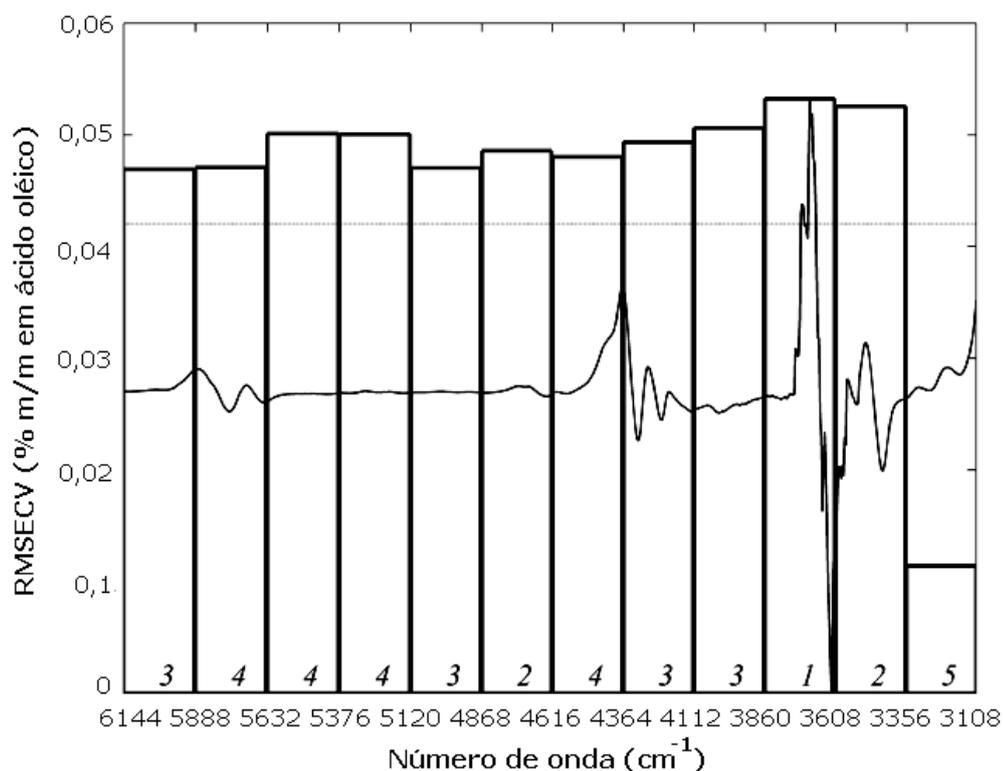
**Figura 3.2.** Espectros derivativos das 70 amostras de óleos vegetais analisados.

Além do pré-processamento aplicado às amostras, foi feito também um pré-processamento nas variáveis, centralizando os dados na média.

### 3.4 - Elaboração dos Modelos PLS, *i*PLS e SPA-MLR

O algoritmo *i*PLS foi empregado dividindo o espectro derivativo apresentado na **Figura 3.2** em 12 intervalos eqüidistantes. Conforme visto na **Seção 1.7.2.1.1**, o critério adotado pelo *i*PLS para a escolha da melhor região é baseado no menor RMSECV<sup>[62]</sup>.

As **Figuras 3.3 a 3.5** mostram os resultados obtidos pela aplicação do *i*PLS para os três parâmetros estudados. Em cada uma dessas figuras é também mostrado um espectro derivativo médio normalizado para uma melhor interpretação dos gráficos. A linha horizontal pontilhada representa o valor do RMSECV obtido pelo modelo PLS que emprega toda a região de trabalho. O número apresentado no interior de cada bloco de intervalo, acima do eixo da abscissa, corresponde ao número de variáveis latentes.



**Figura 3.3.** Intervalos de variáveis selecionados pelo *i*PLS para acidez.

Para a acidez (**Figura 3.3**), a sub-região de menor valor de RMSECV selecionada pelo *i*PLS foi compreendida na faixa de 3356 a 3108  $\text{cm}^{-1}$ , com 5 variáveis latentes. É importante ressaltar que esta região é adequada para análise de acidez porque ela está relacionada à absorção de estiramento C=O de ácido carboxílico, como explicado na **Seção 3.2**.

O *i*PLS indica que o intervalo espectral compreendido entre 5888 e 5636  $\text{cm}^{-1}$  é aquele que produz o menor valor de RMSECV tanto para o índice de refração (**Figura 3.4**) quanto para a viscosidade (**Figura 3.5**). Os modelos *i*PLS construídos apresentaram 6 e 5 variáveis latentes, respectivamente. Vale salientar que a região selecionada para ambos os parâmetros está associada ao estiramento C-H da cadeia alifática, característica dos ácidos carboxílicos. Esta correlação é esperada uma vez que os valores dessas propriedades tendem a aumentar com o crescimento da cadeia alifática e, também, a viscosidade tende a diminuir com o aumento do número de saturações da cadeia<sup>[66]</sup>. Em todos os casos os valores de RMSECV obtidos pelos modelos *i*PLS foram menores do que aqueles obtidos pelo PLS que emprega toda faixa espectral.

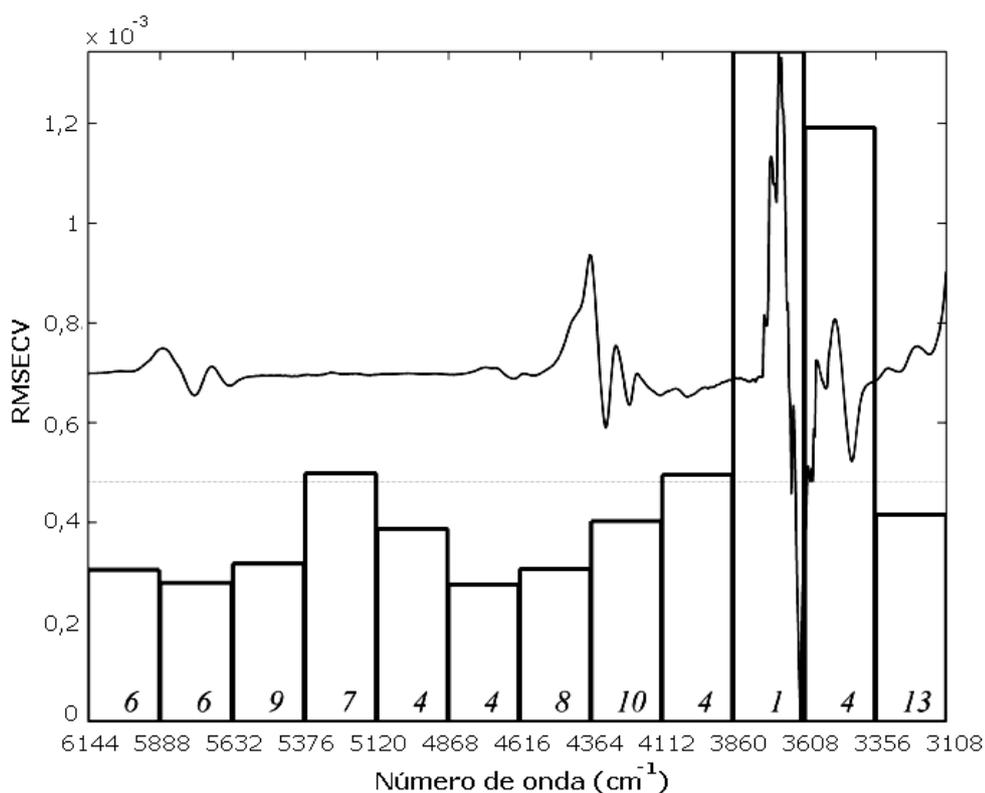


Figura 3.4. Intervalos de variáveis selecionados pelo *i*PLS para índice de refração.

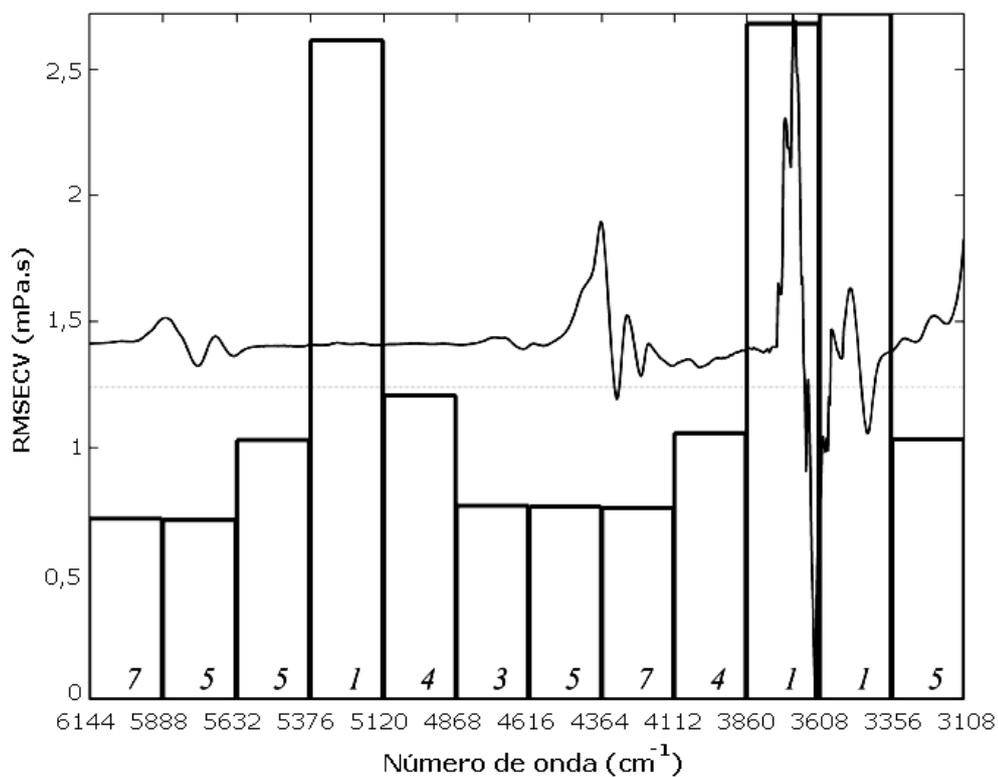
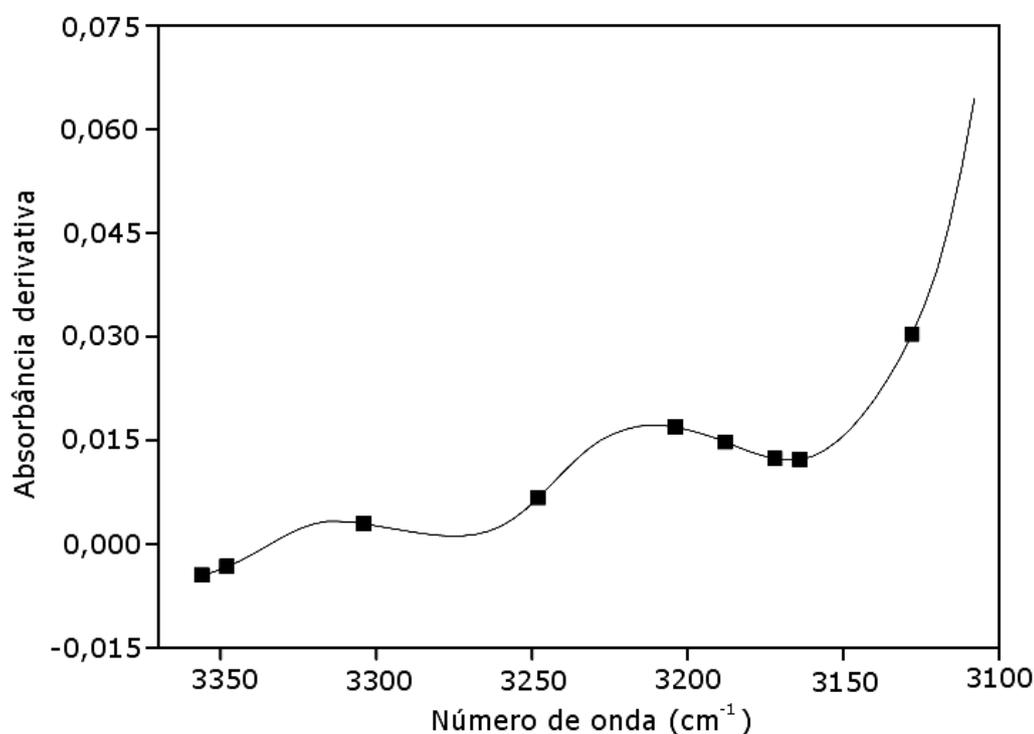


Figura 3.5. Intervalos de variáveis selecionados pelo *i*PLS para viscosidade.

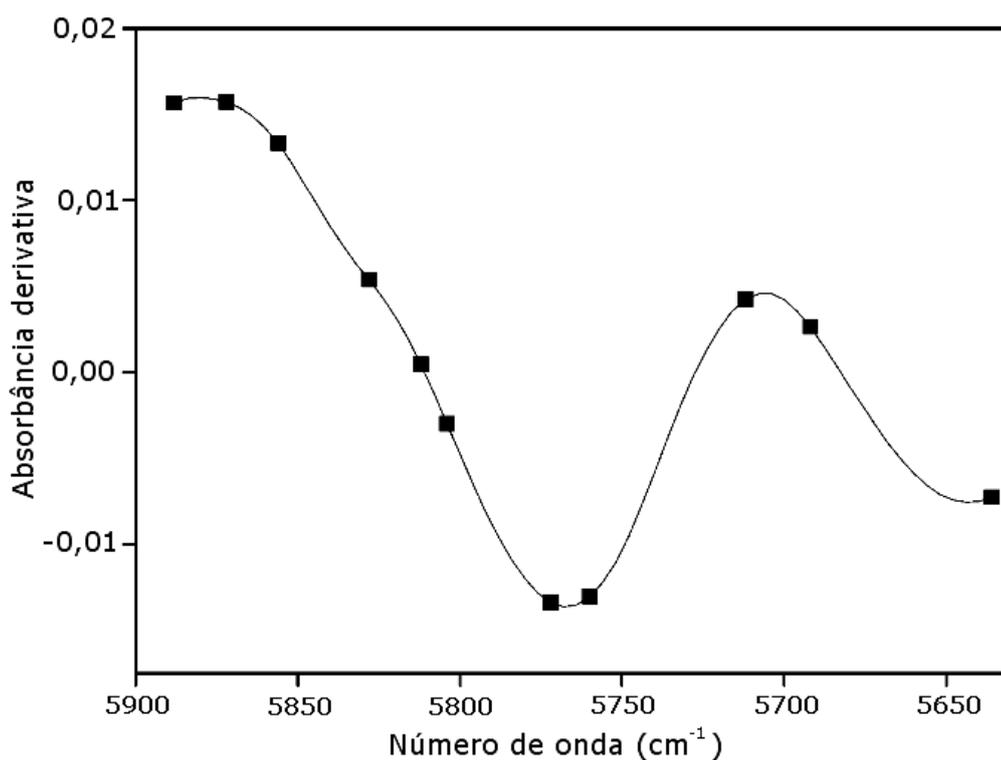
É importante lembrar que os resultados obtidos pelo *i*PLS para os intervalos compreendidos entre 3860 a 3360  $\text{cm}^{-1}$  foram os piores para todos os três parâmetros estudados. De fato, esta faixa inclui a região de 3752 a 3520  $\text{cm}^{-1}$  (**Figura 3.1**) que apresenta saturação no sinal de absorbância, como explanado na **Sessão 3.2**.

Após a escolha das melhores regiões espectrais pelo *i*PLS, o SPA-MLR foi empregado para selecionar os números de onda mais relevantes nessas regiões.

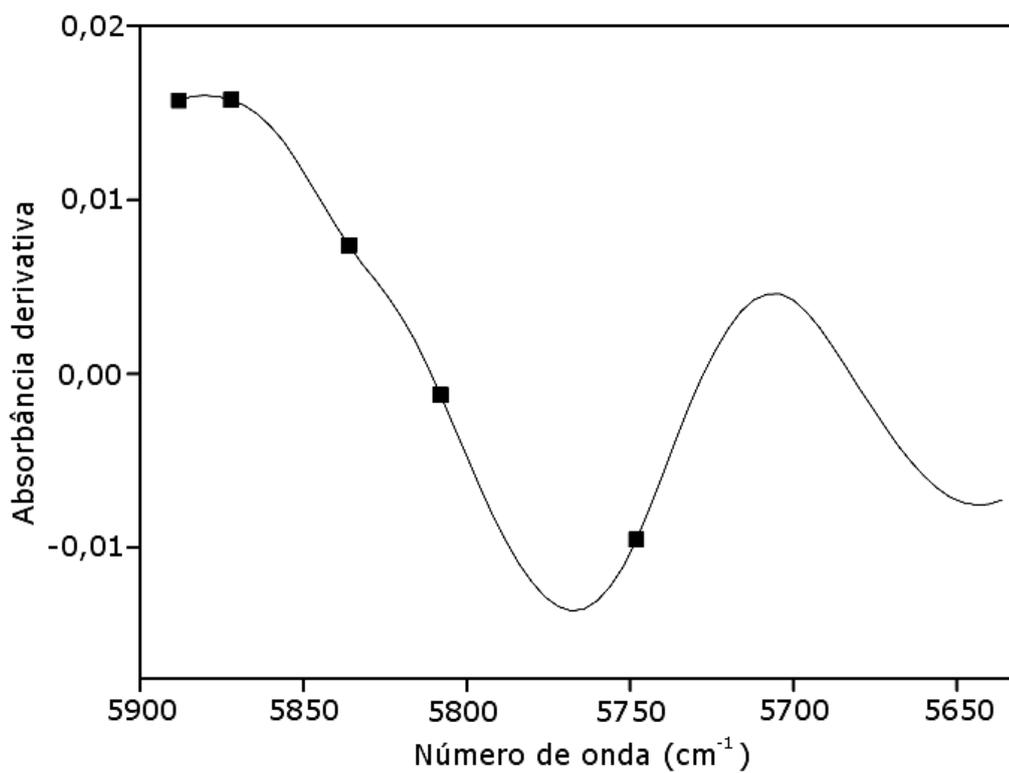
O SPA-MLR selecionou apenas 9 números de onda mais informativos para o parâmetro acidez no intervalo previamente selecionado pelo *i*PLS (3356 a 3108  $\text{cm}^{-1}$ ). Os números de onda selecionados nesta região estão indicados na **Figura 3.6**. Para o índice de refração e viscosidade, o SPA-MLR selecionou 11 e 5 números de onda que estão indicados nas **Figuras 3.7** e **3.8**, respectivamente.



**Figura 3.6.** Variáveis escolhidas pelo SPA-MLR para acidez no intervalo do *i*PLS.



**Figura 3.7.** Variáveis escolhidas pelo SPA-MLR para índice de refração no intervalo *i*PLS.



**Figura 3.8.** Variáveis encolhidas pelo SPA-MLR para viscosidade no intervalo do *i*PLS.

Para efeito de comparação, o conjunto de 25 amostras foi empregado para avaliar a habilidade de predição dos modelos. A **Tabela 3.1** apresenta os valores dos resultados de RMSEP e correlação ( $r$ ), assim como o número de variáveis latentes dos modelos PLS e  $i$ PLS e de números de onda dos modelos SPA-MLR, respectivamente.

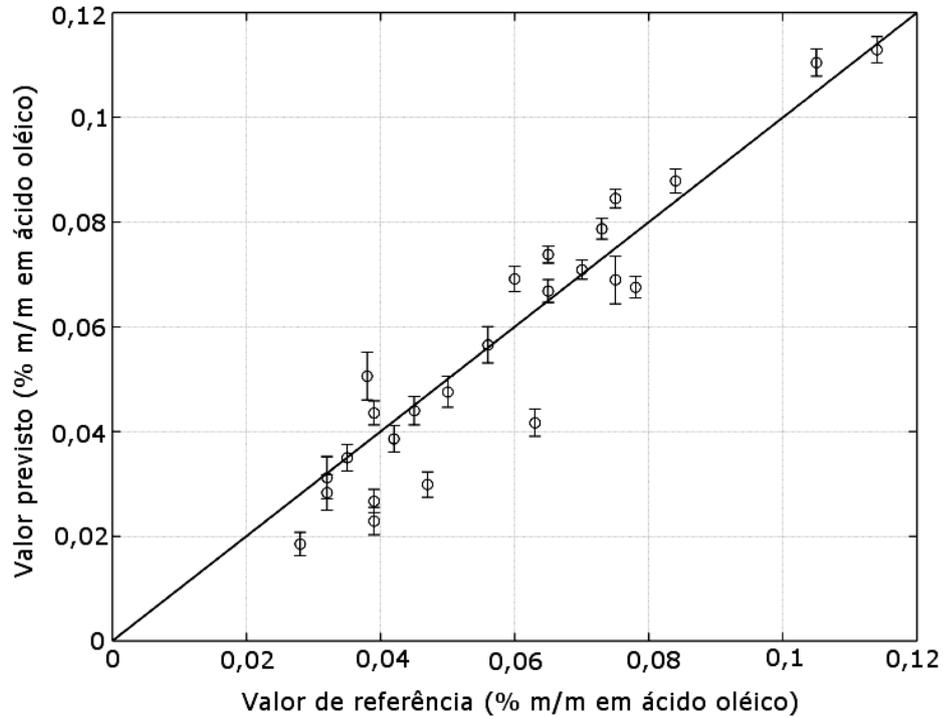
**Tabela 3.1.** Valores de RMSEP e correlação ( $r$ ) dos modelos PLS,  $i$ PLS e SPA-MLR\*.

| Parâmetro          | Faixa de calibração                   | PLS (toda a faixa) |            | $i$ PLS    |            | SPA-MLR     |            | PLS (após remoção da região saturada) |            |
|--------------------|---------------------------------------|--------------------|------------|------------|------------|-------------|------------|---------------------------------------|------------|
|                    |                                       | RMSEP              | $r_{pred}$ | RMSEP      | $r_{pred}$ | RMSEP       | $r_{pred}$ | RMSEP                                 | $r_{pred}$ |
| Acidez             | (0,025 – 0,294 % p/p em ácido oléico) | 0,012 (28)         | 0,86       | 0,010 (5)  | 0,94       | 0,009 (9)   | 0,94       | 0,014 (9)                             | 0,85       |
| Índice de refração | 1,4690 – 1,4730                       | 0,0006 (13)        | 0,94       | 0,0003 (6) | 0,97       | 0,0003 (11) | 0,98       | 0,0006 (3)                            | 0,91       |
| Viscosidade        | 62,0 – 71,4 (mPa.s)                   | 1,0 (13)           | 0,92       | 0,8 (5)    | 0,95       | 0,8 (5)     | 0,96       | 1,0 (2)                               | 0,91       |

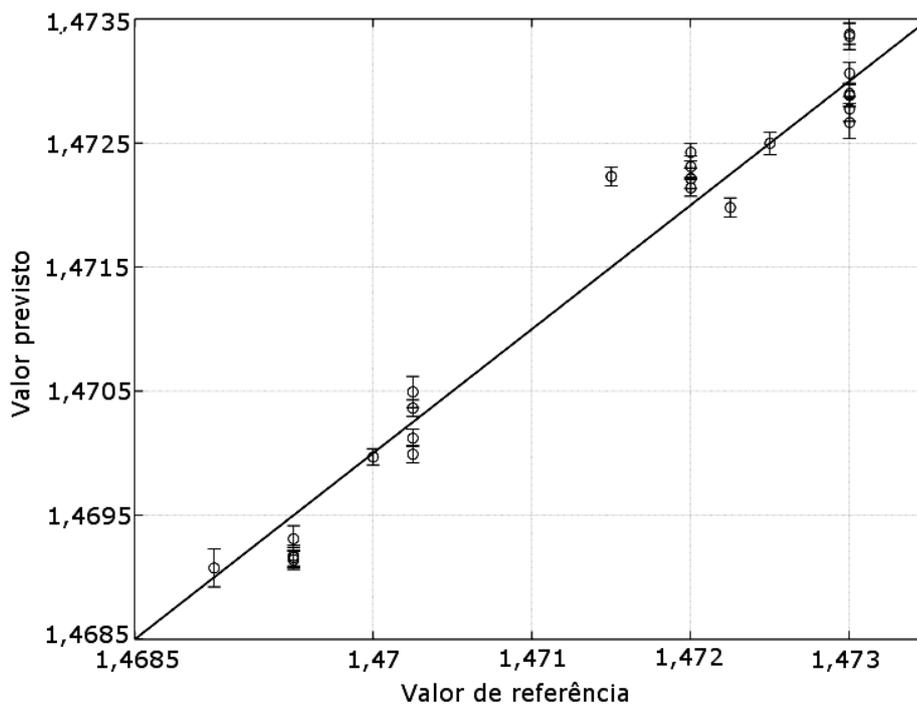
\*O nº de variáveis latentes do PLS e  $i$ PLS e os nºs de onda do SPA-MLR estão indicados entre parênteses.

De acordo com a **Tabela 3.1**, os números de variáveis latentes obtidos pelo  $i$ PLS foram consideravelmente menores quando comparados com aqueles obtidos pelo PLS empregado em toda região espectral. De fato, muitas fontes de variação não informativas, por exemplo, aquela da região saturada, não pertence à região escolhida pelo  $i$ PLS. Entretanto, as informações relevantes usadas para o conjunto de predição foram mantidas, como mostram os valores do RMSEP e  $r$  para todos os parâmetros após o emprego do  $i$ PLS. Pode-se ver, também, que resultados mais satisfatórios do RMSEP (acidez) e  $r$  (índice de refração e viscosidade) foram alcançados após o emprego do SPA-MLR. As duas últimas colunas da tabela mostram resultados ligeiramente melhores para os valores do RMSEP (com menos variáveis latentes) obtidos pelo PLS, após remoção da região saturada (3752 a 3520  $\text{cm}^{-1}$ ), se comparados com os valores obtidos quando o PLS foi empregado em toda região.

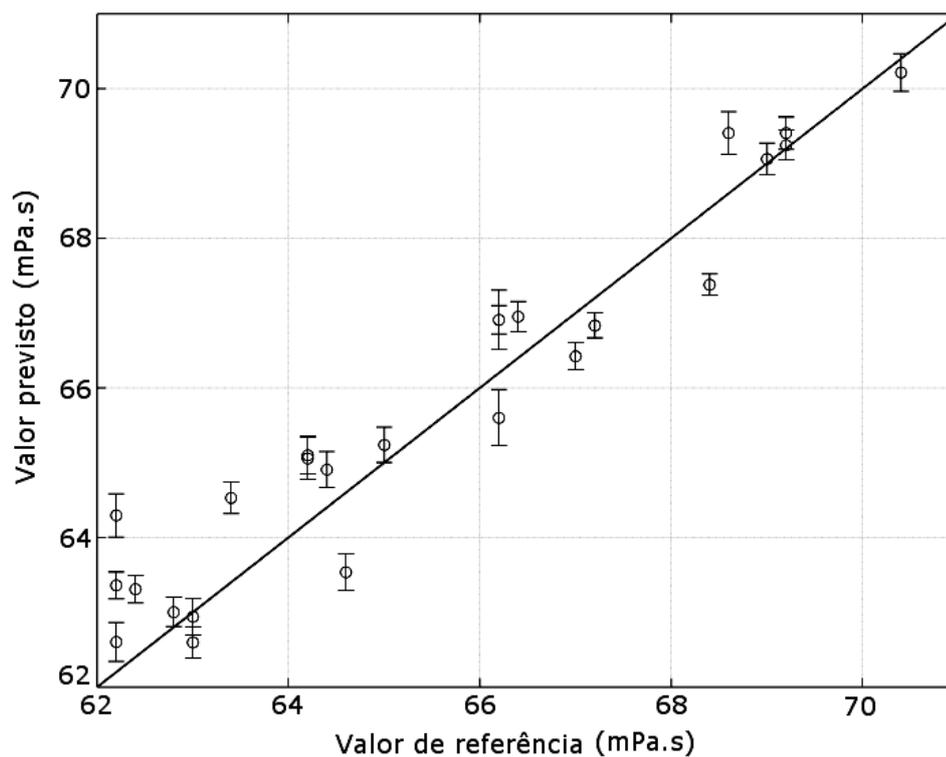
As **Figuras 3.9 a 3.11** mostram os valores previstos pelo modelo SPA-MLR *versus* os valores de referência para os 3 parâmetros estudados.



**Figura 3.9.** Valor previsto pelo SPA-MLR *versus* o valor de referência para acidez.



**Figura 3.10.** Valor previsto pelo SPA-MLR *versus* o valor de referência para o índice de refração.



**Figura 3.11.** Valor previsto pelo SPA-MLR *versus* o valor de referência para viscosidade.

Pode-se observar nas **Figuras 3.9** a **3.11** que todos os pontos estão distribuídos em torno da linha da bissetriz, mostrando que não há erro sistemático nas previsões dos modelos SPA-MLR dos três parâmetros estudados.

# **CAPÍTULO IV**

## **CONCLUSÕES**

#### 4.1 - Conclusões

Este trabalho propôs uma nova metodologia para a determinação simultânea de parâmetros de qualidade, tais como: índice de acidez, índice de refração e viscosidade em quatro tipos de óleos vegetais comestíveis (soja, milho, girassol e canola). Para este propósito, foi empregada a combinação da espectrometria NIR com um modelo simples por regressão linear múltipla (MLR) dos parâmetros estudados. Como partida, empregou-se o algoritmo das projeções sucessivas (SPA) para selecionar as variáveis mais informativas selecionadas previamente pelo algoritmo dos mínimos quadrados parciais por intervalos (iPLS). Os resultados mostraram que os modelos SPA-MLR obtidos mostraram-se eficientes e robustos, visto que apresentaram baixos valores de RMSEPs (0,009 % m/m; 0,0003 e 0,8 mPa.s) e bons valores de correlação (0,94, 0,98 e 0,96) para acidez, índice de refração e viscosidade, respectivamente. Com efeito, uma calibração simples pode ser realizada com sucesso para cada parâmetro, sem a necessidade de desenvolver um modelo separado para cada tipo de óleo vegetal.

#### 4.2 - Propostas Futuras

Pretende-se, como continuidade desse trabalho:

- 1. Ampliar o modelo desenvolvido para a determinação de outros parâmetros de qualidade, tais como: índice de iodo, índice de peróxido e densidade, bem como em outros tipos de óleos vegetais;*
- 2. Melhorar a capacidade de predição dos modelos elaborados utilizando técnicas que permitam remover possíveis ruídos presentes nos espectros, como a Transformada Wavelet, por exemplo.*

**CAPÍTULO V**  
**REFERÊNCIAS**

1. HARTMAN L.; ESTEVES W. Tecnologia de óleos e gorduras vegetais. São Paulo, 1982.
2. APARÍCIO, R.; APARICIO-RUÍZ R. Authentication of vegetable oils by chromatographic techniques. *Journal of Chromatography A*, v.881, P.93-104, 2000.
3. VOET D.; VOET J. G.; PRATT C. W. Fundamentos de Bioquímica. Porto Alegre, 2000.
4. BOBBIO, F. O.; BOBBIO, P. A. Introdução à Química de Alimentos. 2ª ed., Editora e Livraria Varela LTDA, São Paulo: 1989.
5. IBGE: [http://www.ibge.gov.br/home/presidencia/noticias/noticia\\_visualiza.php?id\\_noticia=859&id\\_pagina=1](http://www.ibge.gov.br/home/presidencia/noticias/noticia_visualiza.php?id_noticia=859&id_pagina=1) (acessado em 24/01/2007).
6. Ministério da Ciência e Tecnologia: Estudo da Competitividade da Indústria Brasileira - Competitividade na Indústria de Óleos Vegetais. Nota Técnica Setorial do Complexo Agroindustrial. Disponível em: [http://www.mct.gov.br/upd\\_blob/2225.pdf](http://www.mct.gov.br/upd_blob/2225.pdf) (acessado em 20/12/2006)
7. ZHANG, G.; NI, Y.; CHURCHILL J.; KOKOT S. Authentication of vegetable oils on the basis of their physico-chemical properties with the aid of chemometrics. *Talanta*, v.70, p.293-300, 2006.
8. ANVISA: [http://www.anvisa.gov.br/legis/resol/482\\_99.htm](http://www.anvisa.gov.br/legis/resol/482_99.htm) (acessado em 27/11/2006).
9. PETROBRAS: [http://www2.petrobras.com.br/Petrobras/portugues/perfil/Perfil\\_biodisel.htm](http://www2.petrobras.com.br/Petrobras/portugues/perfil/Perfil_biodisel.htm) (acessado em 24/01/2007).
10. EMBRAPA: <http://www.iea.usp.br/iea/online/midiateca/politicaspubblicascasgazzoni.pdf> (acessado em 30/12/2006).
11. EMBRAPA: [http://www.cnpso.embrapa.br/soja\\_alimentacao/index.php?pagina=23](http://www.cnpso.embrapa.br/soja_alimentacao/index.php?pagina=23) (acessado em 10/12/2006).
12. CAMPESTRE: <http://www.campestre.com.br/oleo-de-soja.shtml>. (acessado em 12/12/2006).
13. CAMPESTRE: <http://www.campestre.com.br/oleo-de-milho.shtml> (acessado em 12/12/2006).

- 
14. RODRIGUES, J. N.; GIOIELLI L. A.; ANTON C. Propriedades físicas de lipídios estruturados obtidos de misturas de gordura do leite e óleo de milho. *Ciênc. Tecnol. Aliment.*, v.23(2) p.226-233, 2003.
  15. REGINA R.; SOLFERINI O. Produção de cultivares de ingredientes de alto valor nutricional: Características e benefícios. Disponível em: [http://www.gtconsult.com.br/ingles/artigos/producao\\_cultivares.pdf](http://www.gtconsult.com.br/ingles/artigos/producao_cultivares.pdf) (acessado em 24/07/06).
  16. ABOISSA: <http://www.aboissa.com.br/girassol/compq.htm> (acessado em 07/07/2006).
  17. CAMPESTRE: <http://www.campestre.com.br/oleo-de-girassol.shtml> (acessado em 15/07/2006).
  18. EMBRAPA: <http://www.cnpt.embrapa.br/culturas/canola/definicao.htm> (acessado em 21/09/2006).
  19. <http://www.aaug.com.br/dicas.php> (acessado em 07/01/2007).
  20. PEREIRA, A. F. C.; PONTES, M. J. C.; GAMBARRA NETO, F. F.; SANTOS, S. R. B.; GALVÃO, R. K. H.; ARAÚJO, M. C. U. NIR spectrometric determination of quality parameters in vegetable oils using iPLS and variable selection. *Food Research International* (2007), doi: 10.1016/j.foodres.2007.12.013.
  21. AOCS (Ca 5a-40): American Oil Chemists' Society. 5ª ed., 1998.
  22. MORETTO, E.; FETT, R. Tecnologia de óleos e gorduras vegetais na indústria de alimentos. Editora e Livraria Varela LTDA, São Paulo, 1998.
  23. CERT, A.; W. MOREDA; M. C. PEREZ CAMINOS. Chromatographic analysis of minor constituents in vegetable oils. *J. Chromatogr. A*, v. 881, p.131-148, 2000.
  24. SACCO, A.; BRESCIA, M. A.; LIUZZI, V.; RENIERO, F.; GUILLOU, G.; GHELLI, S.; VAN DER MEER, P. Characterization of italian olive oils based on analytical and nuclear magnetic resonance determinations, *J. Am. Oil Chem. Soc.* v.77, p.619-625, 2000.
  25. LORENZO, I. M.; PAVÓN, J. L. P.; LAESPADA, M. E. F.; PINTO, C. G.; CORDERO, B. M. Detection of adulterants in olive oil by

- headspace-mass spectrometry, *J. Chromatogr. A*, v.945, p.221-230, 2002.
- 26.DAHLBERG, D. B.; LEE, S. M.; WENGER, S. J.; VARGO, J. A. Classification of vegetable oils by FT-IR. *Applied Chemometrics*, v.51, p.1118-1124, 1997.
- 27.LEE, DONG-SUN; NOH, BONG-SOO; BAE, SUN-YOUNG; KIM, K. Characterization of fatty acids compositions in vegetables oils by gas chromatography and chemometrics. *Analytica Chimica Acta*, v.358, p.163-175, 1998.
- 28.MARTÍN, Y. G.; PAVÓN, J. L. P.; CORDERO, B. M.; PINTO, C. G. Classification of vegetable oils by linear discriminant analysis of Electronic Nose data. *Analytica Chimica Acta*, v.384, p.83-94, 1999.
- 29.BARTHUS, R. C.; POPPI, R. J. Determination of the total unsaturation in vegetable oils by Fourier transform Raman spectroscopy and multivariate calibration. *Vibrational Spectroscopy*, v.29, p.99-105, 2001.
- 30.IÑÓN, F. A.; GARRIGUES, J. M.; GARRIGUES, S.; MOLINA, A.; de la GUARDIA, M. Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance-Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, v.489, p.59-75, 2003.
- 31.CHRISTY, A. A.; KASEMSUMRAN, S.; DU, Y; OZAKI, Y. The detection and quantification of adulteration in olive oil by Near-Infrared Spectroscopy and Chemometrics. *Analytical Sciences*, v.20, p.935-940, 2004.
- 32.DIAZ, T. G.; MERÁS, I. D.; CASAS, J. S.; FRANCO, M. F. A. Characterization of virgin olive oils according to its triglycerides and sterols composition by chemometric methods. *Food Control*, v.16, p.339-347, 2005.
- 33.OSAWA, C. C.; GONÇALVES, L. Ap. G. Titulação potenciométrica aplicada na determinação de ácidos graxos livres de óleos e gorduras comestíveis. *Quím. Nova*, v.29, p.593-599, 2006.

- 
34. CHRISTY A. A.; EGEBERG, Per K. Quantitative determination of saturated and unsaturated fatty acids in edible oils by infrared spectroscopy and chemometrics. *Chemometrics and Intelligent Laboratory Systems*, v.82, p.130-136, 2006.
  35. GIACOMELLI, L. M.; MATTEA, M.; CEBALLOS, C. D. Analysis and Characterization of edible oils by chemometric methods. *J. Am. Oil Chem. Soc.*, v.83, p.303-308, 2006.
  36. ARMENTA, S.; GARRIGUES, S.; de la GUARDIA, M. Determination of edible oil parameters by near infrared spectrometry. *Analytica Chimica Acta*, v.596, p.330-337, 2007.
  37. BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M.B. Chemometrics: A Practical Guide. John Wiley & Sons, New York, 1998.
  38. PASQUINI, C. Near Infrared Spectroscopy: Fundamentals, practical aspects and analytical applications. *J. Brazilian Chemical Society*, v.14 (2), p.198-219, 2003.
  39. MASSART, D. L.; VANDEGINSTE, B. G. M.; BUYDENS, S. J.; LEWI, P. J.; SMEYERS-VERBEKE, J., Handbook of Chemometrics and Qualimetrics: Parte B, Elsevier, Amsterdam, 1997.
  40. SMALL, G. W. Chemometrics and Near Infrared Spectroscopy: Avoiding the Pitfalls. *Trends and Analytical Chemistry*, v.25 (11), p.1057-1066, 2006.
  41. CEN, H; he, y. Theory and Application of Near Infrared Reflectance Spectroscopy in Determination of Food Quality. *Trends in Food Science & Technology*, v.18 (2), p.72-83, 2007.
  42. SKOOG, D. A.; LEARY, J. J. Principles of instrumental analysis. Fourth Edition, Saunders College, 1992.
  43. BLANCO, M.; VILLARROYA, I. NIR Spectroscopy: a rapid-response analytical tool. *Trend. Anal. Chem.*, v.21 (4), p.240-250, 2002.
  44. BOKOBZA, L. Near Infrared Spectroscopy. *J. Near Infrared Spectrosc.*, v.6, p.3-17, 1998.
  45. WORKMAN Jr., J. J. Interpretative spectroscopy for near infrared. *Appl. Spectrosc.*, rev.31 (3), p.251-320, 1996.

- 
46. FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Quimiometria I: Calibração multivariada, um tutorial, *Quím. Nova*, v.22, p.724-731, 1999.
47. BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. 25 anos de Quimiometria no Brasil. *Quím. Nova*, v.29, p.1401-1406, 2006.
48. GALVÃO, R. K. H.; ARAUJO, M. C. U.; SALDANHA, T. C. B.; VISANI, V.; PIMENTEL, M. F. Estudo comparativo sobre filtragem de sinais instrumentais usando transformada de Fourier e de Wavelet. *Quím. Nova*, v.24 (6), p.874-884, 2001.
49. GALVÃO, R. K. H.; JOSÉ, G. E.; DANTAS FILHO, H. A.; ARAÚJO, M. C. U.; SILVA, E. C.; PAIVA, H. M.; SALDANHA, T. C. B.; SOUZA, E. S. O. N. Optimal wavelet filter construction using X and Y data. *Chemometrics and Intelligent Laboratory Systems*, V.70, P.1-10, 2004.
50. BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. Como fazer experimentos, 3ª ed., Editora Unicamp, Campinas – SP: 2007.
51. BRERETON, R. G. Introduction to multivariate calibration in analytical chemistry. *Analyst*, v.125, p.2125-2154, 2000.
52. ARAÚJO, M. C. U.; SALDANHA, T. C. B.; GALVÃO, R. K. H.; YONEYAMA, T.; CHAME, H. C.; VISANI, V. The successive projections algorithm for variable selection in spectroscopy multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems*, v.57 (2), p.65-73, 2001.
53. DANTAS FILHO, H. A.; GALVÃO, R. K. H.; ARAUJO, M. C. U.; SILVA, E. C.; SALDANHA, T. C. B.; JOSÉ, G. E.; PASQUINI, C.; RAIMUNDO JR., I. M.; ROHWEDDER, J. J. R. A Strategy for selecting calibration samples for multivariate modelling. *Chemometrics and Intelligent Laboratory Systems*, v.72, p.83-91, 2004.
54. JOUAN-RIMBAUD, D.; MASSART, D. L.; NOORD, O. E. Random correlation in variable selection for multivariate calibration with a genetic algorithm. *Chemometrics and Intelligent Laboratory Systems*, v.35, p.213-220, 1996.

- 
55. PONTES, M. J. C.; GALVÃO, R. K. H.; ARAUJO, M. C. U.; MOREIRA, P. N. T.; PESSOA NETO, O. D.; JOSÉ, G. E.; SALDANHA, T. C. B. The successive projections algorithm for spectral variable selection in classification problems. *Chemometrics and Intelligent Laboratory Systems*, v.78, p.11-18, 2005.
56. FEUDALE, R. N.; WOODY, N. A.; TAN, H.; MYLES, A. J.; BROWN, S. D.; FERRÉ, J. Transfer of multivariate calibration models: A review. *Chemometrics and Intelligent Laboratory Systems*, v.64, p.181-192, 2002.
57. HONORATO, F. A.; GALVÃO, R. K. H.; PIMENTEL, M. F.; BARROS NETO, B.; ARAÚJO, M. C. U.; CARVALHO, F. R. Robust modeling for multivariate calibration transfer by the successive projections algorithm. *Chemometrics and Intelligent Laboratory Systems*, v.76, p.65-72, 2005.
58. THOMAS, E. V.; HAALAND, D. M., Comparison of multivariate calibration methods for quantitative spectral analysis, *Anal. Chem.*, v.62, p.1091-1099, 1990.
59. GELADI, P.; KOWALSKY, B. R. Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, v.185, p.1-17, 1986.
60. OLIVEIRA, F. C. C.; SOUZA, A. T. P. C.; DIAS, J. A.; DIAS, S. C. L.; RUBIM, J. C. A escolha da faixa espectral no uso combinado de métodos espectroscópicos e quimiométricos. *Quím. Nova*, v.27, p.218-225, 2004.
61. COSTA FILHO, P. A.; POPPI, R. J. Algoritmo genético em química. *Quím. Nova*, v.22, p.405-411, 1999.
62. NØRGAARD, L.; SAUDLAND, A.; WAGNER, J.; NIELSEN, J. P.; MUNCK, L.; ENGELSEN, S. B. Interval partial least-squares regression (i-PLS): A comparative chemometric study with an example from near infrared spectroscopy. *Applied Spectroscopy*, v.54 (3), p.413-419, 2000.

- 
63. BORIN, A.; POPPI, R. J. Application of mid spectroscopy and iPLS for the quantification of contaminants in lubricating oil. *Vibrational Spectroscopy*, v.37, p.27-32, 2005.
64. GALVÃO, R. K. H.; ARAÚJO, M. C. U.; SILVA, E. C.; JOSÉ, G. E.; SOARES, S. F. C.; PAIVA, H. M. Cross-validation for the selection of spectral variables using the successive projections algorithm. *J. Braz. Chem. Soc.*, v.18 (8), p.1580-1584, 2007.
65. GALVÃO, R. K. H.; ARAÚJO, M. C. U.; JOSÉ, G. E.; PONTES, M. J. C.; SILVA, E. C.; SALDANHA, T. C. B. A method for calibration and validation subset partitioning. *Talanta*, v.67, p.736-740, 2005.
66. MORRISON, R.; BOYD, R. Química Orgânica. 13ª edição (traduzida da 6ª ed. original), Fundação Calouste Gulbenkian, Lisboa: 1996.