

MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
*CAMPUS* CAMPO MOURÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA DE ALIMENTOS

LARISSA ROCHA DOS SANTOS

**MÉTODOS ALTERNATIVOS PARA ANÁLISE RÁPIDA DE  
PARÂMETROS DE QUALIDADE DA SOJA**

DISSERTAÇÃO

CAMPO MOURÃO

2017

LARISSA ROCHA DOS SANTOS

**MÉTODOS ALTERNATIVOS PARA ANÁLISE RÁPIDA DE  
PARÂMETROS DE QUALIDADE DA SOJA**

Dissertação apresentada ao Programa de Pós-Graduação em Tecnologia de Alimentos, da Universidade Tecnológica Federal do Paraná, como requisito parcial para obtenção do título de Mestre em Tecnologia de Alimentos.

Orientador: Prof. Dr. Paulo Henrique Março

CAMPO MOURÃO

2017

Dados Internacionais de Catalogação na Publicação

S237m

Santos, Larissa Rocha dos

Métodos alternativos para análise rápida de parâmetros de qualidade da soja. / Larissa Rocha dos Santos – 2017.  
70 f. : il. ; 30 cm.

Orientador: Prof. Dr. Paulo Henrique Março  
Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologia de Alimentos. Campo Mourão, 2017.  
Inclui bibliografias.

1. Soja. 2. Espectroscopia no infravermelho. 3. Métodos Estatísticos. 4. Alimentos – Dissertações. I. Março, Paulo Henrique, orient. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologia de Alimentos. III. Título.

CDD: 664

Biblioteca Câmpus Medianeira  
Fernanda Cristina Gazolla Bem dos Santos CRB: 9-1735



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA DE  
ALIMENTOS

**PPGTA** UTFPR  
Programa de Pós-Graduação  
em Tecnologia de Alimentos

---

**TERMO DE APROVAÇÃO**

MÉTODOS ALTERNATIVOS PARA ANÁLISE RÁPIDA DE PARÂMETROS DE  
QUALIDADE DA SOJA

Por:

**LARISSA ROCHA DOS SANTOS**

Essa dissertação foi apresentada às 09:00 horas, do dia 24 de fevereiro de 2017, como requisito para obtenção do título de Mestre em Tecnologia de Alimentos, Linha de Pesquisa Processos Tecnológicos na Indústria de Alimentos, Programa de Pós-Graduação em Tecnologia de Alimentos, da Universidade Tecnológica Federal do Paraná. A candidata foi arguida pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho APROVADO.

---

Prof. Dr. Paulo Henrique Março (Orientador - PPGTA)

---

Profa. Dra. Ailey Aparecida Coelho Tanamati (Membro Interno - UTFPR)

---

Profa. Dra. Fernanda Vitória Leimann (Membro Interno - PPGTA)

---

Prof. Dr. Rafael Carlos Eloy Dias (Membro Externo)

\* A via original com as assinaturas encontra-se na secretaria do programa.

## AGRADECIMENTOS

À Deus, pela minha vida e por me dar a oportunidade de estar findando mais esta etapa tão importante da minha vida.

Aos meus pais, Eduardo e Edna, meus maiores exemplos. Obrigada por cada incentivo, orientação e pelas orações em meu favor. A minha irmã Laisa e a todos os meus familiares que sempre estiveram presentes, ainda que à distância, e que de alguma forma, incentivaram-me na constante busca pelo conhecimento.

Agradeço ao meu orientador, Professor Paulo Henrique Março, pela receptividade, paciência, confiança, incentivo e amizade, exemplo profissional e pessoa excepcional, pelos ensinamentos a mim transmitidos que muitas vezes transcenderam as exigências do seu papel e da sua responsabilidade, que contribuíram e muito para minha formação.

À Professora Patrícia Valderrama, que apesar de não ser oficialmente minha coorientadora, colaborou como tal. Agradeço pelo apoio, pela contribuição intelectual e por me ajudar em todas as dúvidas na realização da Dissertação.

Aos professores do Programa de Pós-Graduação em Tecnologia de Alimentos, pelos ensinamentos transmitidos e contribuição na minha formação acadêmica.

Às alunas de iniciação científica Marcela Zangirolami e Núbia Oliveira, por me ajudarem sempre que solicitadas durante a realização das análises.

Agradeço imensamente aos amigos que o mestrado me proporcionou conhecer, Melina Almeida, Jéssica Thaís, Fernanda Rúbio, Amanda Salgado, Alexandre Inácio, Carlos Koslinski, Fernanda Lansa, Gustavo Makimori, Luiza Mariano, Tatiane Vieira, Valquíria Maeda, Priscila Freitas e Thaysa Moya. Todos em algum momento me ajudaram, escutaram, incentivaram, com certeza fizeram toda a diferença nos meus momentos difíceis e deixou meus dias mais felizes. Obrigada pelas cervejinhas brindadas, pelas gordices divididas, caronas oferecidas, enfim por todos momentos maravilhosos que passamos juntos.

Aos meus queridos amigos Mariana Oliveira, Fernanda Menechine, Franciele Leila, Vandrielle Franco, Grace Kelly, Carolina Benedicti e vizinhos Rodrigo Barozi, Aline Perbeline e Tamires Barlati, que sempre estiveram por perto,

dispostos a me ajudar, ouvindo minhas angústias e dividindo momentos felizes, com vocês pude desfrutar momentos de descontração, aprendizado, motivação e amizade, além de dividir momentos de incertezas, dúvidas e tristezas. Obrigada por torcerem por mim e me incentivarem não só na vida profissional, mas em todos os assuntos.

Ao meu namorado Guilherme Sanchez por todo incentivo, apoio, paciência e compreensão durante esse período.

À Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) pelo auxílio financeiro.

Ao CNPq (processo 485640/2013-9) pelo financiamento do equipamento utilizado no desenvolvimento deste trabalho.

Aos membros participantes da banca pela disposição e considerações.

Enfim, a todos que de alguma forma contribuíram para a conclusão deste trabalho.

“O que me preocupa não é nem o grito dos corruptos, dos violentos, dos  
desonestos, dos sem caráter, dos sem ética...  
O que me preocupa é o silêncio dos bons.”

*Martin Luther King Jr.*

## RESUMO

Dada a importância mundial da cultivar soja, é imprescindível a aplicação de metodologias para o monitoramento eficiente dos parâmetros físico-químicos que determinam a qualidade dos grãos com agilidade e confiabilidade adequadas. Entretanto, os métodos analíticos empregados para as análises tradicionais envolvem técnicas demoradas, utilizam vários equipamentos e reagentes, além de gerarem resíduos químicos. Desta forma, o desenvolvimento de metodologias alternativas para esta finalidade pode trazer benefícios tanto para as indústrias e órgãos reguladores quanto para os analistas. Este estudo propõe a utilização de Espectroscopia de Infravermelho Próximo (NIR) associada a métodos quimiométricos para a construção de modelos multivariados para previsão do percentual de lipídios totais, índice de acidez, teor de clorofila, proteína bruta e umidade em soja. Na construção dos modelos foram avaliadas 300 amostras de soja *Glycine max* (L.) Merrill. Os dados espectrais foram processados por meio do método de Mínimos Quadrados Parciais (PLS). Os resultados sugerem que os modelos desenvolvidos podem ser utilizados como uma metodologia alternativa para determinar parâmetros físico-químicos e poderiam ser aplicados no controle de qualidade em indústrias de soja.

**Palavras-chave:** Espectroscopia no infravermelho; *Partial Least Square*; Métodos quimiométricos.



## ABSTRACT

Given the worldwide importance of soybean cultivars, it is essential to apply methodologies for the efficient monitoring of the physico-chemical parameters that determine the grain quality with adequate agility and reliability. Nonetheless, the analytical methods used in the traditional analysis involves time-consuming techniques, usage of various equipment and reagents besides generating chemical residues. Considering that, the development of alternative methodologies for this purpose can bring benefits to both industries and regulatory bodies as for the analysts. This study proposes the use of Near Infrared Spectroscopy (NIR) associated with chemometric methods for the construction of multivariate models to predict the percentage of total lipids, acidity index, chlorophyll content, crude protein and moisture in soybean. For this, 300 samples of Glycine max (L.) Merrill soybean were evaluated. The spectral data were processed by the method of Partial Least Squares (PLS). The results suggest that the developed model can be used as an alternative methodology to determine the physical-chemical parameters and could be applied in quality control in the soybean industries.

**Keywords:** Infrared spectroscopy; Partial Least Square; Chemometric methods.

## LISTA DE FIGURAS

<b>Figura 1:</b> Construção da matriz <b>X</b> para calibração multivariada (adaptado de VALDERRAMA, 2005). .....	41
<b>Figura 2:</b> Espectros de infravermelho próximo obtidos a partir de soja triturada (A) sem pré-processamento e (B) após aplicação de primeira derivada.....	45
<b>Figura 3:</b> Amostras anômalas para o modelo PLS de índice de acidez. ....	46
<b>Figura 4:</b> (A) Ajuste do modelo para índice de acidez; (B) Ajuste do modelo para lipídios totais; (C) Ajuste do modelo para proteína bruta e (D) Ajuste do modelo para teor de umidade. (x) Amostras de Calibração (●) Amostras de Validação. 50	50
<b>Figura 5:</b> Resíduos do modelo PLS para determinação de (A) índice de acidez; (B) lipídios totais; (C) proteína bruta e (D) umidade. (x) Amostras de Calibração (●) Amostras de Validação. ....	51
<b>Figura 6:</b> (A) Espectros NIR das amostras de soja sem pré-tratamento e (B) Espectros NIR pré-processados com correção de espalhamento multiplicativo e alisados pela aplicação do algoritmo de <i>Savitski-Golay</i> . ....	53
<b>Figura 7:</b> (A) Ajuste e (B) Resíduos do modelo PLS para determinação do teor de clorofila. (x) Amostras de Calibração (●) Amostras de Validação. ....	55

## LISTA DE TABELAS

<b>Tabela 1:</b> Variações nos valores obtidos pelos métodos de referência para amostras de soja trituradas .....	44
<b>Tabela 2:</b> Conjuntos de amostras utilizados para a construção dos modelos e outliers detectados e removidos para cada parâmetro.....	46
<b>Tabela 3:</b> Figuras de mérito do modelo PLS para os parâmetros físico-químicos da soja.....	48
<b>Tabela 4:</b> Figuras de mérito do modelo PLS para o parâmetro teor de clorofila da soja.....	54

## LISTA DE ABREVIATURAS E SIGLAS

CONAB	Companhia Nacional de Abastecimento
CV	Validação Cruzada (do inglês, <i>Cross Validation</i> )
J	Joule
MLR	Regressão Linear Múltipla (do inglês, <i>Multiple Linear Regression</i> )
MSC	Efeito Espalhamento Multiplicativo (do inglês, <i>Multiplicative Scatter Correction</i> )
N <sub>2</sub>	Nitrogênio
NaOH	Hidróxido de Sódio
NIR	Espectroscopia no Infravermelho Próximo (do inglês, <i>Near Infrared Spectroscopy</i> )
NM	Nanômetro
NO <sub>2</sub>	Óxido Nítrico
PCA	Análise de Componentes Principais (do inglês, <i>Principal Component</i> )
PCR	Regressão por Componentes principais (do inglês, <i>Principal Components Regression</i> )
PLS	Mínimos Quadrados Parciais (do inglês, <i>Partial Least Squares</i> )
R <sup>2</sup>	Coeficiente de Determinação
RMSEC	Raiz Quadrada do Erro Médio Quadrático de Calibração (do inglês, <i>Root mean Square Error of Calibration</i> )
RMSECV	Raiz Quadrada do Erro Médio Quadrático de Validação Cruzada (do inglês, <i>Root mean Square Error of Cross Validation</i> )
RMSEP	Raiz Quadrada do Erro Médio Quadrático de Predição (do inglês, <i>Root mean Square Error of Prediction</i> )
RPD	Relação de Predição de Desvio (do inglês, <i>Residual Prediction Deviation</i> )
SD	Desvio Padrão (do inglês, <i>Standard Deviation</i> )
SEP	Erro Padrão de Predição (do inglês, <i>Standard Error of Prediction</i> )

## SUMÁRIO

1. INTRODUÇÃO.....	15
2. OBJETIVOS.....	18
2.1 OBJETIVO GERAL.....	18
2.2 OBJETIVOS ESPECÍFICOS .....	18
3. REVISÃO BIBLIOGRÁFICA .....	19
3.1 SOJA .....	19
3.2 CONTROLE DE QUALIDADE DOS GRÃOS .....	20
3.2.1 Índice de acidez.....	20
3.2.2 Umidade .....	21
3.2.3 Lipídios totais.....	22
3.2.4 Proteína .....	22
3.2.5 Teor de clorofila.....	23
3.3 ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO .....	24
3.3.1 Métodos de medição .....	25
3.3.2 Vantagens e desvantagens do método NIR .....	26
3.4 QUIMIOMETRIA .....	27
3.4.1 Calibração Multivariada .....	27
3.4.2 Mínimos Quadrados Parciais (PLS).....	28
3.4.3 Detecção de amostras anômalas – <i>Outliers</i> .....	30
3.5 VALIDAÇÃO ANALÍTICA E FIGURAS DE MÉRITO.....	31
3.5.1 Exatidão.....	31
3.5.2 Linearidade.....	32
3.5.3 Coeficiente de correlação (R) .....	32
3.5.4 Precisão.....	32
3.5.5 Ajuste.....	33
3.5.6 Relação de Predição do Desvio (RPD).....	33
3.5.7 Intervalo de confiança.....	34
3.5.8 Extensão da faixa de trabalho .....	34
3.5.9 Sensibilidade e sensibilidade analítica .....	34
3.5.10 Robustez .....	35
3.5.11 Limite de detecção e quantificação.....	35
4. MATERIAL E MÉTODOS .....	36

4.1	PREPARO DAS AMOSTRAS.....	36
4.2	MÉTODOS DE REFERÊNCIA .....	36
4.2.1	Determinação do índice de acidez.....	36
4.2.2	Determinação de umidade.....	37
4.2.3	Determinação do teor de lipídios totais.....	38
4.2.4	Determinação de proteína bruta .....	39
4.2.5	Determinação do teor de clorofila .....	39
4.3	ANÁLISE ESPECTROSCÓPICA NA REGIÃO DO INFRAVERMELHO PRÓXIMO .....	40
4.4	PROCESSAMENTO DOS DADOS .....	40
5.	RESULTADOS E DISCUSSÕES.....	44
6.	CONCLUSÕES.....	57
7.	REFERÊNCIAS .....	58

## 1. INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill), tem grande importância no cenário mercadológico não apenas brasileiro mas também mundial (CAVALCANTE; DE SOUSA; HAMAWAKI, 2011; KLAHOLD et al., 2006), muito graças a versatilidade de aplicação de seus produtos na alimentação humana e animal e ao seu valor econômico. O Brasil está entre os maiores produtores de soja do mundo, sendo esta leguminosa cultivada em várias regiões do país (HIRAKURI; LAZZAROTTO, 2014; MELLO FILHO et al., 2004).

A comercialização de derivados da soja demanda um rigoroso controle dos grãos e dos padrões de comercialização (ARAUJO, 2014). Entretanto, procedimentos analíticos tradicionais utilizados no controle de qualidade são relativamente demorados, dispendiosos, de custo significativo, necessitam de mão de obra qualificada e apresentam caráter destrutivo das amostras, além de gerarem resíduos químicos. Portanto, o desenvolvimento de metodologias que possam aprimorar aspectos analíticos como a velocidade das respostas e aspectos não-destrutivos são requeridos (MA et al., 2017).

Em busca de métodos de rotina rápidos para a quantificação de várias substâncias em amostras de alimentos, os métodos ópticos vem ganhando cada vez mais popularidade (BEVILACQUA et al., 2013). O uso de métodos ópticos apresenta vantagens tais como baixo custo a médio-longo prazo, simplicidade operacional, não destruição das amostras analisadas, além de oferecer análises rápidas, sem necessidade (ou com um mínimo) de preparo de amostras, apresentando caráter não invasivo e não gerarem resíduos (PASQUINI, 2003; PAULA et al., 2013).

Dentre os métodos ópticos destaca-se a espectroscopia na região do infravermelho, que se baseia na absorção de energia radiante, resultando em vibração molecular (PAULA et al., 2013). Um espectro, diferentemente das análises convencionais, resulta em um vetor de resposta, enquanto as análises convencionais apresentam um valor escalar para cada parâmetro. Desta forma, para avaliar os resultados obtidos a partir de métodos ópticos, a forma mais viável é por meio de métodos multivariados de análise, mais conhecidos como

métodos quimiométricos. Estes, são provenientes de métodos matemáticos e estatísticos aplicados com o intuito de se extrair informações não triviais a partir de um conjunto de dados complexos (MARÇO, 2009; NETO; SCARMINIO; BRUNS, 2006; VALDERRAMA, 2009).

O tipo de modelo para calibração multivariada depende da estrutura dos dados. Por exemplo, no caso de estruturas lineares, o método quimiométrico mais utilizado é a regressão por Mínimos Quadrados Parciais (PLS, do inglês *Partial Least Squares*).

Desta forma, a Espectroscopia no Infravermelho associada a métodos quimiométricos pode ser indicada como um meio rápido e tão confiável quanto o método de referência no controle da qualidade de alimentos (MA et al., 2017; RESTAINO; FASSIO; COZZOLINO, 2011). Esta estratégia tem sido aplicada a uma grande diversidade de estudos envolvendo análise de grãos, principalmente no setor agrícola, devido à relativa facilidade de implementação desta tecnologia de baixo custo em um ambiente industrial, seja por um sistema *on-line*, em laboratórios de controle de qualidade ou em linha de produção (HAUGHEY et al., 2013).

Aplicações ópticas para avaliação de grãos comestíveis utilizando a região do infravermelho próximo vêm ganhando mais espaço a cada dia. Por exemplo, testes de grãos em laboratório por Espectroscopia no Infravermelho Próximo (NIR, do inglês *Near Infrared*), foram originalmente propostos para a determinação do conteúdo da matéria seca e proteína em cereais e sementes oleaginosas (MASSIE; NORRIS, 1965). Devido ao seu potencial para avaliação rápida de atributos de qualidade de alimentos, o NIR já substituiu diversos ensaios de rotina em laboratórios agrícolas, como por exemplo, as análises de solo realizadas pela EMBRAPA Solos, a qual avalia os parâmetros do solo relacionados com a fertilidade utilizando-se o infravermelho médio com calibração multivariada (DIAS, 2016).

Diversos autores estudaram a possibilidade de medidas diretas de proteína, carboidrato, lipídios, umidade, peso seco e amido de diversos grãos através de Espectroscopia de Infravermelho Próximo (HUANG et al., 2008). Welle e seus colaboradores (2003) e Montes et al. (2006) mediram teor de matéria seca e parâmetros de qualidade do milho com Espectroscopia no Infravermelho Próximo acoplado diretamente na máquina de colheita. Long et al.



(2008), descreveram uma abordagem para inspecionar a proteína do grão de trigo *on-line* com o uso de análise de NIR. Brás et al. (2005), estudaram o uso do Infravermelho Médio e Próximo para o desenvolvimento de calibração multivariada para prever propriedades de qualidade de farinha de soja, como teor de proteína e umidade. Heman e Hsieh (2016), mediram o teor de umidade de arroz em casca por Espectroscopia Visível e do Infravermelho Próximo. Ferreira e seus colaboradores (2014), pesquisaram a composição centesimal de soja por meio do NIR associado a métodos de calibração multivariada. Com isso, podemos constatar que o infravermelho próximo representa uma técnica analítica emergente, que está desfrutando de crescente popularidade na indústria de alimentos (POREP; KAMMERER; CARLE, 2015).

Muitos modelos multivariados propostos para a análise de alimentos não são validados para uso em análise de rotina, uma vez que a análise das figuras de mérito destes, pode sugerir que os modelos produzidos não devam ser aplicados (BOTELHO; MENDES; SENA, 2013). Na maioria dos casos a inaptidão dos modelos está relacionada com o processo amostral, no qual a quantidade de amostra utilizada para a construção dos modelos é inadequada, não refletindo a situação real de demanda em um laboratório de análise de rotina.

Neste trabalho, foram desenvolvidos modelos de calibração multivariada para previsão de parâmetros físico-químicos em amostras de soja destinadas à produção de óleo e farelo. Nesta proposta, todos os modelos construídos foram validados e, considerando-se os modelos observados na maioria dos trabalhos citados neste estudo, empregou-se um número bastante elevado de amostras.

## 2. OBJETIVOS

### 2.1 OBJETIVO GERAL

Este trabalho teve como objetivo geral oferecer uma metodologia alternativa para determinação rápida e não invasiva de teores de clorofila, lipídios totais, proteína bruta, índice de acidez e umidade em amostras de grãos de soja triturados, baseados em espectros ópticos no Infravermelho Próximo e calibração multivariada para análise dos dados.

### 2.2 OBJETIVOS ESPECÍFICOS

- Avaliar por meio dos métodos de referência os parâmetros índice de acidez, umidade, lipídios totais, proteína bruta e clorofila em amostras de soja trituradas;
- Adquirir espectros de infravermelho próximo (NIR) por reflectância difusa das amostras de soja trituradas;
- Desenvolver modelos de calibração multivariada (Mínimos Quadrados Parciais - PLS) a partir dos espectros obtidos por NIR, correlacionando com os parâmetros obtidos por metodologias oficiais (AOCS - *American Oil Chemists' Society*);
- Validar os métodos de PLS por meio do cálculo e avaliação das figuras de mérito.

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1 SOJA

A soja é uma semente oleaginosa pertencente à família *Leguminosae*, subfamília *Papilionoidae* e ao gênero *Glycyne*, L. A forma cultivada *Glycine max* (L.) Merrill, é uma das mais importantes leguminosas do Brasil e do mundo, sendo de grande expressão econômica (CUNHA et al., 2015; MANDARINO, 2012).

Originária do continente asiático, mais precisamente da China, há pelo menos cinco mil anos, sendo considerada umas das culturas mais antigas da civilização. Juntamente com o arroz, o trigo, o centeio e o milheto, foi tida como um dos cinco grãos sagrados (PAIVA; ALVES; HELENO, 2006; PAULETTO; FOGAÇA, 2012). Somente no século XX passou a ser cultivada nos Estados Unidos, sendo introduzida no Brasil por volta de 1900 (PAULA, 2007).

Vários são os fatores que podem influenciar na constituição do grão de soja, entre eles, a época do plantio, tipo de cultivar, localização geográfica e condições climáticas (GARCÍA-REBOLLAR et al., 2016). Suas vagens contêm em torno de 3 a 4 sementes, sendo formada basicamente de 40% de proteína, 20% de lipídios, 5% de minerais e 34% de carboidratos (CUNHA et al., 2015; GONÇALVES et al., 2014).

O Brasil está entre os maiores produtores de soja do mundo, sendo a leguminosa cultivada em várias regiões do país, atrás apenas dos Estados Unidos que ocupa o primeiro lugar (CUNHA et al., 2015; HIRAKURI; LAZZAROTTO, 2014). É uma das mais importantes *commodities* brasileira, responsável por mais de 56% da área cultivada do país (CONAB, 2016).

Para a safra 2016/17, estima-se uma produção de 104,6 milhões de toneladas de soja em grãos. O consumo interno total deve chegar a 45,9 milhões de ton, e aproximadamente 58,7 milhões de toneladas devem ser exportadas, sendo a China uma das principais importadoras da soja brasileira (CONAB, 2016).

### 3.2 CONTROLE DE QUALIDADE DOS GRÃOS

A soja domina o mercado mundial de proteína vegetal e de óleo comestível. O óleo de soja surgiu como um subproduto do processamento do farelo de soja, e tornou-se um dos líderes mundiais no comércio de óleos (MACIEL et al., 2016). Logo, a qualidade dos grãos é um parâmetro importante não só para os produtores mas também para a indústria, sendo que em muitos países o preço do grão é determinado pelas suas características físico-químicas (HUANG et al., 2008).

Devido à importância econômica dessa matéria-prima, faz-se necessário um rigoroso controle de qualidade dos grãos de soja, como os teores de proteína bruta, lipídios totais, clorofila, umidade e índice de acidez que estão diretamente correlacionados com as características do óleo e farelo de soja produzido (GREGGIO; BONINI, 2014).

#### 3.2.1 Índice de acidez

Depois de colhidos, os grãos de soja são armazenados até o processo de industrialização. Entretanto, os grãos podem sofrer alterações em sua composição em razão do ambiente de armazenagem (GREGGIO; BONINI, 2014). A estocagem em silos comuns pode comprometer a integridade dos grãos devido a processos oxidativos e ações enzimáticas diversas, o que se intensifica dependendo do tempo de estocagem: quanto maior o tempo de armazenagem, maior será o índice de acidez. A acidez elevada gera maior custo do processo de refino do óleo degomado, ocasionado por uma maior adição de insumos e uma maior perda de matéria prima (óleo) (BORDIGNON, 2009).

O índice de acidez de um grão de soja é um dos principais parâmetros observados em uma indústria de óleo, visto que este está relacionado com a qualidade e valor agregado do produto final. Este índice evidencia o estado de conservação do óleo presente no grão. Sabe-se que a luz e o aquecimento aceleram a decomposição dos triacilgliceróis, e a rancidez é quase sempre acompanhada pela formação de ácidos graxos livres (SOARES; BIAGGIONI; FRANÇA NETO, 2005).

Durante a formação até a fase de maturação fisiológica dos grãos de soja, esse índice apresenta-se em torno de 0,3% a 0,5%. Quando os grãos estão em condições de colheita, o processo degradativo se inicia. Consideram-se toleráveis níveis de até 0,7% de ácidos graxos livres, uma vez que a tolerância do mercado de óleo de soja estabelece um padrão de no máximo 0,05%. Logo, quando os grãos apresentam índices de acidez elevados, necessitam de neutralização, processo no qual adicionam-se produtos de custo significativo ao processamento de óleo vegetal (OLIVEIRA et al., 2016).

### 3.2.2 Umidade

Em sementes oleaginosas, outro parâmetro de controle de qualidade fundamental é a umidade, já que é conhecido que a estabilidade desses alimentos diminui com o aumento do teor de umidade (KAIJSER; DUTTA; SAVAGE, 2000).

A recomendação técnica é de que a colheita ocorra quando a umidade estiver em torno de 12-14%, e que durante a armazenagem, devem ser controlados, tanto a umidade do próprio grão, quanto a temperatura e umidade relativa do ambiente (SINNECKER, 2002).

Grãos e/ou sementes oleaginosas possuem a propriedade de ganhar ou perder umidade, por equilíbrio higroscópico, devido às condições de armazenamento. O equilíbrio higroscópico acontece quando em uma mesma temperatura, a umidade dos grãos entra em equilíbrio com a umidade relativa do ambiente (BORDIGNON, 2009).

Grãos com alto teor de água, propiciam condições favoráveis para a infestação de ácaros, fungos e insetos-praga, além de influenciar no metabolismo dos grãos, propiciando variações físico-químicas e fisiológicas (SOARES; BIAGGIONI; FRANÇA NETO, 2005).

Esses fatores em conjunto com o estresse térmico e danos mecânicos, corroboram para a diminuição do percentual de gordura, proteína, carboidratos, aumentam o teor de acidez e consomem matéria seca (GREGGIO; BONINI, 2014).

### 3.2.3 Lipídios totais

A soja é uma oleaginosa que possui considerável teor de óleo (em torno de 20%). Do grão é possível extrair um óleo comestível de boa qualidade e, por isso, a análise de lipídios totais se faz necessária e indispensável, já que a partir desta análise se gerencia o processo industrial para obtenção do óleo (GREGGIO; BONINI, 2014).

Condições climáticas, região e tipo de cultivar são fatores que influenciam na composição lipídica do óleo de soja. Geralmente, o grão de soja contém cerca de 48 até 59% de ácido linoleico (18:2), de 17 a 30% de ácido oleico (18:1), e em torno de 3,5 a 8% de ácido linolênico (18:3). O alto teor de ácidos graxos insaturados tornam o óleo de soja mais susceptível a oxidação. Os ácidos graxos saturados majoritários são: ácido palmítico (16:0), presente em teores que variam de 9,90 até 12,2% e de 3 até 5,4% de ácido esteárico (18:0) (VICARI, 2013).

Os lipídios dos grãos são passíveis de deterioração quando armazenados sem os devidos cuidados, gerando prejuízos para indústria. Por isso se faz necessário avaliar a qualidade da matéria-prima e do óleo bruto, uma vez que estes influenciarão na qualidade do produto final, o óleo de soja refinado (FARONI et al., 2009).

### 3.2.4 Proteína

A soja é classificada principalmente pelo seu alto teor de proteína, sendo considerada a principal fonte de proteína vegetal da alimentação animal, e muito utilizada em alimentos para humanos, além de ser um alimento de valor energético expressivo, com proteína de alta qualidade (ZARKADAS et al., 2007).

Proteínas são macromoléculas orgânicas constituídas por uma ou mais cadeias de aminoácidos, que possuem um número considerável de sítios de sorção de água (BORDIGNON, 2009). Na soja, suas proteínas possuem um equilibrado balanço de aminoácidos essenciais: isoleucina, leucina, lisina, metionina, fenilalanina, treonina e triptofano (DOS SANTOS et al., 2005). Grãos

de soja contém cerca de 40% de proteínas em base seca, divididas em quatro grupos: as enzimas envolvidas no metabolismo; proteínas estruturais, incluindo as ribossomal e cromossômica; proteínas da membrana; e as proteínas de armazenamento. As proteínas presentes em maior quantidade são a proteína de reserva glicínica e  $\beta$ -conglícinina, lipoxigenases, inibidores de tripsina Kunitz, inibidores de protease de baixo peso molecular, lectina e uréase (DE MORAES et al., 2006; ZARKADAS et al., 2007).

Variações nos teores de proteína são determinados por fatores genéticos e ambientais. A temperatura, a disponibilidade hídrica e o suprimento de nitrogênio, principalmente no período de enchimento dos grãos, são exemplos de fatores ambientais. É necessário ter controle do teor de proteína da soja e observar o seu comportamento durante o armazenamento, uma vez que esse parâmetro é determinante para a constituição do farelo de soja, o qual é determinante para seu valor comercial (PÍPOLO et al., 2015).

### 3.2.5 Teor de clorofila

A coloração verde está ligada à presença de clorofila em grãos colhidos ainda imaturos, e isso pode interferir na qualidade do óleo extraído. Outro aspecto a ser considerado é que grãos imaturos apresentam percentual de 2% a 3% menor de óleo em relação ao grão maduro, o que acarretará em uma maior acidez, além de ter um custo maior de refino, conseqüentemente reduzindo o seu valor comercial (FRANÇA-NETO et al., 2012; MANDARINO, 2012).

Devido à importância econômica da soja, se faz necessário um rigoroso controle dos grãos e dos padrões de comercialização. Entretanto, a indústria frequentemente tem problemas relacionados ao recebimento de grãos esverdeados (ARAUJO, 2014). Por isso é importante monitorar a presença de semente esverdeada durante a colheita, o ideal é que esse nível seja menor que 0,09% (KRZYŻANOWSKI et al., 2008). Porém, a Instrução Normativa Nº 11, de 15 de Maio de 2007 estabelece como limite máximo 8% de grãos esverdeados, entretanto caso a soja for pra exportação, esse limite irá variar de acordo com o país importador, tendo como exemplo a legislação dos EUA que permite apenas 1% de grãos verdes na soja (NUNES, 2013).

Clorofilas são pigmentos naturais presentes nas plantas, nos cloroplastos das folhas e em outros tecidos vegetais fotossintéticos (LI et al., 2010; STREIT et al., 2005). Sua incidência em sementes de soja se deve a vários fatores como genótipo, estágio de maturação, condições climáticas, condições de estresse (oscilações drásticas de temperatura, ocorrência de geadas intensas, doenças que afetam as raízes, hastes e folhas da planta), déficit hídrico durante o estágio reprodutivo, presença de pragas, entre outros (MANDARINO, 2012; SINNECKER, 2002).

Assim, óleos extraídos de grãos oleaginosos verdes apresentam elevado teor de clorofila, responsável por conferir ao produto uma coloração escura, comercialmente inadequada. Além de visualmente indesejável, por ser uma molécula pró-oxidante a clorofila diminuí a estabilidade oxidativa do óleo, levando à sua rancificação, o que reduz a eficiência do processo de hidrogenação, além de demandarem maior volume de agentes clarificantes (BORRMANN, 2009; SINNECKER, 2002).

### 3.3 ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO

Em 1800, o astrônomo alemão Frederick William Herschel (1738-1822) descobriu a radiação infravermelha pela passagem da luz solar através de um prisma (REICH, 2005). A Espectroscopia de Infravermelho Próximo recebeu esse nome por se tratar da região do espectro eletromagnético imediatamente superior à região visível, a “mais próxima” ao visível (350-750 nm) (LIMA; BAKKER, 2011). Esta região corresponde ao intervalo de comprimentos de onda que vão de 750 até 2500 nm (PASQUINI, 2003).

Entretanto, foi a partir de 1960 que a Espectroscopia de Infravermelho Próximo (NIR, do inglês *Near Infrared*), começou a ser aplicada de forma analítica (REICH, 2005), a qual ganha força principalmente para determinações quantitativas de produtos agrícolas e alimentares, e também nos mais variados setores industriais, como em indústrias de petróleo, têxtil, carvão, cosméticos, polímeros, química, tintas e farmacêuticas (ROGGO et al., 2007).

A espectroscopia de infravermelho é uma técnica baseada nas vibrações moleculares (HUANG et al., 2008), onde empregam-se fótons com energia



variante entre  $2,65 \times 10^{-19}$  a  $7,96 \times 10^{-20}$  J, que é suficiente para intensificar a vibração das ligações entre os átomos (PASQUINI, 2003). Esta forma de espectroscopia mede a absorção refletida ou transmitida da luz incidente sobre uma amostra em uma determinada região de comprimentos de onda (FIGUEIREDO, 2008). O espectro contém informação relacionada aos grupos funcionais, e estas absorções apresentam-se com alto grau de sobreposição, havendo a necessidade da aplicação de métodos quimiométricos para interpretação dos resultados (MA et al., 2017).

A frequência de uma vibração depende das massas relativas dos átomos ligados, das constantes de força das ligações e da geometria dos ligantes. Além disso, estão relacionadas com vibrações e ligações químicas e ao tamanho, peso e quantidade de átomos presentes em cada molécula. Desta forma, cada frequência de vibração pode ser associada a um grupo específico de ligações químicas (JAMRÓGIEWICZ, 2012).

### 3.3.1 Métodos de medição

A espectroscopia NIR pode ser realizada por reflectância difusa, transmitância, transflectância ou absorção. A escolha do modo de medição é determinada pelas características ópticas da amostra. Geralmente, materiais transparentes são medidos pelo modo de transmitância (MAGALHÃES, 2014), onde o princípio baseia-se na absorção da radiação pela amostra, e o decréscimo da intensidade do sinal captado pelo sensor/detector em uma determinada frequência sugere quais grupos podem estar absorvendo energia para vibrações. Através da diferença da energia emitida e a energia recebida pelo detector, consegue-se verificar as frequências que a amostra mais absorveu (LEITÃO, 2012).

Quando se trata de amostras sólidas (inteiras ou pó), o modo de reflectância difusa é o mais indicado. A radiação incide na amostra, espalha sobre a superfície, e penetra alguns milímetros na amostra, onde uma parte é absorvida e a radiação não absorvida é refletida pela amostra (SILVA, 2011).

As medidas de transflectância funcionam como uma combinação dos métodos de transmitância e reflectância. Esse modo de aquisição pode ser

empregado para amostras líquidas, viscosas e em suspensão. Os espectros são obtidos primeiramente por reflectância, onde a radiação incidente atravessa a amostra e, em seguida, é refletida, atravessando novamente a amostra e sendo então captada pelo detector, gerando um espectro de resposta (MAGALHÃES, 2014; SILVA, 2011).

### 3.3.2 Vantagens e desvantagens do método NIR

A técnica analítica de Espectroscopia de Infravermelho Próximo oferece vantagens como rapidez na obtenção dos dados, apresenta caráter não destrutivo e não invasivo (indicada para uso *in-process*) e por isso, frequentemente, não necessita de preparação prévia da amostra, o que gera economia de tempo e reagentes, minimizando erros do analista (por permitir a diminuição do número de medidas padrão) (HUANG et al., 2008; MA et al., 2017; PASQUINI, 2003; POREP; KAMMERER; CARLE, 2015).

Outro aspecto importante da técnica é que esta propicia uma análise multicomponente de uma grande variedade de matrizes, uma vez que as moléculas que compõem a amostra precisam apenas apresentar variação no momento dipolar para serem influenciadas por esta ferramenta e produzir o espectro (REICH, 2005).

A Espectroscopia de Infravermelho Próximo apresenta algumas desvantagens com relação à implementação por depender em alguns casos de métodos quimiométricos. Assim, por ser uma metodologia aplicada de forma indireta (via calibração multivariada), demanda comparação com método de referência, tendo sua confiabilidade relacionada com a calibração e com erros analíticos que esta apresenta, além de ser considerada como uma técnica de baixa sensibilidade, pois restringe a determinação de substâncias presentes em concentrações reduzidas (JAMRÓGIEWICZ, 2012; LUYPAERT; MASSART; VANDER HEYDEN, 2007).

A espectroscopia NIR é uma técnica interessante para a investigação científica e vem ganhando espaço entre as técnicas mais utilizadas, principalmente quando associadas a quimiometria (POREP; KAMMERER; CARLE, 2015).

### 3.4 QUIMIOMETRIA

A quimiometria é uma área da química analítica que se refere à aplicação de métodos matemáticos e estatísticos para planejar, otimizar ou selecionar procedimentos de medidas e experimentos, bem como extrair o máximo da informação química relevante, com a análise de dados químicos (JACONI, 2011; ROGGO et al., 2007). Estes procedimentos começaram a ser implementados formalmente em meados dos anos 70, consolidando-se com a chegada do computador nos laboratórios químicos (NETO; SCARMINIO; BRUNS, 2006). Segundo Jaconi (2011), a quimiometria surgiu devido à imprescindibilidade de se desenvolver novos métodos estatísticos e matemáticos para analisar a grande quantidade de dados produzidos por instrumentos analíticos comandados por computadores, permitindo a utilização da estatística multivariada para maior extração de informações de dados químicos.

#### 3.4.1 Calibração Multivariada

A calibração multivariada tem como princípio estabelecer uma relação quantitativa entre dois conjuntos de dados de medição. Dentre as possibilidades, a aplicação de calibração multivariada permite a utilização de espectros em lugar de metodologias convencionais (POREP; KAMMERER; CARLE, 2015). Além disso, vários parâmetros podem ser determinados simultaneamente a partir de um único espectro com o auxílio de ferramentas que proporcionam análises multivariadas (BEVILACQUA et al., 2013). Por exemplo, ferramentas para calibração multivariada proporcionam uma correlação entre uma matriz de dados (como um conjunto de espectros) e as concentrações de analitos determinadas por métodos de referência (FODOR et al., 2011).

Métodos de calibração multivariada abordam não apenas um ponto de dados espectrais para a calibração, mas toda a estrutura espectral. Assim, é possível que se tenha informação sobre múltiplas regiões da amostra, permitindo-se conclusões mais abrangentes (FODOR et al., 2011).

No entanto, a aplicação desta ferramenta envolve várias etapas, onde uma depende e influencia a outra. Entre elas, as importantes são: (i) seleção de

um conjunto de calibração e outro para validação; (ii) pré-tratamento dos dados; (iii) verificação da existência de representatividade entre os conjuntos de calibração e validação; (iv) investigação das heterogeneidades e de possíveis não-linearidades nos dados; (v) construção do modelo de calibração; (vi) detecção de *outliers* além da (vii) validação do modelo (JOUAN-RIMBAUD et al., 1999).

Assim como na calibração univariada, a calibração multivariada relaciona a variável dependente (**Y**) com a variável independente (**X**):

$$\mathbf{Y} = f(\mathbf{X}) = \mathbf{X} \cdot \mathbf{b} \quad (1)$$

Onde, **b** corresponde ao coeficiente de regressão do modelo. No entanto, ao contrário das análises univariadas – nas quais obtêm-se apenas um **b** – na calibração multivariada obtêm-se um índice **b** para cada comprimento de onda. As variáveis independentes são utilizadas em conjunto com os coeficientes de regressão para estimar valores para a variável dependente. Deste modo, utilizam-se amostras com variáveis dependentes conhecidas no conjunto de validação para comparar os valores previstos pelo modelo e os valores obtidos pela metodologia padrão, o que permitirá a avaliação de desempenho do modelo obtido (ARAUJO, 2007; HELLAND, 1988). Assim, mais de uma resposta instrumental é relacionada com a propriedade de interesse.

#### 3.4.2 Mínimos Quadrados Parciais (PLS)

O método de regressão por mínimos quadrados parciais (*Partial Least Squares* – PLS), foi introduzido em 1975 por Herman O. A. Wold no campo da econometria, sendo depois implementado por seu filho, Svante Wold na área de dados químicos, onde tornou-se uma ferramenta imprescindível em muitos campos da química e no controle de processos industriais (GELADI; KOWALSKI, 1986; TOSCAN, 2016). É reputado como o método de regressão mais aplicado para a construção de modelos de calibração multivariada de dados de primeira ordem (ARAUJO, 2007; BRERETON, 2000; VALDERRAMA, 2009).

O PLS é um método de calibração multivariada para dados que obedecem comportamento linear (ALBRECHT et al., 2008; JACONI, 2011). Uma de suas vantagens é que é considerado um método capaz de produzir modelos robustos. Assim, os parâmetros do modelo não sofrem alterações significativas quando novas amostras são acrescentadas ao conjunto de calibração, podendo ser aplicado em misturas complexas (GELADI; KOWALSKI, 1986). Desta forma, os modelos produzidos não demandam conhecimento de todos os componentes que integram a amostra, pois possibilita previsão de parâmetros da amostra mesmo na presença de interferentes, desde que estes estejam presentes na etapa de construção do modelo (ARAUJO, 2007; TOSCAN, 2016).

Este método fundamenta-se em correlacionar duas matrizes de dados, uma contendo as novas medidas, **X** (variáveis independentes), e outra com os valores da propriedade de interesse medida pelo método de referência, **Y** (variáveis dependentes) (ARAUJO, 2007; ERNESTINA; SABIN, 2007; ROGGO et al., 2007; TOSCAN, 2016). As matrizes **X** e **Y** são decompostas conforme as Equações a seguir (BRERETON, 2000):

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (2)$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{F} \quad (3)$$

Onde: **P** e **Q** são os pesos de **X** e **Y**, respectivamente; **T** e **U** são os escores de **X** e **Y**, nesta ordem e **E** e **F** correspondem aos erros, ou seja, o quanto o modelo perde de informação com a redução de variáveis em relação aos dados originais.

Os escores e pesos equivalem às coordenadas das amostras no sistema de componentes principais e a contribuição relativa de cada variável original para a formação destes. No modelo PLS, os valores dos escores nas matrizes **T** e **U** são modificados até que o melhor modelo linear seja estabelecido entre eles e, assim, obtém-se a maior covariância entre **X** e **Y**. Desta forma, realiza-se uma série de etapas que são repetidas para cada componente principal estimada por meio de um processo iterativo. Então, no PLS existe um compromisso entre a explicação da variância em **X** e a obtenção da maior correlação com **Y**, as

componentes principais passam assim a ser denominadas variáveis latentes (SOUZA et al., 2013; VALDERRAMA; BRAGA; POPPI, 2009).

O número de variáveis latentes do modelo é determinado por meio de uma validação cruzada (CV – do inglês, *Cross Validation*), baseado no menor erro de previsão, através do método de *contiguous blocks* (“blocos contínuos”). Neste método, 10 amostras de calibração são retiradas de cada vez e a propriedade de interesse é estimada pelo modelo construído com as amostras restantes; a média dos erros de previsão é calculada para cada variável latente, sendo que o número de variáveis latentes que apresentar menor erro de previsão é escolhido (SOUZA et al., 2013).

Nos conjuntos de calibração e validação do modelo, podem estar presentes amostras muito diferentes da média geral das amostras (amostras anômalas, também conhecidas como *outliers*). Este tipo de amostra apresenta divergências significativas, se comportando de forma discrepante das demais amostras do conjunto. Logo, é essencial a identificação e, em alguns casos, recomenda-se a remoção destes elementos durante as etapas de calibração e validação, afim de se construir modelos mais eficientes, precisos e com melhor capacidade de previsão (MARTENS; NAES, 1992; VALDERRAMA, 2005).

### 3.4.3 Detecção de amostras anômalas – *Outliers*

Amostras anômalas, também conhecidas como *outliers*, é o termo utilizado para definir amostras com comportamento diferente das demais amostras presentes nos conjuntos de calibração e de validação. Geralmente, a presença destas amostras propicia a construção de modelos com altos valores de erros, baixa capacidade de previsão, influenciando significativamente na previsão dos resultados. Deste modo, a detecção de *outliers* é de suma importância e, muitas vezes, a remoção destas amostras conduz à construção de modelos mais precisos, eficientes, assegurando a qualidade preditiva dos modelos (MARTENS; NAES, 1992; VALDERRAMA, 2009).

Existem algumas metodologias para a detecção de *outliers*, sendo as mais comuns baseadas na identificação de amostras com grande influência no modelo (*leverage*), altos resíduos não modelados nos dados espectrais e altos

resíduos na variável dependente (SILVA, 2011). O *leverage* indica o quão distante uma amostra está em relação à média do conjunto de dados, e significa o “peso relativo” que essa amostra exerce em relação as demais (MARTENS; NAES, 1992; VALDERRAMA, 2009). Assim, o *leverage* ( $h$ ) é a medida da influência de cada amostra no modelo e pode ser definido de acordo com a Equação (4) (BOTELHO; MENDES; SENA, 2013):

$$h_i = \mathbf{t}_{A,i}^T (\mathbf{T}_A^T \mathbf{T}_A)^{-1} \mathbf{t}_{A,i} \quad (4)$$

Onde  $\mathbf{T}$  representa os escores das amostras de calibração,  $\mathbf{t}_i$  é o vetor de escores da amostra  $i$  e  $A$  é número de variáveis latentes.

### 3.5 VALIDAÇÃO ANALÍTICA E FIGURAS DE MÉRITO

Depois de se construir o modelo de calibração, este deve ser validado. A validação de um método analítico é realizada a fim de se garantir uma metodologia exata, que possa ser reproduzida, comprovando que o método proposto contempla as normas estabelecidas por agências reguladoras e órgãos de fiscalização nacionais e internacionais. Este procedimento pode ser atestado pela determinação de parâmetros conhecidos como figuras de mérito. As principais são: linearidade, exatidão, precisão, sensibilidade, seletividade, coeficiente de correlação, limites de detecção e quantificação, relação de desempenho do desvio, robustez, intervalos de confiança, teste para erros sistemáticos, extensão da faixa de trabalho ou faixa linear dinâmica (SILVA, 2011; VALDERRAMA, 2009).

#### 3.5.1 Exatidão

Refere-se ao grau de concordância entre o valor de referência e o valor estipulado pelo modelo de calibração. Em quimiometria, este valor é avaliado com base nos resultados da raiz quadrada do erro médio quadrático de previsão (RMSEP – do inglês *Root Mean Squares Error of Prediction*), que é uma

aproximação do erro padrão das amostras de predição (VALDERRAMA; BRAGA; POPPI, 2007a).

### 3.5.2 Linearidade

A estimativa desta figura de mérito em calibração multivariada aplicada em modelo PLS é considerada problemática, visto que as variáveis são decompostas em componentes principais. Desta forma, a linearidade é avaliada qualitativamente por meio do gráfico dos resíduos de previsão para as amostras de calibração e validação. Quando este gráfico apresenta um comportamento aleatório, considera-se que os dados seguem um comportamento linear (VALDERRAMA, 2009; VALDERRAMA; BRAGA; POPPI, 2009).

### 3.5.3 Coeficiente de correlação (R)

O coeficiente de correlação (R) fornece informações sobre a qualidade do ajuste do modelo, onde os valores a serem obtidos variam de 0 a 1,0, sendo que 0 sugere que não há correlação e 1,0 demonstra uma relação linear perfeita entre os valores obtidos pelo método de referência e os valores previstos. Logo, são desejados valores de R o mais próximo de 1,0 possível. A notação geral de R é dada na Equação (5), onde o  $y_i$  são os valores de referência,  $\hat{y}_i$ , os valores previstos, e  $\bar{y}_i$ , a média dos valores de referência (POREP; KAMMERER; CARLE, 2015).

$$R = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

### 3.5.4 Precisão

A precisão representa o grau de concordância entre os resultados obtidos em condições determinadas para uma mesma amostra homogênea, ou seja, é a



avaliação da proximidade dos resultados obtidos em uma série de replicatas autênticas. A precisão pode ser determinada em vários níveis, como: repetibilidade, precisão intermediária, reprodutibilidade e precisão média (VALDERRAMA, 2009).

### 3.5.5 Ajuste

O ajuste é um parâmetro determinado a partir da correlação entre os valores tidos como referência e os valores estimados pelo modelo de calibração multivariada, podendo ser considerado também um indicativo de exatidão do modelo (VALDERRAMA, 2009).

### 3.5.6 Relação de Predição do Desvio (RPD)

A RPD (do inglês, *Residual Prediction Deviation*), é um outro dado estatístico importante para a interpretação dos conjuntos de calibração e validação, e é definida como a razão entre o desvio padrão (DP) dos valores de referência e RMSECV, conforme as Equações 6 e 7 (POREP; KAMMERER; CARLE, 2015):

$$RPD_{cal} = \frac{DP_{cal}}{RMSECV} \quad (6)$$

$$RPD_{val} = \frac{DP_{val}}{RMSEP} \quad (7)$$

Onde RMSECV é o erro quadrático médio de validação cruzada e  $DP_{cal}$  e  $DP_{val}$  são os desvios padrão dos valores de referência dos conjuntos de calibração e validação, respectivamente.

### 3.5.7 Intervalo de confiança

O intervalo de confiança é medido através da amplitude de valores para os quais o método analítico apresenta exatidão, precisão e linearidade adequadas quando aplicado a amostras contendo quantidades de substâncias dentro do intervalo especificado (SILVA, 2011).

### 3.5.8 Extensão da faixa de trabalho

É realizada determinando a espécie de interesse em diferentes concentrações. Por meio dos resultados obtidos, determina-se a faixa de concentração na qual os resultados exibem um nível aceitável de incerteza para o método utilizado (VALDERRAMA, 2009).

### 3.5.9 Sensibilidade e sensibilidade analítica

A sensibilidade (SEN) é estabelecida como a fração de sinal responsável pelo acréscimo de uma unidade de concentração da propriedade de interesse. Esta pode ser expressa conforme Equação (8), em modelos de calibração multivariada, como o PLS (SILVA, 2011).

$$\text{Sensibilidade} = \frac{1}{\|b\|} \quad (8)$$

Onde **b** é o vetor dos coeficientes de regressão com as variáveis latentes do PLS.

A sensibilidade analítica ( $\gamma$ ) geralmente não é abordada em normas ou protocolos de validação. Entretanto, esta figura de mérito apresenta a sensibilidade do método em função da unidade de concentração que é utilizada, sendo definida como a razão entre a sensibilidade e o desvio padrão do sinal de referência ( $\delta x$ ) (SILVA, 2011):

$$\text{Sensibilidade analítica} = \frac{\text{sensibilidade}}{\delta x} \quad (9)$$

O inverso da sensibilidade analítica ( $\gamma^{-1}$ ) possibilita estimar a menor diferença de concentração entre amostras que pode ser distinguida pelo método.

### 3.5.10 Robustez

Consiste em medir a sensibilidade de um método analítico frente a pequenas variações de seus parâmetros e verificar se estas são ou não significativas. O método é considerado robusto quando ele não é afetado por essas pequenas variações (VALDERRAMA, 2009; VALDERRAMA; BRAGA; POPPI, 2009).

### 3.5.11 Limite de detecção e quantificação

O limite de detecção (LD) e quantificação (LQ) de um procedimento analítico estima as menores quantidades da espécie de interesse que podem ser detectadas e determinadas quantitativamente, respectivamente. Para métodos multivariados, o LD e o LQ são estimados a partir da sensibilidade e do desvio padrão do sinal de referência ( $\delta x$ ) de acordo com as Equações a seguir (SILVA, 2011; VALDERRAMA, 2009):

$$\text{Limite de detecção} = 3,3 \delta x \frac{1}{\text{sensibilidade}} \quad (10)$$

$$\text{Limite de quantificação} = 10 \delta x \frac{1}{\text{sensibilidade}} \quad (11)$$

Desta forma as figuras de mérito exatidão, linearidade, ajuste, sensibilidade, sensibilidade analítica, inverso da sensibilidade, limites de detecção e quantificação e relação de desempenho do desvio foram utilizadas neste estudo para caracterizar a qualidade e eficiência do método proposto.

## 4. MATERIAL E MÉTODOS

O trabalho foi desenvolvido nos laboratórios da Universidade Tecnológica Federal do Paraná, *Campus* Campo Mourão, juntamente com o laboratório físico-químico do controle de qualidade de uma Cooperativa Agroindustrial da região.

### 4.1 PREPARO DAS AMOSTRAS

As amostras de soja trituradas foram fornecidas pela Cooperativa, sendo um total de 300 amostras provenientes de diferentes cidades da região noroeste do Estado do Paraná, no período de 29 de março à 07 de junho de 2016. Para as análises, cerca de 250 g de amostra foram secas em estufa de circulação de ar forçada a 130 °C por 1 hora e, na sequência, levadas ao dessecador, permanecendo por cerca de 12 horas. Em seguida as amostras secas foram trituradas em moinho de rotor (Marconi, modelo MA-090CFT), com 0,85 mm de abertura, e encaminhadas para análises por meio dos métodos de referência, na própria indústria beneficiadora dos grãos de soja, para a produção de óleo e farelo. Estas medidas foram realizadas de segunda a sexta-feira, no primeiro turno, sem repetição, entretanto, os resultados dos valores de referência foram duplicados.

### 4.2 MÉTODOS DE REFERÊNCIA

#### 4.2.1 Determinação do índice de acidez

Este método determina a acidez livre do óleo removido da semente de soja por hexano a temperatura ambiente. Para isso, aproximadamente 25 g da amostra de soja moída foram adicionadas a 50 mL de hexano e submetidos a agitação constante por 1 hora em agitador magnético, para extração do óleo. Na

sequência, o sobrenadante foi filtrado, onde o extrato obtido foi levado a estufa de circulação de ar a 100 °C, afim de se evaporar o solvente residual. Em seguida, pesou-se aproximadamente 7,0 g do óleo extraído, adicionando-se 50 mL de álcool isopropílico 99%, neutralizado, e 1 mL de fenolftaleína a 1%. Realizou-se a titulação com hidróxido de sódio (NaOH) 0,25 mol.L<sup>-1</sup>, até coloração rósea persistente por aproximadamente 1 min (AOCS adaptado, 2009).

O índice de acidez corresponde à quantidade (em mg) de base necessária para neutralizar os ácidos graxos livres presentes em 1 g de óleo. Os resultados foram expressos em porcentagem, utilizando a seguinte equação:

$$\% \text{ Acidez livre} = \frac{V \times 0,25 \times F \times 282 \times 100}{M \times 1000} \quad (12)$$

Onde:

V= volume de solução de hidróxido de sódio 0,25 mol.L<sup>-1</sup> gasto na titulação (mL);

F= fator de correção da solução de hidróxido de sódio;

0,25= molaridade requerida da solução de NaOH;

M= massa da amostra de óleo utilizado (g);

282= peso equivalente de ácido livre no qual os resultados são expressos.

A acidez livre é usualmente expressa com % de ácido oleico.

#### 4.2.2 Determinação de umidade

O teor de umidade da soja moída foi medido seguindo-se o protocolo da AOCS (2009), onde 5 g de amostra são adicionados em cadinho, previamente calcinado, seco em estufa a 105 °C até peso constante. Os cálculos foram realizados de acordo com a expressão:

$$\% \text{ Umidade} = \frac{100 \times M_p}{M_i} \quad (13)$$

Onde:

Mp= massa perdida pela amostra durante o experimento (g);

Mi= massa inicial da amostra (g).

#### 4.2.3 Determinação do teor de lipídios totais

Para a extração do óleo, pesou-se 2 g de amostra de soja moída, em papel filtro, o qual foi acondicionado dentro de um cartucho extrator de celulose 43x123mm. Em seguida, o cartucho foi inserido no extrator de Soxhlet. Em balões de fundo chato (tarados e pesados previamente em balança analítica), foram adicionados aproximadamente 250 mL de hexano. O volume do solvente foi mantido razoavelmente constante durante a extração, e quando necessário foi adicionado mais solvente para ajustar qualquer perda devido à evaporação. O período de extração foi de 5 horas e, ao término, resfriaram-se e desconectaram-se os balões. Em seguida, estes foram levados à chapa aquecedora, para evaporar o hexano. Para garantir a eliminação de solvente remanescente na amostra, os balões foram levados para estufa de circulação de ar por 1 hora a 130 °C, resfriados na sequência e transferidos para dessecador até atingir a temperatura ambiente e, finalmente, pesados em balança analítica (AOCS adaptado, 2009).

A equação abaixo foi utilizada para determinação do percentual de lipídios totais:

$$\% \text{ Lipídios totais: } \frac{P_f - P_i}{P_a} \times 100 \quad (14)$$

Onde:

Pf= Peso final;

Pi= Peso inicial;

Pa= Peso da amostra.

#### 4.2.4 Determinação de proteína bruta

A metodologia de referência utilizada para a determinação de proteína bruta foi o método de Dumas (AOCS, 2009). Este se baseia no princípio da combustão, onde, todo o processo é realizado em um só equipamento: combustão, redução, separação e detecção (ESKALB, 2014). Consiste na oxidação total da matéria orgânica, em uma temperatura de aproximadamente 800 °C, sendo todo o nitrogênio transformado em gases de óxido nítrico (NO<sub>2</sub>), e logo após são reduzidos a gás nitrogênio (N<sub>2</sub>), que é detectado por condutividade térmica (NETO; SIMEONE; GUIMARÃES, 2012).

Para as medidas, quantidades entre 0,1950 a 0,2050 g de soja moída foram pesados em cápsula de estanho, as quais foram colocadas no equipamento analisador de nitrogênio (LECO, modelo FP-628), para digestão oxidativa com oxigênio puro (99,99%). O teor de proteína bruta foi calculado multiplicando-se o teor de nitrogênio encontrado pelo fator de correção 6,25.

#### 4.2.5 Determinação do teor de clorofila

Aproximadamente 25 g da amostra de soja moída foram empacotadas em papel filtro, colocado em um cartucho extrator de celulose e levado ao extrator de soxhlet com hexano, por cerca de 5 horas. O hexano da amostra foi evaporado em chapa aquecedora, levado a estufa de circulação de ar forçada a 130 °C por 1 hora, para facilitar a remoção total do solvente, e resfriado em dessecador até a temperatura ambiente. Em seguida, o óleo obtido foi submetido à leitura de absorbância em colorímetro espectrofotométrico automático (marca Lovibond, modelo Tintometer PFXi-880/F) nos comprimentos de onda 630, 670 e 710 nm, utilizando-se cubeta de 50 mm de caminho óptico (AOCS adaptado, 2009).

A clorofila quantificada foi expressa através do seguinte cálculo:

$$C = \frac{A_{670} - [(A_{630} + A_{710}) / 2]}{0,0964 \times L} \times 1000 \quad (15)$$

Onde:

A= Absorbância do óleo no respectivo comprimento de onda em nanômetros;

L= Espessura da cubeta (cm);

C= Concentração de mg de clorofila/kg de óleo (ppm);

0,0964= Fator específico do espectrofotômetro.

### 4.3 ANÁLISE ESPECTROSCÓPICA NA REGIÃO DO INFRAVERMELHO PRÓXIMO

Os espectros das amostras de soja moída, foram obtidos com o espectrômetro de infravermelho próximo (MicroNIR™ 1700, JDSU *Uniphase Corporation*), adquiridos à temperatura ambiente, por refletância difusa, sendo cada amostra avaliada em duplicata. Este equipamento oferece varredura *online* para a região que vai de 910 até 1676 nm, com um intervalo constante de 6,2 nm. A medida do branco foi realizada utilizando-se um padrão de reflectância NIR (*Spectralon™*) com um coeficiente de reflexão difusa de 99%, enquanto que uma referência escura (zero – para simular não-reflexão) foi obtida com a lâmpada apagada. Os dados foram exportados para se aplicar a análise quimiométrica.

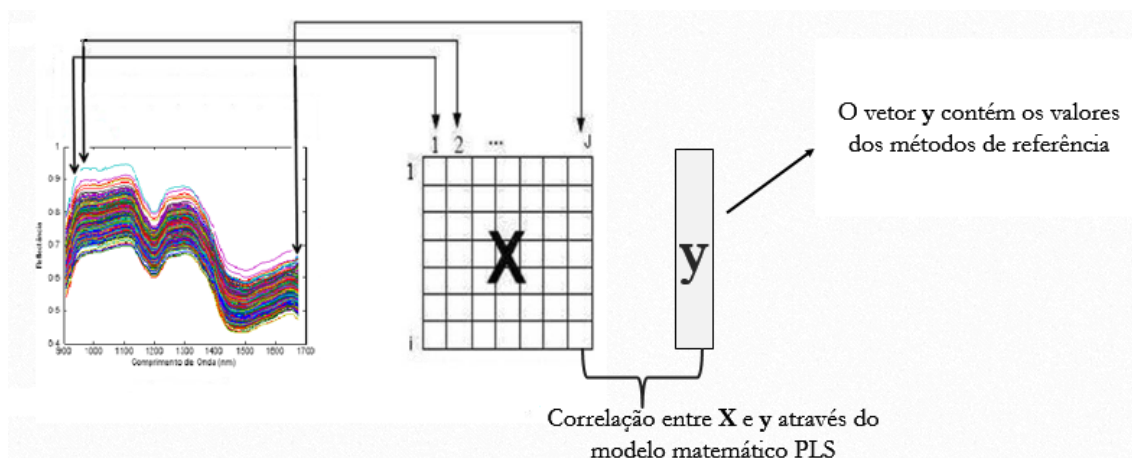
### 4.4 PROCESSAMENTO DOS DADOS

Após a obtenção dos espectros, foi empregado o método quimiométrico de Mínimos Quadrados Parciais (PLS). Para o tratamento dos dados foi empregado o software Matlab R2013a com as ferramentas do pacote PLS-Toolbox 7.8 (*Eingevector Research Inc.*), fornecido pela EMBRAPA Solos do Rio de Janeiro.

Os espectros foram organizados no formato de uma matriz, em que cada linha da matriz **X** corresponde a um espectro e cada coluna corresponde à absorbância nos diferentes comprimentos de onda. A matriz **X** foi então correlacionada a um vetor **y** que continha os valores dos parâmetros obtidos pelo



método padrão, e esta correlação entre  $X$  e  $y$  foi realizada através do método de PLS (OTTO, 1999), conforme apresentado na Figura 1.



**Figura 1:** Construção da matriz  $X$  para calibração multivariada (adaptado de VALDERRAMA, 2005).

Neste método, uma matriz de respostas instrumentais  $X$  e um vetor contendo os valores obtidos a partir de uma metodologia de referência  $y$  é decomposto em um produto de duas matrizes, denominadas de *scores* e *loadings* (BRERETON, 2000; VALDERRAMA et al., 2014). O modelo PLS é construído a partir de uma relação linear obtida entre os *scores* de  $X$  e os *scores* de  $y$  (GELADI; KOWALSKI, 1986; VALDERRAMA et al., 2014).

Para construção de modelos PLS são necessárias etapas de calibração e de validação dos modelos construídos (KOWALSKI; BEEBE, 1987). Portanto, as amostras foram separadas em dois conjuntos (amostras de validação e calibração), com auxílio do algoritmo de 'Kenston' (KENNARD; STONE, 1969).

Dessa forma, para os modelos previstos, foram selecionadas para o conjunto de calibração dois terços das amostras, e um terço foi destinado para a validação. O pré-tratamento utilizado foi o algoritmo de *Savitski-Golay* (janela com 13 pontos, polinômio 1º grau e 1ª derivada) (SAVITSKY; GOLAY, 1964). Já para o modelo de previsão do teor de clorofila, os espectros foram processados para ajustar a linha base pelo algoritmo de correção de espalhamento multiplicativo e, em seguida, alisados pela aplicação do algoritmo de *Savitski-Golay* (janela com 9 pontos, polinômio de 1º grau). Todos os dados espectrais foram centrados na média antes da calibração.

Os modelos foram construídos utilizando-se validação cruzada através de blocos contínuos de 10 amostras. Neste caso, 10 amostras são deixadas de fora no processo de construção do modelo e a seguir essas amostras são previstas pelo modelo construído. Esse processo se repete até que todas as amostras tenham sido deixadas de fora e previstas. As amostras anômalas (*outliers*) foram identificadas e removidas quando necessário através da análise dos gráficos de  $Q_{\text{residual}}$  versus *leverage* (SILVA et al., 2014).

A validação dos modelos de calibração multivariada foi realizada pela determinação e avaliação das figuras de mérito: exatidão, linearidade, ajuste, sensibilidade, sensibilidade analítica, limites de detecção e quantificação. Os cálculos foram realizados com base em trabalhos descritos na literatura (VALDERRAMA; BRAGA; POPPI, 2007; VALDERRAMA; BRAGA; POPPI, 2009). A relação de desempenho do desvio (RPD, *Residual Prediction Deviation*) foi avaliada de acordo com Botelho et al. (2013).

A exatidão expressa o grau de concordância entre o valor estimado pelo modelo multivariado e o valor tido como verdadeiro obtido pelo método de referência. Esse parâmetro é expresso através da Raiz Quadrada do Erro Médio Quadrático de Calibração (RMSEC) e de Previsão (RMSEP) (MARTENS; NAES, 1992):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{nv} (y_i - \hat{y}_i)^2}{nv}} \quad (16)$$

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{nc} (y_i - \hat{y}_i)^2}{nc}} \quad (17)$$

Em que:  $y_i$  é o valor de referência,  $\hat{y}_i$  é o valor previsto pelo modelo multivariado,  $nv$  é o número de amostras na validação e  $nc$  é o número de amostras na calibração.

O ajuste é um parâmetro do modelo multivariado que é estimado a partir da correlação entre os valores tidos como “verdadeiros” (obtidos pelo método de referência) e os valores estimados pelo modelo de calibração multivariada. Este processo é realizado por mínimos quadrados, determinando a reta que melhor

se ajusta aos valores de referência e aos valores estimados pelo modelo para as amostras de calibração (MARTENS; NAES, 1992).

Avaliar a linearidade é um problema em calibração multivariada, pois as variáveis são decompostas em *scores* e *loadings*. Qualitativamente, o gráfico dos resíduos para as amostras de calibração e validação pode indicar se os dados seguem um comportamento linear se a distribuição dos resíduos for aleatória (MARTENS; NAES, 1992).

A avaliação das figuras de mérito em conjunto com os métodos de análise multivariada definiu o potencial do modelo para índice de acidez, teores de lipídios totais, proteína bruta, clorofila e umidade.

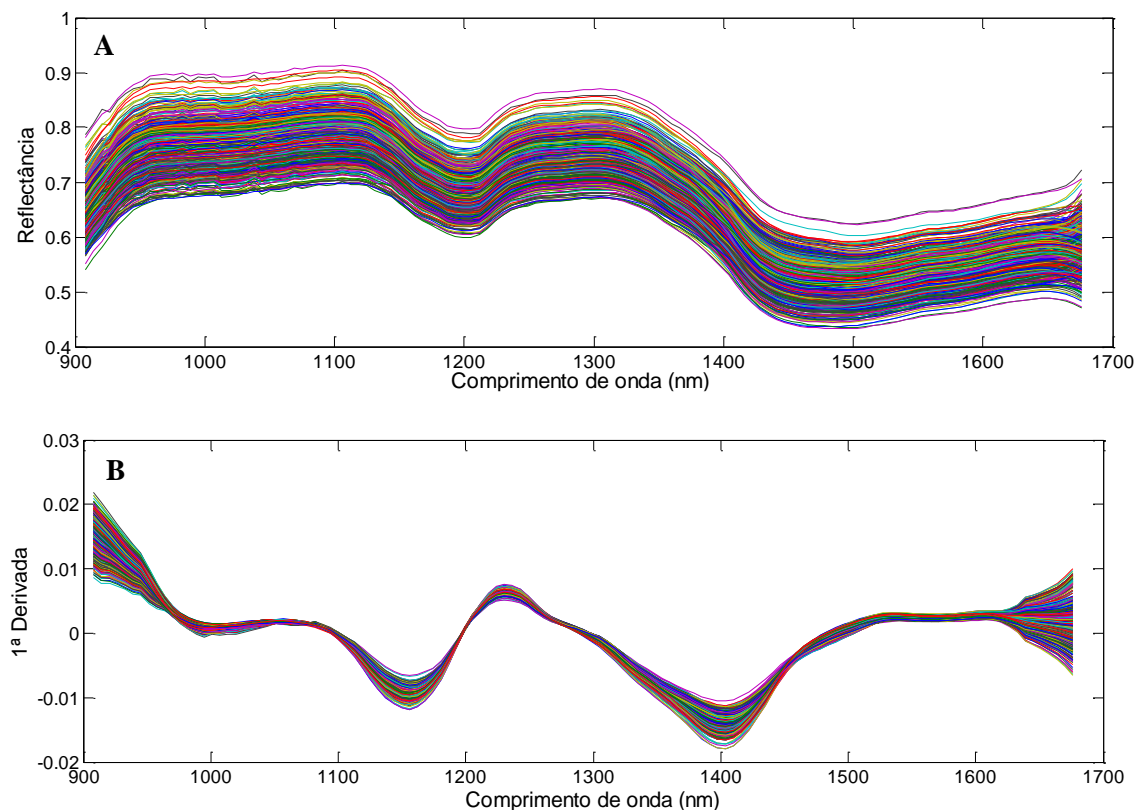
## 5. RESULTADOS E DISCUSSÕES

As medidas dos teores de clorofila se apresentaram com grandes variações e, dessa forma, os resultados referentes a este parâmetro serão relatados separadamente mais adiante. Os resultados obtidos via método de referência para os parâmetros índice de acidez, lipídios totais, proteína bruta e umidade estão apresentados na Tabela 1, informando apenas a variação das medidas (máximo e mínimo para cada parâmetro). Para a determinação do índice de acidez, os resultados variaram de 0,22 a 4,36%. Os percentuais de proteína bruta ficaram entre 33,09 e 40,11%. O teor de lipídios totais teve uma variação entre 16,04 e 26,33% enquanto que, a porcentagem de umidade ficou entre 1,17 a 9,76%. Os resultados obtidos pelas análises de referência estão apresentados de forma integral no Anexo I. Essa variação nos resultados pode ser proveniente de aspectos relacionados à heterogeneidade da matéria-prima utilizada, que era originária de diferentes regiões. Assim, fatores como variações climáticas, perfil do solo, déficit hídrico, colheita, transporte e armazenamento dos grãos, podem influenciar significativamente a composição química (GREGGIO; BONINI, 2014).

**Tabela 1:** Variações nos valores obtidos pelos métodos de referência para amostras de soja trituradas

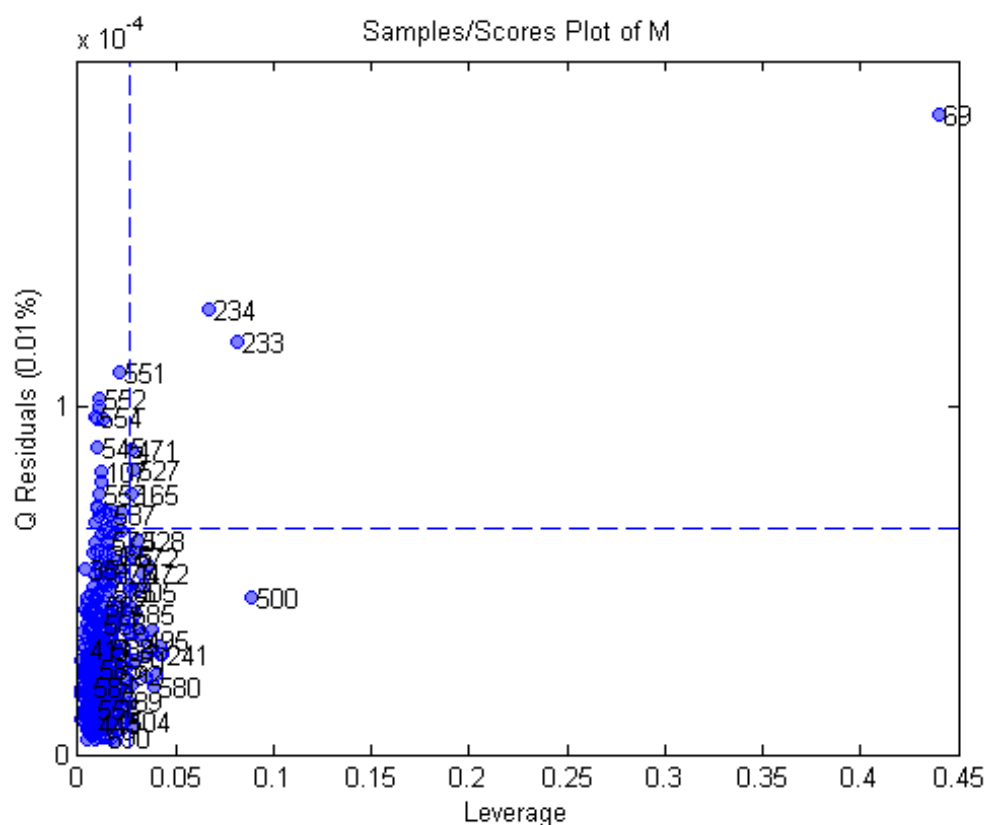
<b>Índice de acidez</b> <b>(%)</b>	<b>Proteína bruta</b> <b>(%)</b>	<b>Lipídios totais</b> <b>(%)</b>	<b>Umidade</b> <b>(%)</b>
0,22 - 4,36	33,09 - 40,11	16,04 - 26,33	1,17 - 9,76

Foram adquiridos espectros das amostras de soja moída, os quais estão apresentados na Figura 2 (A) sem processamento e (B) após primeira derivada via algoritmo de *Savitsky-Golay*. No caso das análises de clorofila, um número menor de amostras foi avaliado e, portanto, este parâmetro será tratado em separado.



**Figura 2:** Espectros de infravermelho próximo obtidos a partir de soja triturada (A) sem pré-processamento e (B) após aplicação de primeira derivada.

Para se obter os modelos de calibração multivariada, procedeu-se a detecção e remoção de amostras anômalas. A detecção de amostras anômalas foi realizada avaliando-se os valores de *leverage* e *Q residuals* (referente ao resíduo espectral), simultaneamente. O “*leverage*” pode ser interpretado como a distância entre uma amostra e o centro do modelo, ou seja, o quão distante um ponto está quando comparado com o conjunto total. Um *leverage* próximo de zero indica que a variável correspondente tem pouca influência no modelo, já variáveis com altos valores de *leverages* demonstram que aquele ponto tem grande influência no modelo. Desta forma, se a amostra apresenta ao mesmo tempo altos valores de *leverage* e de resíduo, pode-se dizer que a amostra é muito diferente das demais e influencia significativamente no modelo de modo a causar erro, já que o resíduo deflagra a diferença entre o valor projetado para a amostra na nova dimensão calculada e seu valor real, conforme exemplo que pode ser visualizado a partir da Figura 3.



**Figura 3:** Amostras anômalas para o modelo PLS de índice de acidez.

O conjunto de amostras consideradas *outliers* foi diferente para cada parâmetro avaliado. Assim, os conjuntos de calibração e validação demandaram um número diferente de amostras, como apresentado na Tabela 2. De acordo com os guias de validação brasileiros e internacionais, *outliers* podem ser removidos até um máximo de 22,2% do número total de amostras (BOTELHO; MENDES; SENA, 2013), e esse limite foi respeitado.

**Tabela 2:** Conjuntos de amostras utilizados para a construção dos modelos e *outliers* detectados e removidos para cada parâmetro.

Parâmetro	Conjunto de Calibração	Conjunto de Validação	<i>Outliers</i>
Índice de acidez	348	175	69
Proteína bruta	342	170	80
Lipídios totais	346	154	92
Umidade	349	158	85

Observou-se que as amostras consideradas anômalas se tratavam de amostras com teores extremos, ou seja, baixos ou altos níveis dos parâmetros

avaliados, apresentando-se diferentes do restante do conjunto. Como o modelo de calibração é linear, amostras desse tipo são mais informativas do que propriamente anomalias. No entanto, acabam por distorcer a capacidade de previsão do modelo. A única forma de mantê-las no modelo seria avaliar um número ainda maior de amostras contemplando teores similares aos removidos, o que não foi objetivo deste estudo.

Para a construção dos modelos PLS, o vetor  $\mathbf{y}$  foi composto pelos valores de referência dos métodos utilizados na determinação de cada parâmetro, enquanto que a matriz  $\mathbf{X}$  foi composta pelos espectros de cada amostra.

Para a construção dos modelos de calibração, foram utilizados os conjuntos de dados de calibração apresentados na Tabela 2. A separação das amostras entre os conjuntos de calibração e validação foi realizado pelo algoritmo de *Kennard-Stone* (KENNARD; STONE, 1969). Esse algoritmo garante que as amostras mais representativas estejam no conjunto de calibração, através de uma varredura uniforme dos dados espectrais (BOTELHO; MENDES; SENA, 2013).

O algoritmo de *Kennard-Stone* seleciona as amostras com base em suas distâncias, de modo que a primeira amostra selecionada é a que apresenta a maior distância em relação à amostra mais central. A segunda amostra a ser escolhida será a que apresentar maior distância em relação à primeira amostra selecionada. A amostra seguinte a ser selecionada apresentará maior distância em relação à última amostra escolhida, e assim sucessivamente até atingir o número de amostras desejadas (VALDERRAMA, 2009).

Os modelos foram centrados na média utilizando-se 12 variáveis latentes, que correspondiam a mais de 98% da variância dos dados originais (espectrais e concentrações de referência). O número de variáveis latentes (VL) foi determinado através dos resultados da raiz quadrada do erro quadrático médio de validação cruzada (RMSECV) para as amostras de calibração, obtido por validação cruzada em blocos contínuos de 10 amostras, e levando-se em consideração a percentagem de variância explicada no vetor  $\mathbf{y}$ .

A escolha do número de variáveis latentes utilizado no modelo é uma etapa de extrema importância, pois influencia diretamente nos resultados. A escolha de um número menor de variáveis latentes poderá gerar resultados não satisfatórios, uma vez que toda a informação disponível dos dados originais não

estará sendo utilizada. Tal situação é chamada de “*subajuste*”. Em contra partida, ao se usar um número de variáveis latentes superior ao necessário ocorrerá a modelagem de ruídos, ocasionando problemas no modelo de modo que tal modelo poderá ser utilizado apenas para aquela situação e não para previsões futuras (ARAUJO, 2007; BARTHUS, 1999). Tais modelos são denominados “*sobreajustados*”.

A performance dos modelos desenvolvidos foi verificada a partir da análise dos valores de algumas figuras de mérito. Foram avaliadas as figuras de mérito de exatidão, através do erro médio quadrático de calibração (RMSEC) e previsão (RMSEP), ajuste (coeficiente de correlação - R), coeficiente de determinação ( $R^2$ ), sensibilidade, sensibilidade analítica, inversa da sensibilidade analítica, limite de detecção, limite de quantificação e relação de desempenho do desvio ( $RPD_{cal}$  e  $RPD_{val}$ ).

Na Tabela 3 são apresentadas as figuras de mérito, calculadas para cada modelo avaliado.

**Tabela 3:** Figuras de mérito do modelo PLS para os parâmetros físico-químicos da soja.

Figuras de mérito	Índice de acidez (%)	Lipídios totais (%)	Proteína bruta (%)	Umidade (%)
$R_{cal}$	0,8589	0,8811	0,8879	0,8721
$R_{val}$	0,8541	0,8290	0,8963	0,8370
$R^2_{cal}$	0,7378	0,7763	0,7884	0,7606
$R^2_{val}$	0,7295	0,6873	0,8034	0,7006
RMSEC	0,1970	0,7504	0,5135	0,9715
RMSEP	0,1694	0,6747	0,4552	0,9914
$RPD_{cal}$	1,7661	1,6906	2,0377	1,8851
$RPD_{val}$	1,8525	1,6213	2,2596	1,8250
Sensibilidade	$3,346 \times 10^{-4}$	$9,468 \times 10^{-5}$	$1,614 \times 10^{-4}$	$6,747 \times 10^{-5}$
Sensibilidade Analítica	18,0102	5,3872	9,6200	3,5593
Inversa da Sensibilidade Analítica	0,0555	0,1856	0,1039	0,2810
Limite de Detecção	0,1832	0,6126	0,3430	0,9271
Limite de Quantificação	0,5552	1,8562	1,0395	2,8095

$R_{cal}$  – Coeficiente de correlação de calibração;  $R_{val}$  – Coeficiente de correlação de validação;  $R^2_{cal}$  – Coeficiente de determinação de calibração;  $R^2_{val}$  – Coeficiente de determinação de validação; RMSEC – Raiz quadrada do erro médio quadrático de calibração; RMSEP – Raiz quadrada do erro médio quadrático de previsão;  $RPD_{cal}$  –



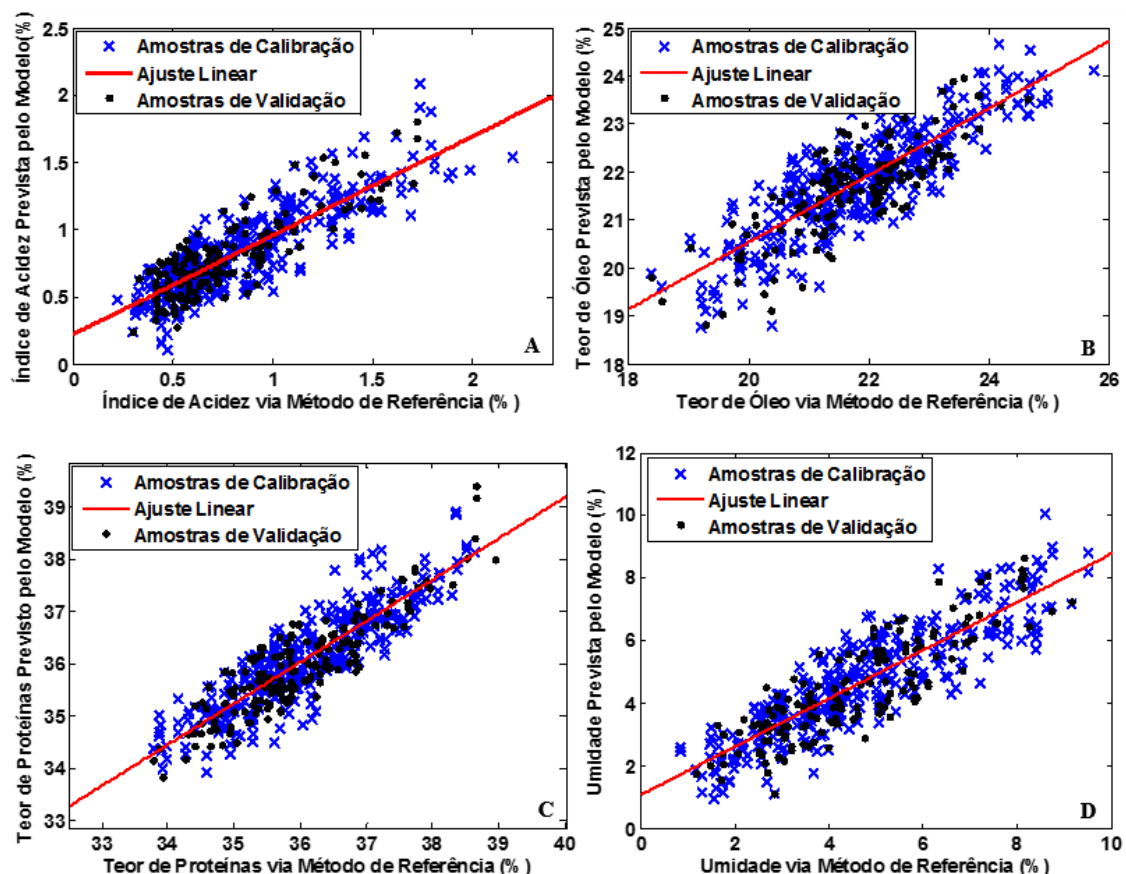
Relação de predição de desvio de calibração;  $RPD_{val}$  – Relação de predição de desvio de validação.

Observando-se a Tabela 3 e levando em consideração que trata-se de uma situação de aplicação real (não simulada em laboratório), todos os modelos apresentaram uma correlação satisfatória entre os valores de referência e os valores previstos pelo modelo via espectros NIR.

Para avaliar o ajuste do modelo construído utiliza-se os parâmetros RMSEC e RMSEP, que são considerados indicadores de exatidão, onde os valores gerados por estes parâmetros devem ser o mais próximo possível entre si. Isso sugere que os valores estimados pelos modelos multivariados apresentaram uma boa concordância com os métodos de referência, e indica que o número de variáveis latentes escolhidos foi adequado para o modelo não havendo sobreajuste ou subajuste (VALDERRAMA; BRAGA; POPPI, 2007b).

Os valores de RMSEP e RMSEC são parâmetros globais que incorporam tanto erros sistemáticos quanto aleatórios, oriundos do método de referência. Os erros de regressão foram considerados baixos para todos os modelos, sugerindo a possibilidade de se prever os parâmetros avaliados utilizando-se espectroscopia NIR sem se afastar da variação de erros encontrada pelos métodos de referência, isso também é um indicativo de que o número de variáveis latentes escolhido foi adequado para os modelos propostos.

Um outro indicador de exatidão é a regressão entre os valores de referência e os valores estimados pelo modelo, ou seja, o ajuste (R). O ajuste obtido para os modelos está apresentado na Figura 4. Os resultados podem ser considerados adequados já que os métodos de referência apresentaram uma ampla variação dos resultados obtidos, e sabe-se que o coeficiente de correlação está diretamente relacionado com os erros embutidos nos métodos de referência. Logo, se o método padrão apresenta altos níveis de erros analíticos estes estarão presentes nos modelos de calibração multivariada, diminuindo os valores de coeficiente de correlação.



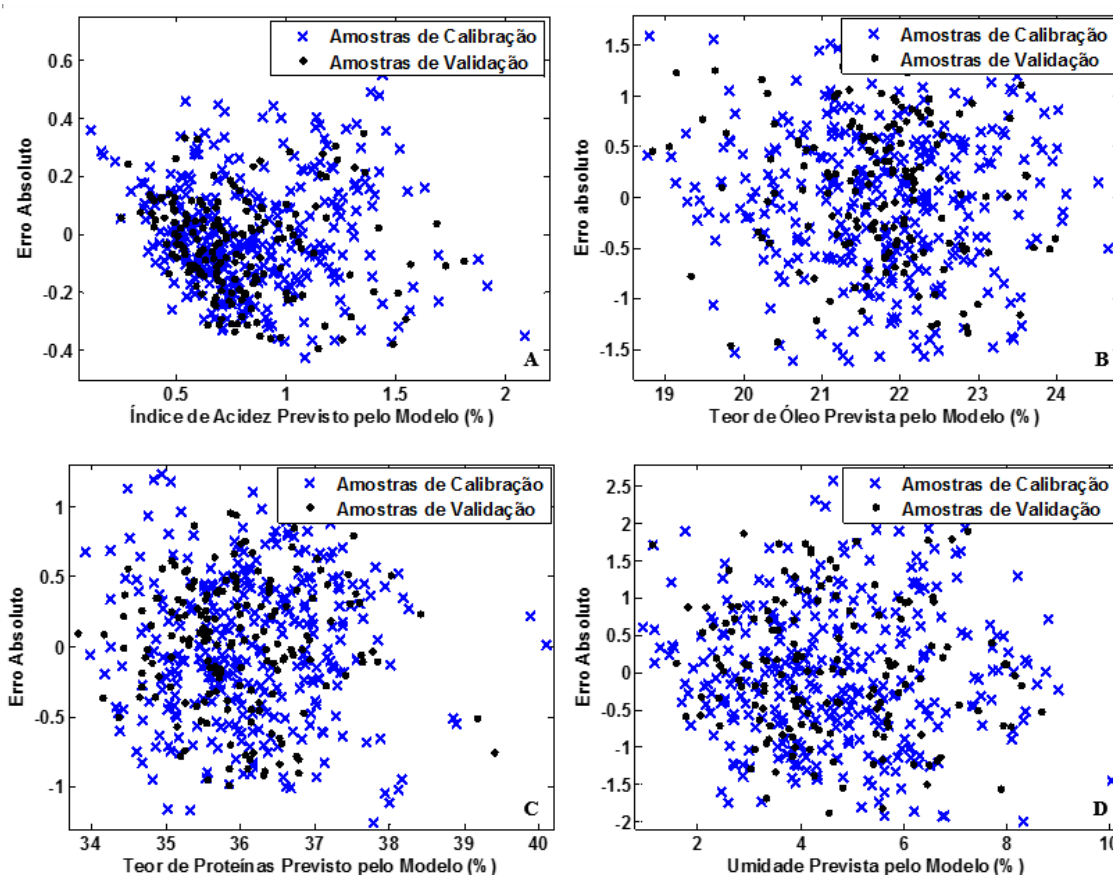
**Figura 4:** (A) Ajuste do modelo para índice de acidez; (B) Ajuste do modelo para lipídios totais; (C) Ajuste do modelo para proteína bruta e (D) Ajuste do modelo para teor de umidade. (x) Amostras de Calibração (●) Amostras de Validação.

Os coeficientes de determinação ( $R^2$ ) para calibração apresentaram valores de 0,7378 para o modelo de índice de acidez, 0,7763 para lipídios totais, 0,7884 para proteína bruta, e 0,7700 para o modelo de umidade. Resultados de coeficiente de determinação similares já foram relatados na literatura quando o método de referência, na calibração multivariada são dessa natureza (FERREIRA et al., 2014; VALDERRAMA; BRAGA; POPPI, 2007a, 2007b).

Ferreira e seus colaboradores (2014), ao analisar a composição centesimal de soja em grão por Espectroscopia de Infravermelho Próximo, obtiveram resultados de coeficientes de determinação de 0,72 para umidade; 0,88 para proteína bruta; e 0,67 para lipídios, valores próximos dos que foram encontrados nesse trabalho. No entanto, foram analisadas somente 40 amostras de soja em laboratório de pesquisa, o que se difere significativamente de uma situação de análise de rotina de uma indústria. Outro trabalho realizado recentemente por Leite (2015), relata sobre teores de lipídios totais e umidade em sementes de soja por NIR, no qual foram encontrados valores de coeficiente

de determinação de inferiores ao deste trabalho: 0,68 para umidade e 0,53 para lipídios totais.

A linearidade dos modelos PLS foi avaliada pela observação dos resíduos. Para este parâmetro, considera-se que os dados avaliados seguem modelo linear se os resíduos apresentarem dispersão aleatória. Assim, pode-se afirmar que o modelo PLS foi ideal para todos os casos avaliados, como pode ser observado na Figura 5.



**Figura 5:** Resíduos do modelo PLS para determinação de (A) índice de acidez; (B) lipídios totais; (C) proteína bruta e (D) umidade. (x) Amostras de Calibração (●) Amostras de Validação.

O valor de RPD do modelo de calibração (Tabela 3), foi de 1,7661 para índice de acidez, 1,6906 para lipídios totais, 2,0377 para proteína bruta e 1,8851 para umidade. A capacidade preditiva dos modelos foi considerada satisfatória, pois de acordo com a literatura, são considerados modelos de calibração de boa qualidade preditiva aqueles que apresentam valores para RPD acima de 2,4, e são classificados satisfatórios os que apresentam valores entre 2,4 e 1,5 (BOTELHO et al., 2013).

Considerando-se a faixa de variação dos valores obtidos pelo método de referência de cada parâmetro, a sensibilidade e sensibilidade analítica apresentaram resultados satisfatórios. A sensibilidade é definida como a fração do sinal responsável pelo acréscimo de uma unidade de concentração à propriedade de interesse, onde os modelos obtiveram um valor de  $3,346 \times 10^{-4}$  para índice de acidez,  $9,468 \times 10^{-5}$  para lipídios totais,  $1,614 \times 10^{-4}$  para proteína bruta e  $6,747 \times 10^{-5}$  para umidade. A sensibilidade analítica apresentou um resultado de 18,0102, 5,3872, 9,6200 e 3,5593, para índice de acidez, lipídios totais, proteína bruta e umidade, respectivamente. Este parâmetro apresenta a sensibilidade do método em termos da unidade de concentração que é utilizada, sendo definida como a razão entre a sensibilidade e o desvio padrão do sinal de referência (VALDERRAMA et al., 2009).

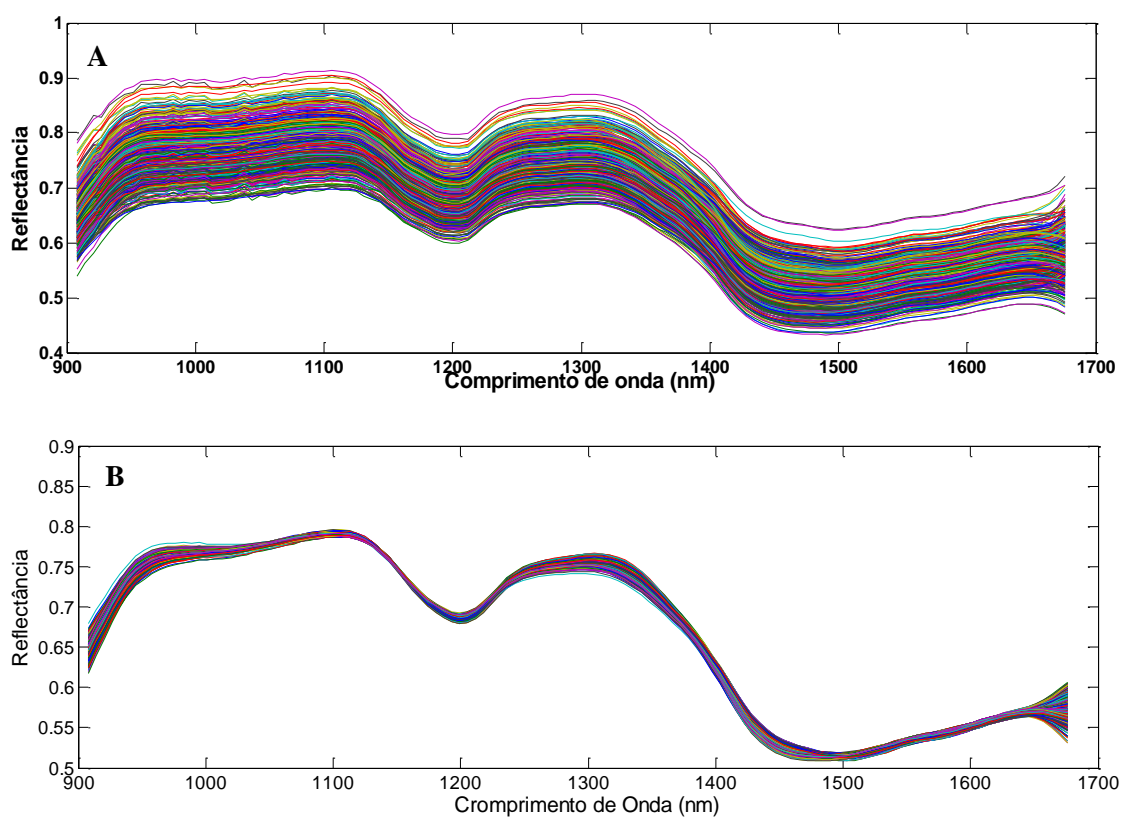
O inverso da sensibilidade analítica (sensibilidade analítica<sup>-1</sup>) permite estabelecer a menor diferença de concentração entre amostras, que pode ser distinguida pelo método, considerando-se o erro aleatório instrumental a única fonte de erros. Sendo assim, é possível fazer a distinção de amostras com diferença de teor de índice de acidez na ordem de 0,0555%, de 0,1856% para lipídios totais, de 0,1039% para proteína bruta e 0,2810% para umidade.

O limite de detecção equivale a menor concentração da substância de interesse que pode ser detectada, mas não necessariamente quantificada; desta forma, a concentração mínima que pode ser detectada confiavelmente pelo modelo empregando infravermelho foi 0,1832% para índice de acidez, 0,6126% para lipídios totais, 0,3430% para proteína bruta e 0,9271% para umidade, indicando que amostras abaixo deste teor podem ser medidas erroneamente. O limite de quantificação representa a menor concentração da substância de interesse que pode ser quantificada com confiabilidade satisfatória. O resultado demonstra que amostras com concentrações inferiores a 0,5552% para índice de acidez, 1,8562% para lipídios totais, 1,0395% para proteína bruta e 2,8095% para umidade, não podem ser diferenciadas.

O modelo construído para previsão dos teores de clorofila apresentou menor capacidade preditiva. Um dos principais motivos se deve ao fato de que os resultados obtidos pelo método de referência para determinação do teor de clorofila apresentaram variação muito alta, de 223 a 8.454 mg/kg (ppm). Além desta discrepância, as concentrações são consideradas baixas para serem

medidas por Espectroscopia no Infravermelho Próximo. No entanto, ainda assim é possível utilizar a técnica para se fazer previsão a respeito dos teores deste parâmetro.

Os espectros obtidos no infravermelho próximo (NIR) da soja triturada estão apresentados na Figura 6 (A) sem tratamento e (B) com correção de espalhamento multiplicativo e, alisados pela aplicação do algoritmo de *Savitski-Golay*. Fez-se a correção de espalhamento multiplicativo nestes dados para tentar melhorar a performance do modelo. No entanto, o efeito desta correção de espalhamento melhorou os parâmetros do modelo em proporções muito pequenas.



**Figura 6:** (A) Espectros NIR das amostras de soja sem pré-tratamento e (B) Espectros NIR pré-processados com correção de espalhamento multiplicativo e alisados pela aplicação do algoritmo de *Savitski-Golay*.

Para se obter o modelo de calibração multivariada para teor de clorofila, realizou-se a detecção e remoção de amostras anômalas. Em seguida, fez-se a separação das amostras entre os conjuntos de calibração e validação pelo algoritmo de *Kennard-Stone* (KENNARD; STONE, 1969). Os espectros foram medidos em duplicata enquanto os valores de referência foram duplicados.

Assim, 206 espectros e respectivos valores de referência foram utilizados para construir o modelo de calibração, enquanto 86 amostras foram utilizadas para a validação.

O desempenho do modelo para avaliação de clorofila nas amostras de soja, foi testado a partir dos valores das figuras de mérito raiz quadrada do erro médio quadrático de calibração (RMSEC) e previsão (RMSEP), ajuste (coeficiente de correlação - R), coeficiente de determinação ( $R^2$ ), sensibilidade, sensibilidade analítica, inversa da sensibilidade analítica, limite de detecção, limite de quantificação e relação de desempenho do desvio ( $RPD_{cal}$  e  $RPD_{val}$ ). Estes resultados são apresentados na Tabela 4.

**Tabela 4:** Figuras de mérito do modelo PLS para o parâmetro teor de clorofila da soja.

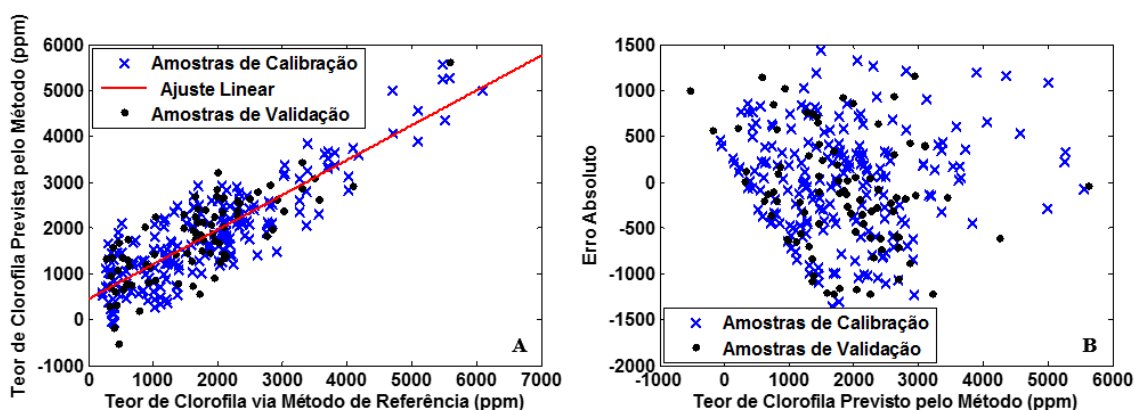
<b>Figuras de Mérito</b>	<b>Teor de clorofila</b>
$R_{cal}$	0,8709
$R_{val}$	0,8073
$R^2_{cal}$	0,7585
$R^2_{val}$	0,6517
RMSEC	623,2972
RMSEP	621,9778
$RPD_{cal}$	1,5956
$RPD_{val}$	1,6295
Sensibilidade	$6,5675 \times 10^{-8}$
Sensibilidade Analítica	0,0033
Inversa da Sensibilidade Analítica	300,6507
Limite de Detecção	992,1474 (ppm)
Limite de Quantificação	3006,50 (ppm)

$R_{cal}$  – Coeficiente de correlação de calibração;  $R_{val}$  – Coeficiente de correlação de validação;  $R^2_{cal}$  – Coeficiente de determinação de calibração;  $R^2_{val}$  – Coeficiente de determinação de validação; RMSEC – Raiz quadrada do erro médio quadrático de calibração; RMSEP – Raiz quadrada do erro médio quadrático de previsão;  $RPD_{cal}$  – Relação de predição de desvio de calibração;  $RPD_{val}$  – Relação de predição de desvio de validação.

Para avaliar o ajuste do modelo construído para teor de clorofila utilizou-se os parâmetros RMSEC e RMSEP, que são considerados indicadores de exatidão. Como os valores para o erro da calibração e da validação do modelo estão muito próximos, sugere-se que o número de variáveis latentes (18) utilizadas foi adequado, não havendo sobreajuste do modelo. O ajuste obtido para este modelo pode ser visualizado na Figura 7, com coeficiente de

correlação de 0,8709 e coeficiente de determinação de 0,7585. Esses resultados podem ser considerados adequados para estas previsões já que o método de referência apresentou ampla variação (VALDERRAMA; BRAGA; POPPI, 2007; VALDERRAMA; BRAGA; POPPI, 2007a; FERREIRA; PALLONE; POPPI, 2013).

A linearidade do modelo PLS foi avaliada através do gráfico de resíduos (Figura 7-A). De acordo com os resultados é possível afirmar que os dados se ajustam a um modelo linear, pois a distribuição dos resíduos se assemelha a uma distribuição aleatória (Figura 7-B).



**Figura 7:** (A) Ajuste e (B) Resíduos do modelo PLS para determinação do teor de clorofila. (x) Amostras de Calibração (●) Amostras de Validação.

A capacidade preditiva do modelo foi considerada satisfatória, pois, de acordo com a literatura, são considerados modelos de calibração de boa qualidade preditiva aqueles que apresentam valores de RPD acima de 2,4, e são classificados satisfatórios os que apresentam valores entre 2,4 e 1,5 (BOTELHO; MENDES; SENA, 2013). O valor de RPD do modelo de calibração (Tabela 4), foi de 1,6, podendo ser considerado adequado em relação ao valor de RPD para a validação desse parâmetro.

A sensibilidade é definida como a fração do sinal responsável pelo acréscimo de uma unidade de concentração à propriedade de interesse (VALDERRAMA, 2009). Em modelos de calibração multivariada baseados no método PLS a sensibilidade pode ser avaliada a partir do vetor de coeficientes de regressão. A sensibilidade analítica apresenta a sensibilidade do método em termos da unidade de concentração que é utilizada, sendo definida como a razão entre a sensibilidade e o desvio padrão do sinal de referência (VALDERRAMA,

2009), apresentando, portanto, a sensibilidade do método em termos de unidades de concentração.

O inverso da sensibilidade analítica (sensibilidade analítica<sup>-1</sup>) permite estabelecer a menor diferença de concentração entre amostras, que pode ser distinguida pelo método. Considerando-se que a faixa de teor de clorofila utilizada no trabalho, variou de 223 a 8.454 mg/kg (ppm), a sensibilidade ( $6,5675 \times 10^{-8}$ ) e sensibilidade analítica (0,0033) apresentaram resultados satisfatórios. O inverso da sensibilidade analítica permite estabelecer a menor diferença de concentração entre as amostras que pode ser distinguida pelo método. Sendo assim, é possível fazer a distinção de amostras com diferença de teor de clorofila da ordem de 300,65 ppm.

Os limites de detecção e quantificação do modelo foram de 992,15 ppm e 3006,50 ppm, respectivamente. Desta forma, a concentração mínima que pode ser medida confiavelmente pelo modelo empregando infravermelho foi de 992,15 ppm, indicando que amostras abaixo deste teor podem ser medidas erroneamente. Como o limite de quantificação ficou em 3006,50 ppm, significa que amostras que apresentem diferenças de concentração inferiores a esta não poderão ser diferenciadas.

Os resultados das figuras de mérito determinadas para os modelos PLS discutidos demonstram que modelos de calibração multivariada podem ser validados, sendo que os parâmetros estimados foram consistentes e de boa concordância com os teoricamente esperados.

De forma geral, os modelos construídos a partir de espectros NIR de amostras de soja e medidas com métodos padrões apresentaram uma correlação considerada satisfatória entre os espectros e os valores de referência, erros relativamente baixos. Entretanto, ainda que este trabalho tenha utilizado um número bastante expressivo de amostras, para melhorar a predição do modelo, faz-se necessário analisar um número ainda maior de amostras e com maior variabilidade dos resultados, aumentando a amplitude dos dados e provavelmente a robustez dos modelos.



## 6. CONCLUSÕES

Os resultados sugerem a possibilidade de utilização de modelos PLS para prever os parâmetros índice de acidez, lipídios totais, proteína bruta, clorofila e umidade de forma satisfatória. Desta forma, esta metodologia pode ser uma alternativa para determinação destes parâmetros em soja, trazendo como vantagens a rapidez, o caráter não destrutivo e não invasivo, o não emprego de reagentes/solventes, sem geração de resíduos tóxicos, sendo uma alternativa aos métodos convencionais, com potencial para ser implementada para medidas em linha de produção.

A previsão dos teores de clorofila demanda metodologias de referência que apresentem menor variação para que os modelos construídos para este fim apresentem respostas mais confiáveis. No entanto, o modelo construído se apresenta como adequado para a medida deste parâmetro considerando que as oscilações obtidas a partir do método de referência são significativamente extensas.

Os resultados deste estudo sugerem que a utilização de Espectroscopia no Infravermelho Próximo aliada a métodos quimiométricos pode ser uma interessante alternativa para o controle de qualidade de grãos de soja em indústrias de alimentos.

## 7. REFERÊNCIAS

ALBRECHT, R. et al. Efficiency of near-infrared reflectance spectroscopy to assess and predict the stage of transformation of organic matter in the composting process. **Bioresource Technology**, v. 99, n. 2, p. 448–455, 2008.

AOCS. **Official Methods and Recommended Practices of the AOCS**. 6th. ed. 2009.

ARAUJO, M. M. V. **Indicadores de qualidade em grãos de soja submetidos a diferentes condições de armazenamento**. 2014. 84 f. Dissertação (Mestrado em Agricultura Tropical). Universidade Federal de Mato Grosso, Faculdade de Agronomia e Medicina Veterinária, 2014.

ARAUJO, T. P. **Emprego de espectroscopia no infravermelho e métodos quimiométricos para a análise direta de tetraciclinas em leite bovino**. 2007. 99 f. Dissertação (Mestrado em Química Analítica) - Universidade Estadual de Campinas, Instituto de Química, Campinas, 2007.

BARTHUS, R. C. **Aplicação de métodos quimiométricos para análises de controle de qualidade de óleos vegetais utilizando espectroscopia no infravermelho e raman**. 1999. 112 f. Dissertação (Mestrado em Química) - Universidade Estadual de Campinas, Instituto de Química, Campinas, 1999.

BEVILACQUA, M. et al. Application of near infrared (NIR) spectroscopy coupled to chemometrics for dried egg-pasta characterization and egg content quantification. **Food Chemistry**, v. 140, n. 4, p. 726–734, 2013.

BORDIGNON, B. C. S. **Relação das condições de armazenamento com qualidade fisiológica de sementes e composição do óleo extraído de cultivares de soja**. 2009. 90 f. Tese (Doutorado em Agronomia) - Universidade Federal de Santa Maria, Centro de Ciências Rurais, Santa Maria, 2009.

BORRMANN, D. **Efeito do déficit hídrico em características químicas e bioquímicas da soja e na degradação da clorofila, com ênfase na formação de metabólitos incolores**. 2009. 125 f. Tese (Doutorado Ciências dos Alimentos) - Universidade de São Paulo, Faculdade de Ciências Farmacêuticas, São Paulo, 2009.

BOTELHO, B. G.; MENDES, B. A. P.; SENA, M. M. Implementation of a robust method for quality inspection control of artisanal minas cheese. A multivariate metrological approach. **Química Nova**, v. 36, n. 9, p. 1416–1422, 2013.

BRÁS, L. P. et al. Multiblock PLS as an approach to compare and combine NIR and MIR spectra in calibrations of soybean flour. **Chemometrics and Intelligent Laboratory Systems**, v. 75, n. 1, p. 91–99, 2005.

BRERETON, R. G. Introduction to multivariate calibration in analytical chemistry. **The Analyst**, v. 125, n. 11, p. 2125–2154, 2000.

CAVALCANTE, A. K.; DE SOUSA, L. B.; HAMAWAKI, O. T. Determination and Evaluation of Oil Content in Soybean Seeds by Nuclear Magnetic Resonance

- Methods and Soxhlet. **Bioscience Journal**, v. 27, n. 1, p. 8–15, 2011.
- CONAB- COMPANHIA NACIONAL DE ABASTECIMENTO. Acompanhamento da safra brasileira de grãos. v. 3, n. 4, p. 154, 2016.
- CUNHA, D. DA S. et al. Soja para consumo humano: Breve abordagem. **Agrarian Academy**, v. 2, n. 3, p. 101–113, 2015.
- DA SILVA, V. A. G. et al. Non-destructive identification of different types and brands of blue pen inks in cursive handwriting by visible spectroscopy and PLS-DA for forensic analysis. **Microchemical Journal**, v. 116, p. 235–243, 2014.
- DE MORAES, R. M. A. et al. Biochemical characterization of high protein soybean lines. **Pesquisa Agropecuaria Brasileira**, v. 41, n. 5, p. 725–729, 2006.
- DIAS, C. **Tecnologia inovadora analisa solos em apenas 30 segundos - Portal Embrapa**. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/17162564/tecnologia-inovadora-analisa-solos-em-apenas30segundos>>. Acesso em: 8 nov. 2016.
- DOS SANTOS, A. L. et al. Estudo do creseimento, desempenho, rendimento de carcaça e qualidade de carne de três linhagens de frango de corte. **Revista Brasileira de Zootecnia**, v. 34, n. 5, p. 1589–1598, 2005.
- ERNESTINA, F.; SABIN, G. P. Otimização de métodos de regressão multivariada para quantificação de sulfametoxazol e trimetoprima em medicamentos. **Revista Produção on line**, v. 32, n. 51, 2007.
- ESKALB. Determinação de nitrogênio e proteína bruta pelos métodos Kjeldahl e DUMAS. **Revista Analytica**, v. 73, p. 54, 2014.
- FARONI, L. R. A. . et al. Armazenamento de soja em silos tipo bolsa. **Journal of the Brazilian Association of Agricultural Engineering**, v. 29, n. 1, p. 91–100, 2009.
- FERREIRA, D. S. et al. Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples. **Food Control**, v. 35, n. 1, p. 227–232, 2014.
- FERREIRA, D. S.; PALLONE, J. A. L.; POPPI, R. J. Fourier transform near-infrared spectroscopy (FT-NIRS) application to estimate Brazilian soybean [Glycine max (L.) Merrill] composition. **Food Research International**, v. 51, n. 1, p. 53–58, 2013.
- FIGUEIREDO, J. **FT-NIR spectroscopy and Laser Diffraction particle sizing of APIs in Pharmaceutical formulations**. 2008. 72 f. Dissertação (Mestrado Integrado em Engenharia Química) - Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisboa., 2008.
- FODOR, M. et al. Development of a rapid, non-destructive method for egg content determination in dry pasta using FT-NIR technique. **Journal of Food Engineering**, v. 107, n. 2, p. 195–199, 2011.
- FRANÇA-NETO, J. B. . et al. Semente Esverdeada de Soja: Causas e Efeitos Sobre o Desempenho Fisiológico. **Circular Técnica 91 Embrapa**, p. 1–16, 2012.
- GARCÍA-REBOLLAR, P. et al. Composition and nutritive value of commercial

- soybean. **Animal Feed Science and Technology**, v. In Press, p. 17 p., 2016.
- GELADI, P.; KOWALSKI, B. Partial Least-Squares Regression - a Tutorial. **Analytica Chimica Acta**, v. 185, p. 1–17, 1986.
- GONÇALVES, L. C. et al. CHEMICAL COMPOSITION AND TECHNOLOGICAL PROPERTIES OF TWO SOYBEAN CULTIVARS. **BBR - Biochemistry and Biotechnology Reports**, v. 3, n. 1, p. 30–40, 2014.
- GREGGIO, E. A.; BONINI, E. A. Soybean Quality and Oil Acidity Rates. **Revista em Agronegócios e Meio Ambiente**, v. 7, n. 3, p. 645–658, 2014.
- HAUGHEY, S. A. et al. The application of Near-Infrared Reflectance Spectroscopy (NIRS) to detect melamine adulteration of soya bean meal. **Food Chemistry**, v. 136, n. 3–4, p. 1557–1561, 2013.
- HELLAND, I. S. On the structure of partial least squares regression. **Communications in Statistics - Simulation and Computation**, v. 17, n. 2, p. 581–607, 1988.
- HEMAN, A.; HSIEH, C.-L. Measurement of moisture content for rough rice by visible and near-infrared (NIR) spectroscopy. **Engineering in Agriculture, Environment and Food**, v. 9, n. 3, p. 280–290, jul. 2016.
- HIRAKURI, M. H.; LAZZAROTTO, J. J. O agronegócio da soja nos contextos mundial e brasileiro. **Documentos 349-Embrapa Soja.**, p. 37, 2014.
- HUANG, H. et al. Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. **Journal of Food Engineering**, v. 87, n. 3, p. 303–313, 2008.
- JACONI, A. **O uso da espectroscopia no infravermelho próximo na quantificação de carbono em solos sob o cultivo de cana-de-açúcar.** 2011. 53 f. Dissertação (Mestrado em Ciências) - Universidade de São Paulo, Centro de Energia Nuclear na Agricultura, Piracicaba, 2011.
- JAMRÓGIEWICZ, M. Application of the near-infrared spectroscopy in the pharmaceutical technology. **Journal of Pharmaceutical and Biomedical Analysis**, v. 66, p. 1–10, 2012.
- JOUAN-RIMBAUD, D. et al. Detection of prediction outliers and inliers in multivariate calibration. **Analytica Chimica Acta**, v. 388, n. 3, p. 283–301, 1999.
- KAIJSER, A.; DUTTA, P.; SAVAGE, G. Oxidative stability and lipid composition of macadamia nuts grown in New Zealand. **Food Chemistry**, v. 71, n. 1, p. 67–70, 2000.
- KENNARD, R. W.; STONE, L. A. Computer Aided Design of Experiments. **Technometrics**, v. 11, n. 1, p. 137–148, 1969.
- KLAHOLD, C. A. et al. Resposta da soja *Glycine max* (L.) Merrill à ação de bioestimulante. **Acta Scientiarum. Agronomy**, v. 28, n. 1972, p. 179–185, 2006.
- KOWALSKI, B. R.; BEEBE, K. R. An Introduction to Multivariate calibration and Analysis. **Analytical Chemistry**, v. 59, n. 17, p. 1007–1017, 1987.
- KRZYŻANOWSKI, F. C. et al. Controle de qualidade. **Circular Técnica**, 54 -

**Embrapa Soja**, p. 12, 2008.

LEITÃO, T. M. D. **Aplicações da espectroscopia de infravermelho próximo em Ciências Farmacêuticas**. 2012. 75 f. Dissertação (Mestrado em Ciências Farmacêuticas) - Universidade de Fernando Pessoa, Faculdade de Ciências da Saúde, Porto, 2012.

LEITE, D. C. **Mapeamento de QLT e análises de espectroscopia para óleo visando aplicação em programas de melhoramento de soja**. 2015. 85 f. Tese (Doutorado em Agronomia) - Universidade Estadual Paulista Julio de Mesquita Filho, Faculdade de Ciências Agrárias e Veterinárias, Jabotical, 2015.

LI, G. J. et al. QTL Analysis for Dynamic Expression of Chlorophyll Content in Soybean (*Glycine max* L. Merr.). **Acta Agronomica Sinica**, v. 36, n. 2, p. 242–248, 2010.

LIMA, A.; BAKKER, J. Near-infrared spectroscopy for monitoring peripheral tissue perfusion in critically ill patients. **Revista Brasileira de Terapia Intensiva**, v. 23, n. 3, p. 341–351, 2011.

LONG, D. S.; ENGEL, R. E.; SIEMENS, M. C. Measuring grain protein concentration with in-line near infrared reflectance spectroscopy. **Agronomy Journal**, v. 100, n. 2, p. 247–252, 2008.

LUYPAERT, J.; MASSART, D. L.; VANDER HEYDEN, Y. Near-infrared spectroscopy applications in pharmaceutical analysis. **Talanta**, v. 72, n. 3, p. 865–883, 2007.

MA, H. et al. Rapid authentication of starch adulterations in ultrafine granular powder of Shanyao by near-infrared spectroscopy coupled with chemometric methods. **Food Chemistry**, v. 215, p. 108–115, 2017.

MACIEL, V. G. et al. Greenhouse Gases Assessment of Soybean Cultivation steps in Southern Brazil. **Journal of Cleaner Production**, v. 131, p. 747–753, 2016.

MAGALHÃES, D. F. **Aplicações da espectroscopia de infravermelho próximo na monitorização de processos farmacêuticos**. 2014. 127 f. Dissertação (Mestrado em Química Tecnológica) - Universidade de Lisboa, Faculdade de Ciências, Departamento de Química e Bioquímica, Lisboa, 2014.

MANDARINO, J. M. G. Grãos verdes: influência na qualidade dos produtos à base de soja - Série Sementes. **Circular Técnica 90 Embrapa**, p. 1–5, 2012.

MARÇO, P. H. **Estudo da influência da radiação e pH no comportamento cinético de antocianinas de plantas do gênero Hibiscus por métodos quimiométricos**. 2009. 214 f. Tese (Doutorado) - Universidade Estadual de Campinas. Departamento de Química Analítica, Campinas, 2009.

MARTENS, H.; NAES, T. **Multivariate Calibration**. John Wiley ed. 1992.

MASSIE, D. R.; NORRIS, K. H. Spectral Reflectance and Transmittance Properties of Grain in the Visible and Near Infrared. **American Society of Agricultural and Biological Engineers**, v. 8, n. 4, 1965.

MELLO FILHO, O. L. et al. Grain yield and seed quality of soybean selected for

high protein content. **Pesquisa Agropecuária Brasileira**, v. 39, n. 5, p. 445–450, 2004.

MONTES, J. M. et al. Near-infrared spectroscopy on combine harvesters to measure maize grain dry matter content and quality parameters. **Plant Breeding**, v. 13, n. 6, p. 143–595, 2006.

NETO, B. DE B.; SCARMINIO, I. S.; BRUNS, R. E. 25 anos de quimiometria no Brasil. **Química Nova**, v. 29, n. 6, p. 1401–1406, 2006.

NUNES, A. L. S. **Quantificação de clorofila em óleo vegetal extraído de soja com diferentes percentuais de grãos imaturos**. 2013. 32 f. Trabalho de Conclusão de Curso. Universidade Tecnológica Federal do Paraná - Curso Superior de Tecnologia em Alimentos. Londrina, 2013.

OLIVEIRA, M. A. et al. Determinação do índice de acidez titulável dos grãos de soja colhidos na safra 2014/15. **Comissão de Pós-Colheita e Segurança Alimentar**, p. 239–242, 2016.

OTTO, M. **Chemometrics - Statistics and Computer application in Analytical Chemistry**. Wiley ed. Weinheim: 1999.

PAIVA, B. M.; ALVES, R. M.; HELENO, N. M. Aspecto socioeconômico da soja. **Informe Agropecuário**, v. 27, n. 230, p. 96, 2006.

PASQUINI, C. Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, v. 14, n. 2, p. 198–219, 2003.

PAULA, E. F. E. et al. Determinação da Qualidade da Carne com Uso da Espectroscopia de Reflectância. **Scientia Agraria Paranaensis**, v. 12, n. 4, p. 301–307, 2013.

PAULA, S. A. DE. **Composição bioquímica e fatores antinutricionais de genótipos de soja**. 2007. 88 f. Dissertação (Mestrado em Bioquímica Agrícola) - Universidade Federal de Viçosa, Viçosa, 2007.

PAULETTO, F. B.; FOGAÇA, A. O. Evaluation of the Centesimal Composition of Tofu and Okara. **Disciplinarum Scientia**, v. 13, n. 1, p. 85–95, 2012.

PÍPOLO, A. E. et al. Teores de óleo e proteína em soja: fatores envolvidos e qualidade para a indústria. **Comunicado Técnico 86 EMBRAPA**, p. 16, 2015.

POREP, J. U.; KAMMERER, D. R.; CARLE, R. On-line application of near infrared (NIR) spectroscopy in food production. **Trends in Food Science & Technology**, v. 46, n. 2, p. 211–230, 2015.

REICH, G. Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. **Advanced Drug Delivery Reviews**, v. 57, n. 8, p. 1109–1143, 2005.

RESTAINO, E.; FASSIO, A.; COZZOLINO, D. Discrimination of meat patés according to the animal species by means of near infrared spectroscopy and chemometrics Discriminación de muestras de paté de carne según tipo de especie mediante el uso de la espectroscopia en el infrarrojo cercano y la quimi. **CyTA - Journal of Food**, v. 9, n. 3, p. 210–213, 2011.

ROGGO, Y. et al. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. **Journal of Pharmaceutical and Biomedical Analysis**, v. 44, n. 3 SPEC. ISS., p. 683–700, 2007.

SAVITSKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627–1639, 1964.

SILVA, M. A. DE M. **Desenvolvimento e Validação de Método Quimiométrico para Determinação de Amoxicilina em Formulação Farmacêutica**. 2011. 66 f. Dissertação (Mestrado em Ciências Moleculares) - Universidade Estadual de Goiás, Unidade Universitária de Ciências Exatas e Tecnológicas, Anápolis, 2011.

SINNECKER, P. **Degradação da clorofila durante a maturação e secagem de sementes de soja**. 2002. 117 f. Tese (Doutorado em Ciências dos Alimentos) - Universidade de São Paulo, Faculdade de Ciências Farmacêuticas, São Paulo, 2002.

SOARES, T. A.; BIAGGIONI, M. A. M.; FRANÇA NETO, J. B. Análise de acidez graxa como índice de qualidade em grãos de soja. **Energia na Agricultura**, v. 20, n. 1, p. 91–102, 2005.

SOUZA, A. M. DE et al. Teaching experiment of chemometrics for multivariate calibration in determination of paracetamol in commercial tablets using Near-Infrared Spectroscopy: A tutorial, part II. **Quimima Nova**, v. 36, n. 7, p. 1057–1065, 2013.

STREIT, N. M. et al. As Clorofilas. **Ciência Rural**, v. 35, n. 3, p. 748–755, 2005.

TOSCAN, A. **Espectroscopia No Infravermelho E Método Leite Da Microrregião De Pato Branco – Pr Leite Da Microrregião De Pato Branco – Pr**. 2016. 107 f. Dissertação (Mestrado em Tecnologia de Processos Químicos e Bioquímicos) - Universidade Tecnológica Federal do Paraná, Pato Branco, 2016.

VALDERRAMA, L. et al. Espectroscopia Uv-Vis e Método Quimiométrico na Avaliação de Adulterações e Fraudes em Azeite de Oliva Extra Virgem. **Revista Brasileira de Pesquisa em Alimentos**, v. 5, n. 2, p. 32–40, 2014.

VALDERRAMA, P. **Avaliação de figuras de mérito em calibração multivariada na determinação de parâmetros de controle de qualidade em indústria alcooleira por espectroscopia**. 2005. 128 f. Dissertação (Mestrado em Química Analítica) - Universidade Estadual de Campinas, Instituto de Química, Campinas, 2005.

VALDERRAMA, P. **Calibração Multivariada de Primeira e Segunda Ordem e Figuras de Mérito na Quantificação de Enantiômeros por Espectroscopia**. 219 f. 2009. Tese (Doutorado em Química Analítica) - Universidade Estadual de Campinas. Instituto de Química, Campinas, 2009.

VALDERRAMA, P.; BRAGA, J. W. B.; POPPI, R. J. Variable Selection, Outlier Detection, and Figures of Merit Estimation in a Partial Least-Squares Regression Multivariate Calibration Model. A Case Study for the Determination of Quality Parameters in the Alcohol Industry by Near-Infrared Spectroscopy. **Journal of Agricultural and Food Chemistry**, v. 55, p. 8331–8338, 2007a.

VALDERRAMA, P.; BRAGA, J. W. B.; POPPI, R. J. Validation of multivariate calibration models in the determination of sugar cane quality parameters by near infrared spectroscopy. **Journal of the Brazilian Chemical Society**, v. 18, n. 2, p. 259–266, 2007b.

VALDERRAMA, P.; BRAGA, J. W. B.; POPPI, R. J. State of the art of figures of merit in multivariate calibration. **Química Nova**, v. 32, n. 5, p. 1278–1287, 2009.

VICARI, A. S. . **Qualidade de óleo de soja refinado embalado em PET (Polietileno Tereftalato) armazenado na presença e ausência de luz**. 2013. 68 f. Dissertação (Mestrado em Biotecnologia) - Universidade Católica Dom Bosco, Campo Grande, 2013.

WELLE, R. et al. Near-Infrared Spectroscopy on Chopper to Measure Maize Forage Quality Parameters Online. **Crop Science**, v. 43, p. 1407–1413, 2003.

ZARKADAS, C. G. et al. Assessment of the protein quality of fourteen soybean [Glycine max (L.) Merr.] cultivars using amino acid analysis and two-dimensional electrophoresis. **Food Research International**, v. 40, n. 1, p. 129–146, 2007.



## ANEXO I

**Quadro 1:** Valores de análises de referência das amostras de soja moída utilizados para a construção dos modelos de calibração e validação do teores de umidade, proteína bruta, lipídios totais, clorofila e índice de acidez.

<b>Amostras</b>	<b>Lipídios totais (%)</b>	<b>Proteína bruta (%)</b>	<b>Índice de Acidez (%)</b>	<b>Clorofila (ppm)</b>	<b>Umidade (%)</b>
A 01	20,38	36,89	0,48		8,02
A 02	21,48	34,49	0,58		5,86
A 03	21,42	35,37	0,74		5,83
A 04	22,06	34,94	0,60		4,36
A 05	22,15	35,99	0,80		7,95
A 06	23,40	35,83	0,53		8,11
A 07	22,65	34,76	0,63		8,09
A 08	19,82	36,82	0,45		4,51
A 09	22,24	36,32	0,72		7,46
A 10	21,41	36,09	1,11		8,37
A 11	22,04	36,82	0,56		5,06
A 12	22,10	35,00	0,51		5,62
A 13	20,27	39,95	0,71		8,50
A 14	21,67	34,85	0,60		6,64
A 15	22,35	34,93	0,54		6,27
A 16	22,83	34,96	0,98		5,06
A 17	21,24	34,97	0,70		7,42
A 18	22,68	36,03	0,87		6,31
A 19	22,22	36,12	0,42		6,87
A 20	23,59	35,50	0,72		5,24
A 21	22,29	35,15	0,39		2,13
A 22	22,34	35,55	0,48		3,13
A 23	21,02	38,08	0,53		3,91
A 24	25,02	36,36	0,61		3,55
A 25	23,84	35,36	1,13		4,61
A 26	22,37	35,39	0,75		3,01
A 27	22,28	35,48	0,93		3,74
A 28	20,48	36,70	0,85		4,05
A 29	22,20	35,61	0,84		4,26
A 30	21,12	35,91	0,54		3,26
A 31	22,14	36,28	0,47		2,80
A 32	22,48	36,06	0,46		3,03
A 33	20,70	36,08	0,86		3,98
A 34	20,09	39,01	1,00		2,93

A 35	19,25	39,85	1,30		4,60
A 36	22,55	35,00	0,48		3,56
A 37	20,35	36,78	0,53		2,81
A 38	24,30	38,35	1,79		3,00
A 39	22,83	38,64	1,62		3,04
A 40	23,58	35,12	0,62		1,72
A 41	22,36	35,04	0,54		4,64
A 42	23,03	35,29	0,59		5,02
A 43	23,39	38,65	1,72		4,76
A 44	22,98	36,88	0,33		2,72
A 45	20,95	37,61	0,22		2,56
A 46	16,04	37,73	0,57		3,02
A 47	21,53	35,60	0,65		4,17
A 48	20,42	34,41	0,51		2,71
A 49	22,32	35,85	0,94		5,30
A 50	20,76	38,37	1,30		4,23
A 51	21,76	33,98	0,72		2,67
A 52	21,23	34,64	0,77		2,51
A 53	18,42	37,06	1,00		3,55
A 54	23,27	36,67	1,06		3,62
A 55	23,03	35,60	1,53		1,75
A 56	20,38	34,29	0,72		3,07
A 57	21,52	35,57	1,02		3,45
A 58	21,32	36,29	1,37		4,11
A 59	21,85	35,21	0,76		2,17
A 60	21,16	36,65	0,98		3,27
A 61	20,88	38,52	1,15		2,48
A 62	22,13	35,58	0,72		2,08
A 63	19,93	34,57	0,48		3,44
A 64	21,87	33,92	0,64		3,02
A 65	21,40	37,95	0,86		2,38
A 66	21,05	37,29	0,93		2,63
A 67	25,45	35,79	0,56		2,66
A 68	24,88	35,15	0,93		3,44
A 69	22,53	35,68	0,64		1,63
A 70	22,10	37,85	1,40		2,97
A 71	23,82	36,24	0,73		1,17
A 72	21,93	40,11	1,74		1,95
A 73	20,73	33,70	0,67		6,09
A 74	25,11	33,88	0,72		5,19
A 75	20,82	35,97	0,80		5,80
A 76	20,53	35,32	0,42		2,64
A 77	20,37	35,67	0,33		2,68
A 78	22,23	34,78	0,89		4,27
A 79	21,44	35,33	0,56		3,04

A 80	21,84	35,43	0,80		4,41
A 81	21,31	34,64	0,51		4,75
A 82	20,08	36,82	0,98		3,60
A 83	21,83	35,96	0,49		3,41
A 84	22,78	34,98	0,59		2,91
A 85	22,37	37,06	1,12		2,43
A 86	23,34	35,61	0,47		2,90
A 87	23,27	37,29	0,63		1,98
A 88	23,34	34,59	0,71		3,62
A 89	20,08	36,91	1,30		2,79
A 90	23,31	38,30	1,01		2,06
A 91	20,97	35,34	1,14		3,55
A 92	22,77	35,61	0,50		4,41
A 93	21,74	34,81	0,41		3,31
A 94	19,34	35,76	0,57		1,48
A 95	22,32	34,26	0,72		3,65
A 96	20,72	35,51	1,96		3,49
A 97	23,50	35,29	0,79		3,76
A 98	21,77	36,03	0,56		1,30
A 99	21,52	35,23	0,58		4,07
A 100	21,51	35,85	0,93		1,71
A 101	22,21	35,50	0,57		3,20
A 102	23,15	37,77	1,11		3,67
A 103	23,38	37,89	1,03		1,65
A 104	20,94	35,34	0,90		1,47
A106	21,42	34,65	0,61		3,01
A 107	23,93	35,13	0,78		1,61
A 108	22,42	37,21	1,05		2,16
A 109	21,32	34,15	0,77		1,88
A 110	22,32	36,27	0,76		2,96
A 111	26,33	34,73	0,71		3,83
A 112	22,04	33,86	0,86		2,04
A 113	21,11	37,56	1,30		2,98
A 114	23,83	35,84	0,89		2,52
A 115	22,16	35,69	0,41		4,16
A 116	23,47	35,28	0,51		3,99
A 117	21,54	34,86	0,52		4,02
A 118	22,66	38,86	1,52		3,02
A 119	23,46	35,08	0,74		1,71
A 120	20,53	34,52	0,93		1,75
A 121	20,83	35,23	0,49		4,10
A 122	22,69	36,75	1,46		2,29
A 123	23,95	38,52	1,70		4,68
A 124	21,31	34,39	1,19		4,12
A 125	21,23	35,37	0,76		2,67

A 126	21,58	35,56	0,81		3,05
A 127	21,52	35,23	1,00		3,21
A 128	21,33	35,70	0,92	2030,00	3,77
A 129	19,87	36,15	0,55	310,00	2,55
A 130	20,76	36,57	0,31	318,00	0,85
A 131	20,10	35,87	0,35	470,00	1,50
A 132	19,85	33,86	0,41	336,00	3,59
A 133	20,87	36,42	0,44	473,00	1,56
A 134	21,17	35,54	0,32	661,00	2,55
A 135	19,28	37,65	0,41	324,00	5,59
A 136	19,22	37,16	0,62	831,00	2,93
A 137	19,48	36,72	0,68	2058,00	3,38
A 138	20,49	35,46	0,82	2763,00	5,33
A 139	20,41	36,39	0,53	1735,00	3,49
A 140	20,24	35,87	0,51	589,00	4,58
A 141	19,64	33,91	0,53	584,00	3,58
A 142	21,78	37,22	1,88	2068,00	2,27
A 143	20,10	35,87	0,35	470,00	3,72
A 144	20,53	35,91	0,39	852,00	4,23
A 145	22,20	33,90	0,50	790,00	3,87
A 146	19,55	35,82	0,52	915,00	4,25
A 147	19,89	34,94	0,47	352,00	2,42
A 148	21,32	36,89	0,68	2019,00	9,56
A 149	23,05	36,06	0,53	5050,00	9,74
A 150	21,22	35,35	0,58	622,00	8,74
A 151	20,88	36,27	0,46	526,00	8,80
A 152	22,16	35,76	0,64	3555,00	8,41
A 153	21,51	36,58	0,87	1185,00	8,29
A 154	21,08	36,57	0,62	1125,00	9,16
A 155	20,21	36,18	0,44	223,00	8,18
A 156	21,11	36,41	0,48	608,00	7,60
A 157	21,65	36,43	0,50	345,00	7,99
A 158	23,10	36,38	0,87	5510,00	8,85
A 159	21,38	37,40	0,51	1979,00	7,21
A 160	19,73	35,40	0,60	2084,00	5,54
A 161	23,73	37,55	1,54	2083,00	5,79
A 162	19,01	34,86	0,68	4128,00	6,71
A 163	19,96	35,70	0,61	2073,00	5,87
A 164	20,54	36,49	1,05	1032,00	8,42
A 165	20,53	33,64	0,44	418,00	6,61
A 166	19,19	36,17	0,39	437,00	5,42
A 167	19,01	34,86	0,68	1754,00	7,49
A 168	21,38	36,37	1,54	2750,00	6,29
A 169	21,72	34,11	0,86	1776,00	6,96
A 170	25,75	34,27	0,61	2073,00	7,19

A 171	25,61	35,40	0,86	1224,00	5,00
A 172	24,16	36,95	1,48	1610,00	5,03
A 173	24,89	34,70	0,62	525,00	5,91
A 174	25,18	36,53	1,05	1720,00	4,44
A 175	24,65	33,09	0,44	380,00	6,23
A 176	24,16	34,78	0,52	472,00	5,96
A 177	23,69	34,30	0,60	467,00	4,92
A 178	24,68	36,38	1,08	2838,00	4,14
A 179	24,99	37,30	1,30	1023,00	3,85
A 180	21,54	37,45	0,58	681,00	6,33
A 181	20,47	37,90	1,46	1005,00	5,05
A 182	22,41	36,50	1,03	1995,00	5,78
A 183	25,27	37,16	1,48	2913,00	3,96
A 184	23,96	37,09	2,20	2137,00	6,30
A 185	20,97	36,23	1,09	2018,00	7,39
A 186	21,94	37,44	1,34	2287,00	6,53
A 187	22,68	35,16	1,10	1163,00	5,94
A 188	22,28	36,47	1,24	2490,00	4,81
A 189	22,12	37,39	1,20	1636,00	6,29
A 190	22,73	35,37	1,03	1495,00	4,56
A 191	22,08	35,41	1,03	1942,00	4,19
A 192	22,33	37,10	1,09	1518,00	5,43
A 193	22,02	37,38	1,40	1871,00	5,24
A 194	21,27	35,64	0,84	1360,00	5,90
A 195	22,03	34,25	1,54	2322,00	4,80
A 196	21,32	35,86	0,78	1267,00	5,12
A 197	21,79	37,12	1,19	2083,00	5,26
A 198	21,94	37,91	1,99	2331,00	4,87
A 199	21,60	36,72	1,32	1952,00	4,53
A 201	20,90	36,09	0,51	1867,00	5,57
A 202	20,81	34,32	0,61	1119,00	5,08
A 203	22,29	35,79	1,44	2923,00	5,17
A 204	21,72	35,35	1,38	2122,00	6,78
A 205	20,69	36,56	0,55	1614,00	6,91
A 206	21,73	35,49	0,60	2299,00	4,18
A 207	20,50	34,18	1,00	4712,00	5,66
A 208	21,26	35,48	0,66	6088,00	6,84
A 209	21,56	33,78	0,47	388,00	3,58
A 210	21,60	35,71	0,67	5477,00	6,01
A 211	18,55	35,62	0,30	396,00	5,38
A 212	22,23	35,79	0,66	5563,00	5,45
A 214	20,64	37,72	1,31	1799,00	4,23
A 217	19,58	37,30	1,38	2242,00	3,59
A 218	19,92	37,08	0,95	3271,00	6,47
A 219	21,26	36,25	0,53	1027,00	5,84

A 220	19,71	35,69	0,71	2334,00	6,89
A 221	19,86	37,27	0,47	2196,00	6,37
A 222	22,06	35,91	0,66	8454,00	6,19
A 223	21,48	35,74	0,54	5311,00	5,61
A 224	20,28	36,78	1,34	1215,00	8,59
A 225	21,86	37,74	1,31	2194,00	8,76
A 226	23,41	33,93	0,57	2100,00	6,32
A 227	22,59	36,96	1,70	1650,00	8,25
A 228	24,69	37,46	1,64	2135,00	8,31
A 229	22,47	34,16	0,63	733,00	8,61
A 230	22,91	36,68	0,98	5094,00	7,41
A 231	23,11	34,41	0,47	404,00	5,24
A 232	21,62	36,95	1,23	2287,00	7,54
A 233	24,48	37,83	1,90	2341,00	7,31
A 234	22,91	36,63	0,58	3641,00	5,86
A 235	24,17	34,69	0,43	375,00	3,66
A 236	23,18	35,54	0,51	300,00	3,45
A 237	22,05	35,83	1,08	1587,00	7,66
A 238	23,15	37,31	0,57	1030,00	2,70
A 239	23,32	34,27	0,44	1004,00	4,20
A 240	24,72	37,20	1,36	2807,00	9,14
A 241	24,18	36,84	2,38	2880,00	2,85
A 242	23,31	35,92	0,59	6814,00	4,81
A 243	23,73	37,50	4,36	3035,00	7,80
A 244	24,66	36,62	1,49	872,00	7,33
A 245	25,08	36,90	1,17	1323,00	7,81
A 246	24,48	37,52	1,57	2849,00	7,05
A 247	23,21	36,50	1,25	697,00	2,88
A 248	23,95	38,10	1,81	2044,00	6,28
A 249	22,18	35,43	1,06	865,00	5,18
A 250	20,33	36,82	1,08	3697,00	5,26
A 251	20,16	36,19	1,29	3379,00	6,57
A 252	22,76	35,97	0,58	1717,00	4,51
A 253	21,86	37,30	2,28	2606,00	6,95
A 254	23,05	34,61	0,46	816,00	5,71
A 255	18,00	34,29	1,08	1379,00	5,59
A 256	22,09	36,69	1,49	1893,00	5,08
A 257	22,56	34,88	2,48	2249,00	6,81
A 258	22,92	34,69	0,51	530,00	5,06
A 259	21,11	37,01	1,38	3261,00	5,13
A 260	22,30	35,63	0,76	4083,00	6,81
A 261	22,32	35,79	2,14	1348,00	5,49
A 263	21,34	37,62	1,49	2420,00	6,66
A 263	20,63	35,69	0,51	451,00	3,69
A 264	23,21	36,77	0,86	2079,00	4,94

A 265	22,66	38,95	2,21	1189,00	4,31
A 266	21,65	36,22	0,43	275,00	4,98
A 267	22,16	35,29	0,61	2481,00	4,32
A 268	25,64	36,32	0,38	939,00	5,14
A 269	24,29	35,68	3,11	3694,00	4,04
A 270	21,79	36,24	0,60	1170,00	5,01
A 271	22,80	36,10	0,90	4185,00	5,47
A 272	19,68	37,72	0,54	264,00	4,76
A 273	21,15	36,81	0,97	1751,00	4,88
A 274	20,63	35,69	0,51	451,00	4,23
A 275	22,61	36,64	1,44	1904,00	4,35
A 276	22,70	35,11	0,83	2375,00	4,38
A 277	20,90	35,93	0,61	5588,00	4,50
A 278	20,73	36,68	1,50	2208,00	4,73
A 279	21,29	36,71	2,05	2791,00	7,57
A 280	21,20	35,86	0,94	3016,00	8,11
A 281	22,55	37,14	1,04	4016,00	7,20
A 282	20,76	37,87	0,81	1688,00	8,30
A 283	21,25	36,87	2,03	1534,00	6,95
A 284	21,89	35,61	1,57	3289,00	7,48
A 285	21,31	36,17	1,69	1624,00	4,85
A 286	19,41	35,15	0,68	702,00	4,09
A 287	22,63	35,46	0,55	353,00	8,15
A 288	22,76	34,65	0,38	342,00	9,50
A 289	22,17	35,32	0,80	2553,00	8,06
A 290	21,06	34,93	0,72	1642,00	7,35
A 291	18,36	35,69	0,50	1291,00	4,02
A 292	19,83	36,17	0,44	1208,00	3,66
A 293	22,89	35,45	0,60	1472,00	8,10
A 294	21,72	36,87	0,90	3485,00	6,92
A 295	21,83	36,61	0,59	1380,00	8,42
A 296	21,89	36,52	0,79	2451,00	8,52
A 297	22,43	35,54	0,70	3367,00	6,64
A 298	22,70	38,31	0,76	827,00	7,85
A 299	22,28	36,74	1,64	2334,00	4,66
A 300	20,30	35,97	0,90	3802,00	4,02
A 301	20,82	35,96	0,62	358,00	7,74