

UNIVERSIDADE FEDERAL DO PARANÁ

LEOMARA FLORIANO RIBEIRO

USO DE ESPECTROSCOPIA NO INFRAVERMELHO E ANÁLISE MULTIVARIADA
PARA PREVISÃO DE ÁCIDOS GRAXOS EM LINHAÇA DOURADA E MARROM

CURITIBA
2012

LEOMARA FLORIANO RIBEIRO

USO DE ESPECTROSCOPIA NO INFRAVERMELHO E ANÁLISE MULTIVARIADA
PARA PREVISÃO DE ÁCIDOS GRAXOS EM LINHAÇA DOURADA E MARROM

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Alimentos, Departamento de Engenharia Química, Setor de Tecnologia, Universidade Federal do Paraná, como requisito parcial à obtenção do Título de Mestre em Engenharia de Alimentos.

Orientador: Prof. Dr. Adauto Bellarmino de Pereira Netto

CURITIBA
2012

R484u Ribeiro, Leomara Floriano
 Uso de espectroscopia no infravermelho e análise multivariada para previsão
 de ácidos graxos em linhaça dourada e marrom. [manuscrito] / Leomara
 Floriano Ribeiro. – Curitiba, 2012.
 118f. . : il. [algumas color.] ; 30 cm.

 Impresso.
 Dissertação (mestrado) - Universidade Federal do Paraná, Setor de
 Tecnologia, Programa de Pós-graduação em Engenharia de Alimentos, 2012.

 Orientador: Aducto Bellarmino de Pereira Netto.

 1. Linhaça. 2. Ácidos graxos. I. Universidade Federal do Paraná. II. Pereira
 Netto, Aducto Bellarmino de. III. Título.

CDD: 613.26

LEOMARA FLORIANO RIBEIRO

**USO DE ESPECTROSCOPIA NO INFRAVERMELHO E ANÁLISE
MULTIVARIADA PARA PREVISÃO DE ÁCIDOS GRAXOS EM
LINHAÇA DOURADA E MARROM**

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Engenharia de Alimentos, da Universidade Federal do Paraná, pela Comissão formada pelos professores:



Orientador: Prof. Dr. ADAUCTO BELLARMINO DE PEREIRA NETTO
Setor de Ciências Biológicas, UFPR



Prof. Dr. CÉLIO PASQUINI
Instituto de Química, UNICAMP



Prof^a. Dr^a. BEATRIZ HELENA LAMEIRO DE NORONHA SALES MAIA
Setor de Ciências Exatas, UFPR

Curitiba, 06 de março de 2012.

Dedico este trabalho à minha família pelo apoio, amor incondicional e compreensão em todos os momentos.

AGRADECIMENTOS

Agradeço primeiramente a Deus por estar presente em todos os momentos da minha vida iluminando o meu caminho, em especial durante estes dois últimos anos.

Aos meus pais, Leonildo e Maria, meus irmãos Rodrigo e Lucas, que de uma forma especial e carinhosa sempre me apoiaram em todos os momentos e, principalmente, me deram força e coragem nos momentos de dificuldades.

Ao Prof. Dr. Aducto Bellarmino de Pereira Netto, pela orientação, apoio e pela confiança em mim depositada durante a realização deste trabalho.

Ao Prof. Dr. Patricio Guillermo Peralta-Zamora e Profa. Dra. Noemi Nagata, pelas sugestões apresentadas, pelos conhecimentos transmitidos, por terem estado sempre à disposição para esclarecer todas as minhas dúvidas e pela amizade.

À Profa. Dra. Beatriz Helena Lameiro de Noronha Sales Maia por transmitir seus conhecimentos e disponibilizar material necessário para realização das análises preliminares no Laboratório de Produtos Naturais e Ecologia Química, Departamento de Química - UFPR.

À Profa. Dra. Silvana Nisgoski pela colaboração nas análises de infravermelho no Laboratório de Anatomia e Identificação da Madeira, Departamento de Engenharia Industrial Madeireira – UFPR.

Ao Prof. Dr. Luiz Pereira Ramos e ao Prof. Dr. Arion Zandoná Filho pelas análises de Cromatografia Gasosa acoplada a Espectrometria de Massas.

Aos demais professores que colaboraram na minha formação e contribuíram de alguma forma com a realização deste trabalho.

As empresas Jasmine e Vitao que gentilmente doaram as amostras utilizadas neste trabalho.

Ao MSc Tecnologia de Alimentos Eriel Forville de Andrade por sempre me auxiliar com seus conhecimentos técnicos e disponibilizar material necessário para as análises em Cromatografia Gasosa.

Aos técnicos de laboratório, Marcelo Zadorecki e Patrícia Kanczewski Iwankiw pela amizade, disponibilidade, cooperação e paciência no decorrer desse trabalho.

À MSc Química Thalita Gilda Santos, pelas análises preliminares de Cromatografia Gasosa acoplada a Espectrometria de Massas.

Ao secretário do Programa de Pós-Graduação em Engenharia de Alimentos – UFPR, Paulo Krainski pelos serviços prestados.

Aos colegas do Programa de Pós-Graduação em Engenharia de Alimentos, em especial aos queridos amigos que conquistei durante este período, Bogdan, Fabiane, Flávia, Larize, Luana, Marlene, Valesca e Vitor, obrigada pela amizade, incentivo e momentos de descontração.

Ao Programa de Pós-Graduação em Engenharia de Alimentos (PPGEAL), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Fundo Paraná – SETI, Governo do Estado do Paraná, pelo auxílio financeiro.

A todos aqueles cujos nomes não foram citados, mas que de alguma forma contribuíram para que esse trabalho fosse realizado.

Muito Obrigada!!!

*"Talvez não tenhamos conseguido fazer o melhor,
mas lutamos para que o melhor fosse feito.
Não somos o que deveríamos ser,
não somos o que iremos ser,
mas, Graças à Deus,
não somos o que éramos antes."*

Martin Luther King

RESUMO

A linhaça, semente do linho (*Linum usitatissimum* Linaceae) contém elevado teor dos ácidos graxos poliinsaturados oleico, linoleico e linolênico, os quais são considerados essenciais, por não serem sintetizados pelo organismo humano. Todavia, quando ingeridos através da alimentação contribuem para a prevenção de doenças cardiovasculares e para regulação hormonal. A identificação e quantificação de ácidos graxos presentes na linhaça é tradicionalmente realizada através de cromatografia gasosa. No entanto, esta técnica é demorada e requer procedimentos de extração prévia dos ácidos graxos. Análises de espectroscopia no infravermelho têm sido empregadas como ferramentas analíticas alternativas rápidas, seletivas, de simples operação e baixo custo. O objetivo deste estudo foi avaliar a possibilidade de utilização de análises de infravermelho, em substituição a análises de cromatografia gasosa, para quantificação dos ácidos graxos oleico, linoleico e linolênico, em sementes e farinhas de linhaça das variedades dourada e marrom. Análise de componentes principais (PCA) e processos de calibração multivariada, particularmente a regressão de mínimos quadrados parciais (PLSR), foram utilizados com o objetivo de desenvolver modelos para validação das determinações dos ácidos graxos oleico, linoleico e linolênico, através de espectroscopia no infravermelho, com Transformada de Fourier, no modo de reflectância difusa. Os dados de espectroscopia infravermelho foram utilizados na faixa entre 9.000 e 4.000 cm^{-1} para o infravermelho próximo (NIR), e de 4.000 e 750 cm^{-1} para o infravermelho médio (MID). Os modelos de regressão para previsão do teor de ácidos graxos em sementes apresentaram coeficientes de determinação mais elevados na região do NIR, com médias de erro padrão relativo de 0,80%, 1,61% e 0,63%, respectivamente, para os ácidos graxos oleico, linoleico e linolênico. Para as farinhas, a região do NIR forneceu modelos de regressão com maiores coeficientes de determinação ($R^2_{\text{val}} = 0,99$) tanto para o ácido oleico quanto para o ácido linolênico, com erro padrão relativo de 1,02% e 1,21%, respectivamente. Já para o teor de ácido linoleico o maior coeficiente de determinação ($R^2_{\text{val}} = 0,88$) foi obtido na região do infravermelho médio, sendo que o erro padrão médio relativo foi de 0,76%. Estes dados demonstram que análises de espectroscopia de infravermelho podem ser utilizadas para predição acurada do teor de ácidos graxos em sementes e farinhas de linhaça. Estes dados ainda demonstram que análises de infravermelho podem ser utilizadas como técnicas alternativas eficazes, rápidas e com menor custo para determinação de ácidos graxos em linhaça, em substituição a tradicionalmente utilizada cromatografia gasosa.

Palavras-chave: *Linum usitatissimum*. Ácidos graxos. Infravermelho médio. Infravermelho próximo. Análise de Componentes Principais. Regressão por Mínimos Quadrados Parciais.

ABSTRACT

Flaxseeds, the seeds from *Linum usitatissimum* (Linaceae) contain high amount of the polyunsaturated oleic, linoleic and linolenic fatty acids, which are considered essential for not being synthesized by the human body. However, when ingested through the diet, these fatty acids help to prevent cardiovascular diseases and also help on hormonal regulation. Identification and quantification of fatty acids in flaxseeds are traditionally performed by gas chromatography. However, this technique is time consuming and requires prior extraction of fatty acids. Infrared spectroscopy analysis has been used as an alternative, selective and fast analytic tool simple operation and low cost. The aim of this study was to evaluate the possibility of using infrared analysis for the replacement of gas chromatography analysis to quantify the fatty acids oleic, linoleic and linolenic acid in flaxseeds and flax seed flours of yellow and brown varieties. Principal component analysis (PCA) and multivariate calibration, particularly the partial least squares regression (PLSR), were used in order to develop models to validate determinations of the oleic, linoleic and linolenic fatty acids by infrared spectroscopy with Fourier transform, in the diffuse reflectance mode. The infrared spectroscopy data were acquired in the range between 9.000 and 4.000 cm^{-1} to the near infrared (NIR), and 4.000 cm^{-1} and 750 to the mid-infrared (MIR). Regression models to predict the fatty acid content in flaxseeds showed higher correlation coefficients in the NIR, with average relative standard error of 0,80%, 1,61% and 0,63%, respectively for the oleic, linoleic and linolenic fatty acids. For flax seed flours, the NIR region of the regression models provided with the highest correlation coefficients ($R^2_{\text{val}} = 0,99$) both as the oleic acid to linolenic acid, with relative standard error of 1.02% and 1.21% respectively. For the linoleic acid content, the highest correlation coefficient ($R^2_{\text{val}} = 0,88$) was obtained in the mid infrared region, and the relative standard error was on average 0,76%. These data demonstrated that infrared spectroscopy analysis can be used for accurate prediction of the fatty acid contents in flaxseeds and flax seed flours. These data also demonstrate that infrared analysis can be used as an efficient, alternative, fast and less expensive technique for the determination of fatty acids in flaxseeds and flax seed flours, replacing the traditionally used gas chromatography.

Keywords: *Linum usitatissimum*. Fatty acids. Mid infrared. Near infrared. Principal Component Analysis. Partial Least Squares Regression.

LISTA DE ILUSTRAÇÕES

FIGURA 1 – LINHO SECCIONADO: FLORES, CAULE, CACHOPAS E SEMENTES.....	24
FIGURA 2 – CONVERSÃO DO SDG EM ENTERODIOL E ENTEROLACTONA POR AÇÃO BACTERIANA.....	27
FIGURA 3 – ESTRUTURA QUÍMICA DOS ÁCIDOS OLEICO E ELAÍDICO.....	31
FIGURA 4 – ESTRUTURA QUÍMICA DA FAMÍLIA ÔMEGA REPRESENTADA PELO ÁCIDO OLEICO, LINOLEICO E ÁCIDO LINOLÊNICO.....	32
FIGURA 5 – METABOLISMO DOS ÁCIDOS GRAXOS LINOLEICO E LINOLÊNICO.....	33
FIGURA 6 – DIFERENTES FORMAS DE VIBRAÇÕES MOLECULARES DE DEFORMAÇÃO.....	36
FIGURA 7 – REPRESENTAÇÃO ESQUEMÁTICA DE REFLECTÂNCIA DIFUSA NO INFRAVERMELHO EM AMOSTRA SÓLIDA.....	38
FIGURA 8 – DIAGRAMA DE CONSTRUÇÃO DO MODELO DE CALIBRAÇÃO.....	43
FIGURA 9 – ORGANIZAÇÃO DOS DADOS PARA CALIBRAÇÃO MULTIVARIADA.....	44
FIGURA 10 – A: GRÁFICO TRIDIMENSIONAL ILUSTRANDO OS EIXOS DAS COMPONENTES PRINCIPAIS e B: GRÁFICO BIDIMENSIONAL DA PC1 X PC2: ESCORES REPRESENTADOS POR (___) E <i>LOADINGS</i> POR (.....)	45
FIGURA 11 – CROMATÓGRAFO A GÁS (SHIMADZU GC 14B).....	56
FIGURA 12 – ESPECTROFOTOMETRO FTNIR/MID.....	57
FIGURA 13 – CROMATOGRAMA DOS PRINCIPAIS ÉSTERES METÍLICOS DE ÁCIDOS GRAXOS OBTIDOS DE AMOSTRA DE FARINHA DE LINHAÇA DOURADA (FLDA 21) POR CROMATOGRAFIA GASOSA ACOPLADA À ESPECTROMETRIA DE MASSAS.....	59
FIGURA 14 – CROMATOGRAMA DOS PRINCIPAIS ÉSTERES METÍLICOS DE ÁCIDOS GRAXOS OBTIDOS DE AMOSTRA DE FARINHA DE LINHAÇA DOURADA (FLDA 21) POR CROMATOGRAFIA GASOSA COM DETECTOR DE IONIZAÇÃO DE CHAMA.....	60
FIGURA 15 – ESPECTROS NIR PARA 78 AMOSTRAS DE LINHAÇA.....	65
FIGURA 16 – GRÁFICO TRIDIMENSIONAL DOS ESCORES DAS COMPONENTES PRINCIPAIS PC1 x PC2 x PC3 NA REGIÃO DO NIR.....	66
FIGURA 17 – GRÁFICO DE ESCORES DA PC1 x PC3 NA REGIÃO NIR.....	67
FIGURA 18 – GRÁFICO DE <i>LOADINGS</i> PARA A PC1 NA REGIÃO NIR.....	68
FIGURA 19 – GRÁFICO DE <i>LOADINGS</i> PARA A PC 3 NA REGIÃO NIR.....	69

FIGURA 20 – ESPECTROS MID PARA 78 AMOSTRAS DE LINHAÇA.....	70
FIGURA 21 – GRÁFICO TRIDIMENSIONAL DOS ESCORES DAS COMPONENTES PRINCIPAIS PC1 x PC2 x PC3 NA REGIÃO DO MID.....	71
FIGURA 22 – GRÁFICO DE ESCORES DA PC2 x PC3 NA REGIÃO MID.....	72
FIGURA 23 – GRÁFICO DE <i>LOADINGS</i> PARA A PC2 NA REGIÃO MID.....	73
FIGURA 24 – GRÁFICO DE <i>LOADINGS</i> PARA A PC3 NA REGIÃO MID.....	74
FIGURA 25 – ESPECTROS NO INFRAVERMELHO NA REGIÃO NIR E MID PARA FARINHAS DE LINHAÇA.....	76
FIGURA 26 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	78
FIGURA 27 – GRÁFICOS DE <i>LEVERAGE</i> x RESÍDUO DE <i>STUDENT</i> OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	80
FIGURA 28 – GRÁFICO DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	80
FIGURA 29 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	81
FIGURA 30 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	83
FIGURA 31 – GRÁFICOS DE <i>LEVERAGE</i> x RESÍDUO DE <i>STUDENT</i> OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	84
FIGURA 32 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	85
FIGURA 33 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	85
FIGURA 34 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	87

FIGURA 35 – GRÁFICOS DE <i>LEVERAGE</i> x RESÍDUO DE <i>STUDENT</i> OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	88
FIGURA 36 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHAS DE LINHAÇA.....	89
FIGURA 37 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA.....	90
FIGURA 38 – ESPECTROS NO INFRAVERMELHO NA REGIÃO NIR E MID PARA SEMENTES DE LINHAÇA.....	91
FIGURA 39 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	93
FIGURA 40 – GRÁFICOS DE <i>LEVERAGE</i> x RESÍDUO DE <i>STUDENT</i> OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	94
FIGURA 41 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	95
FIGURA 42 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	96
FIGURA 43 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	98
FIGURA 44 – GRÁFICOS DE <i>LEVERAGE</i> x RESÍDUO DE <i>STUDENT</i> OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	98
FIGURA 45 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	99
FIGURA 46 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	100

FIGURA 47 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	102
FIGURA 48 – GRÁFICOS DE <i>LEVERAGE</i> x RESÍDUO DE <i>STUDENT</i> OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	102
FIGURA 49 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	103
FIGURA 50 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA.....	104

LISTA DE TABELAS

TABELA 1 – COMPOSIÇÃO QUÍMICA DA SEMENTE DE LINHAÇA.....	25
TABELA 2 – CONCENTRAÇÃO DOS ÁCIDOS OLEICO, LINOLEICO E LINOLÊNICO (g/100g) EM ALIMENTOS DE ORIGEM VEGETAL.....	34
TABELA 3 – REGIÕES ESPECTRAIS DO INFRAVERMELHO.....	35
TABELA 4 – REGIÕES ESPECTRAIS NO MID E NIR RELACIONADAS ÀS LIGAÇÕES QUÍMICAS DE DIVERSOS GRUPOS FUNCIONAIS.....	37
TABELA 5 – CODIFICAÇÃO DAS AMOSTRAS DE LINHAÇA DOURADA E MARROM.....	53
TABELA 6 – TEORES DE ÁCIDO OLEICO, ÁCIDO LINOLEICO E ÁCIDO LINOLÊNICO OBTIDOS A PARTIR DAS AMOSTRAS DE LINHAÇA DOURADA E MARROM (continua).....	61
TABELA 6 – TEORES DE ÁCIDO OLEICO, ÁCIDO LINOLEICO E ÁCIDO LINOLÊNICO OBTIDOS A PARTIR DAS AMOSTRAS DE LINHAÇA DOURADA E MARROM (conclusão).....	62
TABELA 7 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO OLEICO NO NIR E MID EM FARINHA DE LINHAÇA.....	77
TABELA 8 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO OLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA FARINHA DE LINHAÇA.....	82
TABELA 9 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLEICO NO NIR E MID EM FARINHA DE LINHAÇA.....	82
TABELA 10 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA FARINHA DE LINHAÇA.....	86
TABELA 11 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLÊNICO NO NIR E MID EM FARINHA DE LINHAÇA.....	87
TABELA 12 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLÊNICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA FARINHA DE LINHAÇA.....	90
TABELA 13 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO OLEICO NO NIR E MID EM SEMENTE DE LINHAÇA.....	92
TABELA 14 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO OLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA SEMENTE DE LINHAÇA.....	96
TABELA 15 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLEICO NO NIR E MID EM SEMENTE DE LINHAÇA.....	97

TABELA 16 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA SEMENTE DE LINHAÇA.....	100
TABELA 17 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLÊNICO NO NIR E MID EM SEMENTE DE LINHAÇA.....	101
TABELA 18 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLÊNICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA SEMENTE DE LINHAÇA.....	104

LISTA DE ABREVIATURAS

AOCS - *American Oil Chemistry Society*

CG/DIC – Cromatografia gasosa com detector de ionização de chama

CG/EM – Cromatografia gasosa acoplada à espectrometria de massas

CV – Coeficiente de variação

DCM – Dados centrados na média

DHA – Docosaheptaenoico

EPA – Eicosapentaenoico

FAO – *Food and Agriculture Organization*

FAR – Região de infravermelho distante

FLDA – Farinha de linhaça dourada estabilizada – EMPRESA A

FLDB – Farinha de linhaça dourada – EMPRESA B

FLMB – Farinha de linhaça marrom – EMPRESA B

IUPAC - União Internacional de Química Pura e Aplicada

IV – Infravermelho

LD – Limite de Detecção

LQ – Limite de Quantificação

M_i – Número de matrizes

MID – Região de infravermelho médio

MSC – Correção do Espalhamento Multiplicativo

NIR – Região de infravermelho próximo

NIPLS – *Nonlinear Iterative Partial Least Squares*

NIST – *National Institute of Standards and Technology*

p – *Loadings*

PC – Componente principal

PCA – Análise de componentes principais

PLSR – Regressão por mínimos quadrados parciais

R^2_{cal} – Coeficiente de determinação de calibração

R^2_{val} – Coeficiente de determinação de validação

RMSEC – Raiz quadrada do erro médio de calibração

RMSECV – Raiz quadrada do erro médio de validação cruzada

RMSEP – Raiz quadrada do erro médio de previsão

SDG – Secoisolariciresinol diglicosídeo

SDOA – Semente de linhaça dourada orgânica – EMPRESA A

SLDA - Semente de linhaça dourada – EMPRESA A

SLDB - Semente de linhaça dourada – EMPRESA B

SLMB – Semente de linhaça marrom – EMPRESA B

SNV – *Standard Normal Variate*

t – Escores

USDA – *United States Department of Agriculture*

UV-VIS – Região ultravioleta e visível

VLs – Número de variáveis latentes

X – Matriz de dados espectrais (variáveis independentes)

Y – Matriz de dados experimentais (variáveis dependentes)

SUMÁRIO

1	INTRODUÇÃO	20
2	REVISÃO DE LITERATURA	23
2.1	LINHAÇA	23
2.1.1	<i>Histórico do cultivo</i>	23
2.1.2	<i>Caracterização botânica da linhaça</i>	24
2.1.3	<i>Componentes químicos da semente de linhaça</i>	25
2.1.4	<i>Produtos derivados da linhaça</i>	28
2.2	ÁCIDOS GRAXOS	29
2.2.1	<i>Estrutura química dos ácidos graxos</i>	29
2.2.2	<i>Classificação dos ácidos graxos</i>	30
2.2.3	<i>Funções fisiológicas dos ácidos graxos essenciais</i>	32
2.3	ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO E MÉDIO	34
2.4	MÉTODOS DE ANÁLISE MULTIVARIADA	39
2.4.1	<i>Quimiometria</i>	39
2.4.2	<i>Análise Multivariada</i>	39
2.4.3	<i>Construção do modelo de regressão</i>	42
2.4.4	<i>Análise de Componentes Principais (PCA)</i>	44
2.4.5	<i>Regressão por Mínimos Quadrados Parciais (PLSR)</i>	46
2.4.6	<i>Validação</i>	47
2.4.7	<i>Amostras anômalas (Outliers)</i>	49
3	OBJETIVOS	51
3.1	OBJETIVO GERAL	51
3.2	OBJETIVOS ESPECÍFICOS	51
4	MATERIAL E MÉTODOS	53
4.1	MATERIAL	53
4.2	MÉTODOS	53
4.2.1	<i>Extração dos lipídeos totais</i>	53
4.2.2	<i>Preparação de ésteres metílicos</i>	54
4.2.3	<i>Análise através de Cromatografia Gasosa associada à Espectrometria de Massas</i>	54
4.2.4	<i>Análise em Cromatografia Gasosa</i>	55
4.2.5	<i>Análise de espectroscopia no infravermelho próximo e médio</i>	56
4.2.6	<i>Desenvolvimento dos modelos de regressão</i>	57

5	RESULTADOS E DISCUSSÃO	58
5.1	PERFIL LIPÍDICO DAS FARINHAS E SEMENTES DE LINHAÇA	58
5.1.1	<i>Identificação dos Ésteres Metílicos de Ácidos Graxos</i>	59
5.2	ANÁLISE DAS COMPONENTES PRINCIPAIS (PCA)	64
5.2.1	<i>Análise de Componentes Principais de Espectros no Infravermelho Próximo</i>	64
5.2.2	<i>Análise de Componentes Principais no Infravermelho Médio</i>	69
5.2.3	<i>Comparação entre Análise de Componentes Principais (PCA) no NIR e MID</i>	74
5.3	DESENVOLVIMENTO DE MODELOS DE CALIBRAÇÃO MULTIVARIADA PARA PREVISÃO DO TEOR DE ÁCIDOS GRAXOS EM FARINHAS DE LINHAÇA	75
5.3.1	<i>Modelos de calibração multivariada para previsão do teor de ácido oleico para farinha de linhaça</i>	77
5.3.2	<i>Modelos de calibração multivariada para previsão do teor de ácido linoleico para farinha de linhaça</i>	82
5.3.3	<i>Modelos de calibração multivariada para previsão do teor de ácido linolênico para farinha de linhaça</i>	86
5.4	MODELOS DE CALIBRAÇÃO MULTIVARIADA PARA SEMENTES DE LINHAÇA	91
5.4.1.1	<i>Modelos de calibração multivariada para previsão do teor de ácido oleico para semente de linhaça</i>	92
5.4.1.2	<i>Modelos de calibração multivariada para previsão do teor de ácido linoleico para semente de linhaça</i>	97
5.4.1.3	<i>Modelos de calibração multivariada para previsão do teor de ácido linolênico para semente de linhaça</i>	101
5.5	COMPARAÇÃO DOS MODELOS DE CALIBRAÇÃO MULTIVARIADA PARA NIR E MID	105
6	CONCLUSÕES	107
7	SUGESTÕES PARA TRABALHOS FUTUROS	108
	REFERÊNCIAS	109

1 INTRODUÇÃO

Desde o ano 5000 a.C., o linho (*Linum usitatissimum* Linaceae) já era utilizado na produção de roupas e o óleo obtido da semente era empregado para embalsamar múmias. Posteriormente, o cultivo do linho se difundiu por toda a Europa, principalmente para o consumo de suas sementes e derivados (VROMANS, 2006).

Atualmente, duas variedades de linhaça, a semente do linho, são amplamente conhecidas e consumidas: a linhaça dourada e a linhaça marrom. A linhaça dourada é cultivada em regiões de clima frio, enquanto que a linhaça marrom é cultivada em regiões de clima quente e úmido (MARQUES, 2008).

Diversos estudos têm mostrado que o consumo regular de linhaça pode trazer benefícios à saúde humana por prevenir obstruções arteriais, causadoras de doenças cardiovasculares, por reduzir a pressão sanguínea em hipertensos e ainda diminuir o risco de enfermidades como o câncer e arteriosclerose devido à presença de ácidos graxos essenciais como o ácido linolênico e linoleico (PSOTA; GEBAUER; KRIS-ETHERTON, 2006). O consumo de linhaça pode também auxiliar no funcionamento intestinal graças ao elevado teor de fibra dietética e de compostos bioativos como lignanas e compostos fenólicos, que fazem da linhaça um alimento funcional.

A linhaça se diferencia das demais oleaginosas por apresentar elevado teor de ácidos graxos poliinsaturados (73% do total de ácidos graxos), moderado teor de monoinsaturados (13%) e baixo teor de saturados (9%). A linhaça contém entre 30 e 40% de lipídeos totais, dos quais entre 45 e 60% são de ácido linolênico, 15 e 18% de ácido linoleico e, 18 e 20% de ácido oleico (MORRIS, 2001; WHITE, 2007). A semente de linhaça contém ainda de 20 a 25% de proteínas, 20 a 28% de fibra dietética total, 4 a 8% de umidade e 3 a 4% de conteúdo mineral. Todavia, a composição química da semente de linhaça varia de acordo com as características genéticas da planta, condições ambientais de cultivo e tipo de processamento empregado (DAUN *et al.*, 2003).

Para identificar e quantificar os ácidos graxos presentes tanto na semente quanto na farinha de linhaça usam-se as técnicas de cromatografia gasosa acoplada à espectrometria de massas (CG/EM) e, mais comumente a cromatografia gasosa

com detector de ionização de chama (CG/DIC). Estas técnicas permitem determinar compostos em quantidades muito pequenas e, por isso, são consideradas análises sensíveis (HIGSON, 2009). No entanto, estas técnicas requerem operações de pré-tratamento da amostra como extração prévia dos analitos de interesse, utilização de vários reagentes, vidrarias e equipamentos, o que as tornam demoradas e geradoras de resíduos (GELADI, 2003; SINELLI *et al.*, 2010). Neste contexto, a espectroscopia no infravermelho apresenta-se como uma técnica analítica rápida, seletiva, sensível, de simples operação e baixo custo, além de apresentar elevada velocidade analítica, possibilidades de uso em sistemas de controle *on-line*, e de constituir um método não destrutivo e não poluente. Em virtude destas características, esta técnica pode ser utilizada em rotinas de análises químicas (CHIANG, COLEGROVE, 2007; SHIROMA; RODRIGUEZ-SAONA, 2009).

Tanto a espectroscopia no infravermelho próximo quanto no infravermelho médio são técnicas que permitem a determinação de propriedades físicas e químicas de alimentos através de seus respectivos espectros de absorção na região do infravermelho que se estende de 12.800 a 200 cm^{-1} . Estes espectros apresentam informações que têm sido cada vez mais empregadas no controle de qualidade e na caracterização qualitativa e quantitativa de produtos alimentícios (BLANCO *et al.*, 1999). Contudo, a análise dos dados gerados a partir dos espectros de infravermelho requer o uso de métodos que auxiliam na identificação e quantificação de componentes de interesse com base em procedimentos matemáticos e estatísticos. Dentre os métodos podem ser citadas a Análise de Componentes Principais (PCA), técnica quimiométrica que permite extrair, de um determinado conjunto de dados, informações relevantes para o seu entendimento, e a Regressão por Mínimos Quadrados Parciais (PLSR), utilizada na determinação de várias espécies químicas.

Neste trabalho, utilizou-se a metodologia de Análise de Componentes Principais que permitiu a discriminação entre as variedades de linhaça, bem como o desenvolvimento de modelos de calibração multivariada, por meio de PLSR, que permitiram a quantificação dos ácidos graxos oleico, linoleico e linolênico em farinhas e sementes de linhaça, utilizando espectroscopia no infravermelho médio e próximo.

2 REVISÃO DE LITERATURA

2.1 LINHAÇA

2.1.1 Histórico do cultivo

Os indícios mais antigos para o cultivo do linho (*Linum usitatissimum* L.) foram encontrados na região do Egito, datando entre os anos de 6000 a 8000 a.C. As propriedades medicinais da linhaça já eram estudadas por Hipócrates em 500 a.C (COSKUNER; KARABABA, 2007). Naquele mesmo período, o cultivo do linho se difundiu pelo oeste da Europa, por meio de nômades, e posteriormente pelo norte da Europa, África e por todo o mundo (ZOHARY; HOPF, 1993).

A cultura do linho foi trazida ao Brasil em 1550, primeiramente para Santa Catarina, em seguida para São Paulo, Paraná e Rio Grande do Sul. A partir de 1995, estudos a respeito do cultivo da linhaça se intensificaram, com base em suas propriedades nutracêuticas com o intuito de aumentar a produção desta oleaginosa. Até 2005, havia somente o plantio da variedade de linhaça marrom no Brasil, entretanto, a partir de 2006, a variedade dourada, que até então era importada do Canadá, passou a ser produzida também, mas em pequena escala (GALVÃO, 2009).

Segundo a Organização das Nações Unidas para Agricultura e Alimentação – FAO (2010) o maior produtor e exportador de linhaça é o Canadá (423 mil toneladas/ano), seguido por China (350 mil toneladas/ano), Estados Unidos da América (230 mil toneladas/ano), Etiópia (150 mil toneladas/ano) e Índia (146 mil toneladas/ano). Na América do Sul, os principais produtores são Argentina (52 mil toneladas/ano) e Brasil (12 mil toneladas/ano).

2.1.2 Caracterização botânica da linhaça

O linho (*Linum usitatissimum* L.) é uma planta herbácea, pertencente à família Linaceae, com mais de 200 espécies reconhecidas (CUI, 1998). A planta pode ter entre 30 e 130 cm de altura, caule ereto e fibroso com vários ramos, de onde nascem as folhas alongadas e estreitas, e flores de cor azul-clara com cinco pétalas dispostas em panículas terminais eretas (FIGURA 1) (EL-NAGDY *et al.*, 2010).

Os frutos apresentam-se em forma de cápsulas esféricas, também conhecidas como cachopas e, estas podem conter de uma a dez sementes. A semente de linhaça se caracteriza por ser achatada e oval com borda pontiaguda e com dimensões variando de 3,0 a 6,4 mm de comprimento, 1,8 a 3,4 mm de largura e 0,5 a 1,6 mm de espessura. A coloração das sementes pode variar de marrom-avermelhado ao dourado, conforme a quantidade de pigmentos da camada externa (COSKUNER; KARABABA, 2007).



FIGURA 1 – LINHO SECCIONADO: FLORES, CAULE, CACHOPAS E SEMENTES

FONTE: Adaptado de EPAMINONDAS (2009)

NOTA: 1 – corola; 2 – sépala; 3 – pétala; 4 – fruto (cachopa); 5 – seção transversal do fruto; 6 – semente marrom e dourada

As cultivares de linho destinadas à produção de fibra têxtil, se caracterizam por serem plantas mais altas e menos ramificadas, cultivadas especialmente no leste da Europa. No entanto, as cultivares oleaginosas, destinadas a alimentação, possuem muitas ramificações, são de menor porte e tipicamente desenvolvem-se em regiões frias e temperadas do hemisfério norte, como no Canadá e nos Estados Unidos da América (VAISEY-GENSER; MORRIS, 2003). As cultivares oleaginosas mais comumente cultivadas são a linhaça dourada, cultivada em regiões de clima frio, e a linhaça marrom, cultivada em regiões de clima quente e úmido, como no Brasil (MARQUES, 2008).

2.1.3 Componentes químicos da semente de linhaça

A composição química da semente de linhaça varia de acordo com as características genéticas da planta, condições ambientais de cultivo e tipo de processamento empregado (DAUN *et al.*, 2003). Os principais constituintes da semente são lipídeos, principalmente os ácidos graxos linolênico, oleico e linoleico, proteínas, fibra alimentar (solúvel e insolúvel), umidade e cinzas (Tabela 1).

TABELA 1 – COMPOSIÇÃO QUÍMICA DA SEMENTE DE LINHAÇA

Lipídeos (%)	Proteínas (%)	Fibras (%)	Umidade (%)	Cinzas (%)	Referência
30 – 40	20 – 25	20 – 25	4 – 8	3 – 4	Coskuner e Karababa (2007)
45,2	23,4	–	7,4	3,5	Muller <i>et al.</i> (2010)
39	22	–	–	–	Rubilar <i>et al.</i> (2010)

Classificada como uma oleaginosa, a semente de linhaça apresenta em média 30 a 40% de lipídeos. Os cotilédones armazenam a maior parte do óleo presente na semente, sendo encontrado em torno de 98% na forma de triglicerídeos, 0,9% de fosfolipídeos e 0,1% de ácidos graxos livres. As variedades de linhaça são ricas em ácidos graxos poliinsaturados (73%), dentre os quais o ácido linolênico representa cerca de 45 a 60%, o ácido linoleico 15 a 18% e em torno de 20% de

ácido oleico. Em menores proporções há também os ácidos graxos saturados (9%) e monoinsaturados (18%) (WHITE, 2007; MULLER *et al.*, 2010).

As principais proteínas encontradas na semente de linhaça são as globulinas, correspondendo a 18,6%, e as albuminas, com cerca de 17,7% do total (CHUNG; LEI; LI-CHAN, 2005). Em relação aos aminoácidos, tem-se em maior proporção o ácido glutâmico (aproximadamente 19,7 g/100g de proteína), o ácido aspártico (aproximadamente 9,7 g/100g de proteína), a arginina (aproximadamente 9,4 g/100g de proteína), e em menor proporção os aminoácidos essenciais como fenilalanina, isoleucina, leucina, lisina, metionina, treonina, triptofano e valina (OOMAH; MAZZA, 1993).

A semente de linhaça contém entre 35 e 45% de fibras, das quais dois terços constituem as fibras insolúveis, constituídas por celulose, lignina e hemicelulose (TARPILA, WENBERG, TARPILA, 2005). Tais fibras são pouco fermentáveis e agem regulando os movimentos intestinais, por aumentarem o volume fecal (SPILLER, 1994). As fibras solúveis são compostas por pectinas, gomas e mucilagens e, quando na presença de água, formam géis (RUBILAR *et al.*, 2010). As gomas e a mucilagem são obtidas durante a extração da torta da linhaça e são constituídas por xilose, arabinose, glicose, galactose, ácido galacturônico, ramnose e fucose, que correspondem cerca de 8% do peso da semente (CHEN; XU; WANG, 2006). As fibras contribuem para o atraso do esvaziamento gástrico, aumentam a saciedade e agem como absorvedoras de glicose e colesterol, reduzindo seus níveis na corrente sanguínea (SALMERÓN *et al.*, 1997).

Os fitoestrógenos são compostos encontrados nas paredes celulares dos vegetais. Estes compostos apresentam propriedades estrogênicas, agindo na regulação hormonal, incluem-se essencialmente em dois grupos, as isoflavonas e as lignanas (LAMPE, 2003). A semente de linhaça é a maior fonte vegetal conhecida de lignanas, contendo aproximadamente 600–700 vezes mais secoisolariciresinol diglicosídeo (SDG) do que outras fontes vegetais (RAJESHA *et al.*, 2006). Este fitoestrógeno é convertido primeiramente a enterodiol e, posteriormente à enterolactona por ação bacteriana no trato gastrointestinal (FIGURA 2) (OOMAH; MAZZA, 2000).

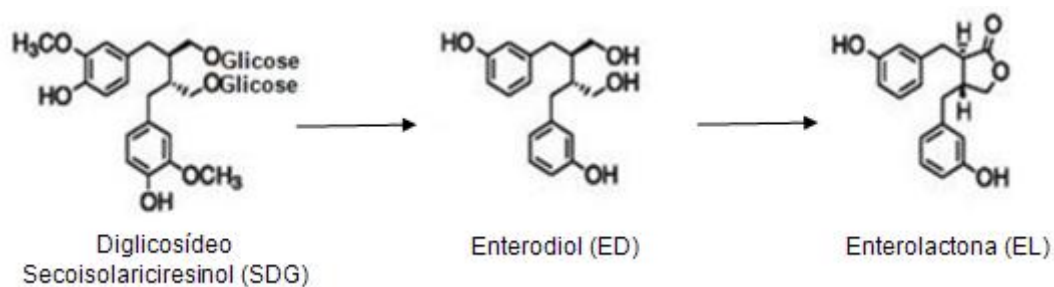


FIGURA 2 – CONVERSÃO DO SDG EM ENTERODIOL E ENTEROLACTONA POR AÇÃO BACTERIANA

FONTE: Adaptado de EPAMINONDAS (2009)

As lignanas demonstram semelhanças estruturais tanto com os estrogênios sintéticos quanto com os naturais e, podem ter propriedade estrogênica e antiestrogênica, já que se ligam aos receptores de estrogênio e, (TARPILA; WENBERG; TARPILA, 2005) desta forma auxiliam na prevenção de cânceres dependentes de estrogênio, inibindo o crescimento de células cancerosas. Os benefícios oriundos do consumo de lignanas residem não só nas propriedades anticarcinogênicas, mas também na capacidade antioxidante destes compostos (RUBILAR *et al.*, 2010), o que reduz o risco de problemas cardiovasculares e o desenvolvimento de diabetes (MUELLER *et al.*, 2010).

Além do SDG, outras lignanas como o matairesinol, isolariciresinol e pinioresinol (STRUIJS *et al.*, 2007) podem ser encontradas na linhaça, embora suas concentrações variem de acordo com a variedade, safra, origem e métodos de processamento da semente (LAMPE, 2003).

Com relação aos minerais presentes na semente de linhaça, que constituem entre 3 e 4% (COSKUNER; KARABABA, 2007), os encontrados em maior quantidade são o potássio, o fósforo, o magnésio e o cálcio. Dentre as principais vitaminas hidrossolúveis estão a niacina (B₃), piroxidina (B₆), ácido pantotênico (B₅), tiamina (B₁), ácido ascórbico (C) e riboflavina (B₂) e, dentre as vitaminas lipossolúveis estão a filoquinona (K1) e o tocoferol (E), sendo que o gama-tocoferol é a forma mais abundante da vitamina E presente na linhaça (MORRIS, 2007).

Dentre os demais constituintes da semente de linhaça tem-se a umidade com cerca de 4 a 8% (COSKUNER; KARABABA, 2007).

2.1.4 Produtos derivados da linhaça

As sementes podem ser utilizadas inteiras ou moídas, como farinhas, sendo incorporadas em produtos de panificação como pães, bolos e biscoitos. Estas sementes ainda podem ser adicionadas em barras de cereais e usadas como substitutos de gordura em produtos cárneos. Outros produtos comerciais oriundos da linhaça são óleo, farelo, torta e goma (DAVIS, 2004).

O óleo da semente de linhaça pode ser obtido por meio da prensagem a frio, sendo destinado ao consumo humano, comercializado na forma encapsulada ou em garrafas de vidro âmbar. Ainda que a prensagem a frio seja muito empregada, esta técnica recupera somente parcialmente o óleo contido nas sementes (BOZAN; TEMELLI, 2002). A extração do óleo pode ocorrer também por prensagem mecânica seguida da adição de solventes orgânicos. Neste caso, a produção é direcionada a indústria de tintas e vernizes (JHALA; HALL, 2010), principalmente devido à capacidade destes óleos formarem filmes com boas propriedades mecânicas, em consequência do elevado teor de ácido linolênico (C18:3), além de apresentarem secagem rápida (LAZZARI; CHIANTORE, 1999).

A extração por prensagem a frio apresenta menor rendimento e maior consumo de energia, enquanto que a extração por solvente envolve maior tempo de extração (ZHANG *et al.*, 2008). Uma alternativa para obtenção do óleo de linhaça, sem perda de qualidade, e com economia de tempo, é o uso de extração supercrítica com CO₂, que ainda apresenta as vantagens de proporcionar uma extração de compostos termolábeis sem degradação e sem resíduos de solventes (PRADHAN *et al.*, 2010).

A partir da extração do óleo obtém-se a torta e o farelo da linhaça. A torta é um subproduto rico em proteína largamente empregado como ração animal ou fertilizante (MULLER *et al.*, 2010), enquanto que o farelo, que apresenta elevado teor de fibras e conteúdo proteico concentrado, pode ser utilizado no consumo humano ou na produção de ração animal, com o benefício de auxiliar os movimentos do trato intestinal (PITA *et al.*, 2006).

A goma da linhaça, que representa cerca de 8% do peso da semente, é facilmente extraída. Esta goma é uma substância viscosa, com alta capacidade de retenção hídrica, semelhante à goma arábica do ponto de vista de aplicações

tecnológicas. A goma da linhaça é capaz de formar géis termo-lábeis e, desta forma, ser empregada na substituição de gomas de elevado custo financeiro (CHEN; XU; WANG, 2006).

2.2 ÁCIDOS GRAXOS

2.2.1 Estrutura química dos ácidos graxos

Os ácidos graxos são constituídos de ácidos carboxílicos com cadeias, geralmente lineares, de hidrocarbonetos alifáticos de comprimento variando entre 4 e 36 carbonos (NELSON; COX, 2002). A maioria dos ácidos graxos naturais apresentam comprimentos de cadeia entre 12 e 24 átomos de carbonos, sendo os mais comuns com 18 átomos de carbono. Há cerca de 1.000 ácidos graxos conhecidos, que diferem entre si de acordo com o número de átomos de carbono da cadeia, posição, configuração e tipo de insaturação. No entanto, apenas 20 ácidos graxos são encontrados na natureza, sendo que destes, os ácidos palmítico, oleico, linoleico e linolênico compõem aproximadamente 80% dos óleos e gorduras (GUNSTONE; HARWOOD; DIJKSTRA, 2007).

A nomenclatura dos ácidos graxos é padronizada pela União Internacional de Química Pura e Aplicada (*International Union of Pure and Applied Chemistry – IUPAC*) que considera para a nomenclatura a estrutura química, como o número de átomos de carbono, presença ou não de insaturação e o grupo funcional. No caso de haver insaturação, esta pode ter configuração *cis* ou *trans* (GRAZIOLA; SOLIS; CURI, 2002; VISENTAINER; FRANCO, 2006).

2.2.2 Classificação dos ácidos graxos

2.2.2.1 Ácidos graxos saturados

Os ácidos graxos saturados formam uma série homóloga de ácidos monocarboxílicos. Estes ácidos são caracterizados por terem um número par de átomos de carbono e serem encontrados tanto em óleos vegetais como em gorduras animais.

Entre os ácidos graxos presentes em óleos vegetais, encontram-se aqueles que têm cadeia curta como o ácido butanoico (ácido butírico – C4:0), os de cadeia média, como o ácido dodecanoico (ácido láurico - C12:0) e o ácido tetradecanoico (ácido mirístico - C14:0) e, os ácidos graxos de cadeia longa como o ácido hexadecanoico (ácido palmítico – C16:0) e ácido octodecanoico (ácido esteárico – C18:0) (GUNSTONE; HARWOOD; DIJKSTRA, 2007). Raramente há ácidos graxos saturados com mais de 24 carbonos nos óleos vegetais comestíveis, mas estes podem ser encontrados nas ceras (OETTERER; REGITANO-D'ARCE; SPOTO, 2006).

2.2.2.2 Ácidos graxos insaturados

Em produtos de origem vegetal, o mais comum é o ácido oleico (ácido *cis*-9-octadecenoico). Embora seu isômero, o ácido elaídico (*trans*-9-octadecenoico) não seja de ocorrência natural, ele é formado no processo de hidrogenação catalítica dos óleos vegetais (KEEFE, 2008). A configuração natural das ligações duplas em ácidos graxos insaturados é a configuração *cis*, porém a configuração *trans* apresenta a cadeia mais linear em relação à configuração *cis*, tornando a molécula mais rígida e termodinamicamente mais estável (FIGURA 3) (DAMODARAN; PARKIN; FENNEMA, 2010).

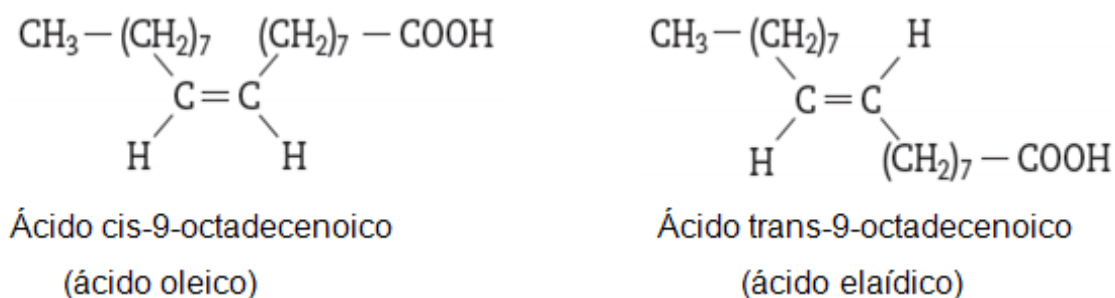


FIGURA 3 – ESTRUTURA QUÍMICA DOS ÁCIDOS OLEICO E ELAÍDICO
 FONTE: SOLOMONS; FRYHLE (2006)

Os ácidos monoinsaturados são mais suscetíveis a oxidação, devido à presença da ligação dupla na cadeia alifática, do que os compostos correspondentes saturados. No entanto, ácidos monoinsaturados são mais resistentes a oxidação que os poliinsaturados (OETTERER; REGITANO-D'ARCE; SPOTO, 2006).

Os ácidos graxos poliinsaturados possuem duas ou mais ligações duplas em sua cadeia carbônica. No entanto, o elevado grau de insaturação causa instabilidade oxidativa e alta reatividade (OETTERER; REGITANO-D'ARCE; SPOTO, 2006). As ligações duplas presentes na cadeia dos ácidos graxos poliinsaturados são separadas por um grupo metileno (-CH=CH-CH₂-CH=CH-) e em alguns casos há ligações duplas conjugadas, ou seja, sem a presença do grupo metileno (SPITELLER, 2006).

Além da classificação oficial da IUPAC, considera-se usualmente a nomenclatura ômega (ω), onde os ácidos graxos insaturados podem ser classificados de acordo com a posição da primeira ligação dupla, a partir do grupo metila, e divididos em famílias, conforme exemplo na FIGURA 4. Os mais conhecidos são os ácidos graxos das famílias ômega 9 (ω -9), ômega 6 (ω -6) e ômega 3 (ω -3) (GUNSTONE, 2004), sendo seus principais representantes o ácido oleico, o ácido linoleico e o ácido linolênico, respectivamente (BENATTI *et al.*, 2004).

O ácido linoleico (*cis* 9,12-octadecadienoico) e o ácido linolênico (*cis* 9,12,15-octadecatrienoico) são considerados ácidos graxos essenciais por não serem sintetizados pelo organismo humano estando presentes em óleos vegetais como o óleo de soja, canola e linhaça, além de outros produtos derivados destas fontes vegetais (OETTERER; REGITANO-D'ARCE; SPOTO, 2006).

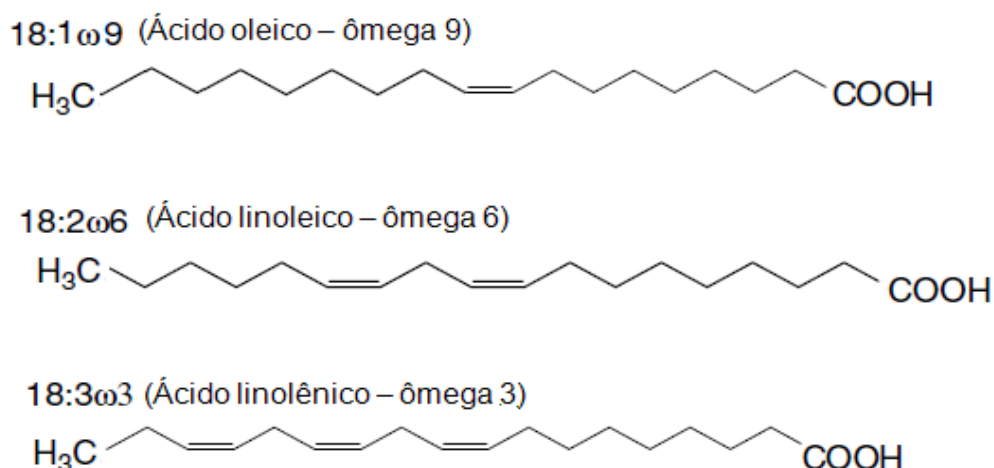


FIGURA 4 – ESTRUTURA QUÍMICA DA FAMÍLIA ÔMEGA REPRESENTADA PELO ÁCIDO OLEICO, LINOLEICO E ÁCIDO LINOLENICO
 FONTE: KEEFE (2008)

2.2.3 Funções fisiológicas dos ácidos graxos essenciais

Os ácidos graxos essenciais têm importante função nutricional e podem ser encontrados em diversas formas, tanto de origem vegetal quanto animal. Estes ácidos exercem funções estruturais das membranas celulares e em diversos processos metabólicos (YOU DIM; MARTIN; JOSEPH, 2000), além de contribuírem para a regulação hormonal. Os ácidos graxos poliinsaturados como o ácido linoleico e linolênico são necessários para a saúde humana por contribuírem para a manutenção da fluidez, da flexibilidade e da permeabilidade seletiva das membranas, além das funções cerebrais e transmissão de impulsos nervosos (WARD; SINGH, 2005). De acordo com Morris *et al.* (2003), uma dieta contendo ácidos graxos poliinsaturados, como ácido linolênico, pode auxiliar na redução do risco do mal de Alzheimer, bem como na prevenção de outras doenças degenerativas.

Os ácidos graxos essenciais também desempenham funções na síntese da hemoglobina e na divisão celular, além de participarem da produção de energia e da transferência do oxigênio atmosférico para o plasma sanguíneo (YEHUDA *et al.*, 2002). Tem-se ainda especulado que o consumo de ácidos graxos poliinsaturados pode contribuir significativamente no combate a doenças cardiovasculares (HU;

WILLETT, 2002). Além disso, os ácidos graxos auxiliam no transporte das vitaminas lipossolúveis como A, D, E e K (WARD; SINGH, 2005).

Os ácidos graxos essenciais são encontrados nas famílias ômega 3 e 6, sendo representados principalmente pelos ácido linolênico e linoleico, que mediante reações enzimáticas produzem ácidos graxos poliinsaturados de cadeia muito longa, com 20 ou mais átomos de carbono, que são extremamente importantes para saúde humana, como os ácidos eicosapentaenoico e docosahexaenoico (YEHUDA *et al.*, 2002). Estas reações ocorrem pela ação de enzimas alongase e dessaturase, por meio dos processos de alongação, onde são adicionados dois átomos de carbono, e/ou dessaturação da cadeia, na qual ocorre inserção de uma ligação dupla. A Figura 5 apresenta o metabolismo dos ácidos graxos linolênico e linoleico, que resultam na produção dos ácidos eicosapentaenoico (EPA C20:5 n-3) e docosahexaenoico (DHA C22:6 n-3), e araquidônico (C20:4 n-6), respectivamente (MORRIS, 2007). A principal função atribuída ao DHA (C22:6 n-3) é o auxílio nas funções cerebrais e da retina, sendo predominante na maioria das membranas celulares desses órgãos. Já o EPA forma a prostaglandina, substância que age como hormônio, mas são ácidos graxos produzidos por quase todas as células do corpo (MORRIS, 2007).

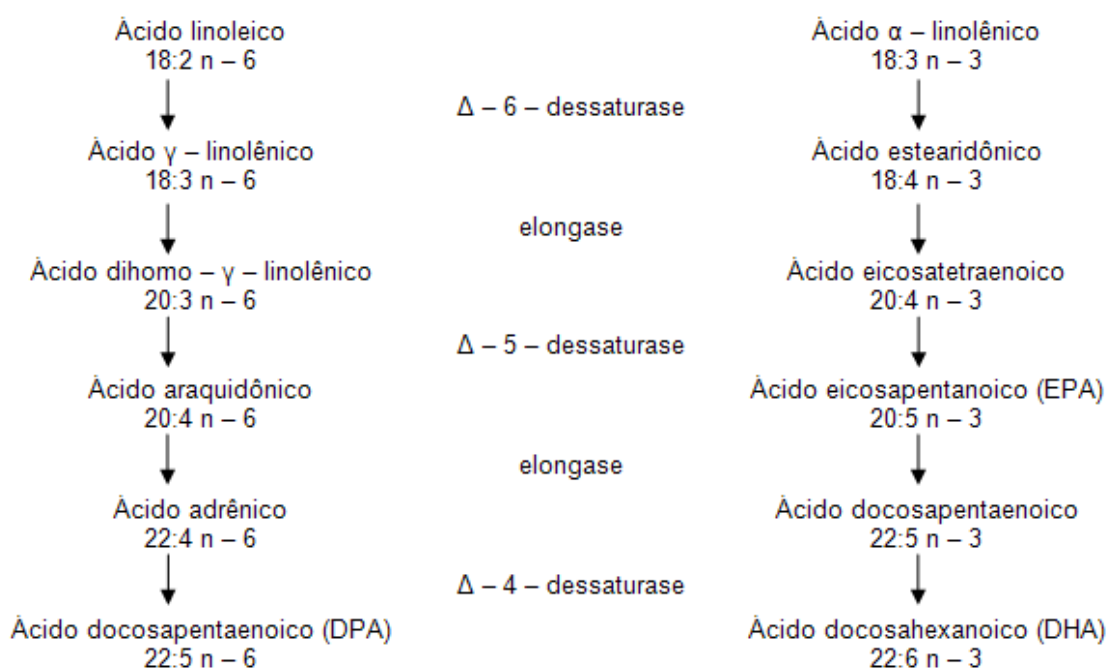


FIGURA 5 – METABOLISMO DOS ÁCIDOS GRAXOS LINOLEICO E LINOLÊNICO

FONTE: Adaptado de KEEFE (2008)

Os ácidos oleico, linoleico e linolênico podem ser encontrados em produtos de origem vegetal como óleos e sementes, e em alguns produtos de origem animal, como peixes. Nos produtos vegetais, o maior teor de ácido linolênico ocorre no óleo e na semente de linhaça (Tabela 2) (USDA - *United States Department of Agriculture*, 2012).

TABELA 2 – CONCENTRAÇÃO DOS ÁCIDOS OLEICO, LINOLEICO E LINOLÊNICO (g/100g) EM ALIMENTOS DE ORIGEM VEGETAL

Óleos	Ácido oleico (C18:1 n-9)	Ácido linoleico (C18:2 n-6)	Ácido linolênico (C18:3 n-3)
Canola	61,74	19,0	9,14
Linhaça	18,32	14,33	53,37
Milho	27,33	53,51	1,16
Oliva	71,27	9,77	0,76
Soja	22,55	50,95	6,79
Grãos	Ácido oleico (C18:1 n-9)	Ácido linoleico (C18:2 n-6)	Ácido linolênico (C18:3 n-3)
Linhaça	7,36	5,90	22,81
Milho	0,42	0,47	0,01
Soja	4,35	9,92	1,33

FONTE: Adaptado de USDA (2012)

2.3 ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO E MÉDIO

A espectroscopia é o estudo das interações da radiação eletromagnética com a matéria (gases, líquidos ou sólidos). As interações da radiação com a matéria podem levar a transições entre os níveis de energia dos átomos e moléculas e, o resultado destas interações depende diretamente da energia incidente, associada ao comprimento de onda da radiação (BRUICE, 2006; SKOOG *et al.*, 2005).

A região do infravermelho corresponde a região do espectro eletromagnético situada entre as regiões do visível e micro-ondas, engloba radiações com número de onda de 12.800 cm^{-1} a 10 cm^{-1} , ou com comprimentos de onda de 780 a $1,0 \times 10^6 \text{ nm}$. O espectro infravermelho pode ser dividido em infravermelho próximo (NIR, do inglês *Near Infrared*), médio (MID, do inglês, *Middle Infrared*) e distante (FAR, do inglês, *Far Infrared*) (Tabela 3) (HOLLER, 2009; SILVERSTEIN *et al.*, 2007).

TABELA 3 – REGIÕES ESPECTRAIS DO INFRAVERMELHO

Região Espectral	Número de onda (cm ⁻¹)	Comprimento de onda (nm)	Frequência (Hz)
Próximo (NIR)	12.800 a 4.000	780 a 2.500	3,8 x10 ¹⁴ a 1,2 x10 ¹⁴
Médio (MID)	4.000 a 200	2.500 a 50.000	1,2 x10 ¹⁴ a 6,0 x10 ¹²
Distante (FAR)	200 a 10	50.000 a 1.000.000	6,0 x10 ¹² a 3,0 x10 ¹¹

Fonte: SKOOG; HOLLER; NIEMAN (2002)

Para absorver a radiação infravermelha, uma molécula deve sofrer uma variação no momento dipolo durante seu movimento rotacional ou vibracional. Apenas sob estas circunstâncias o campo elétrico alternado da radiação pode interagir com a molécula e causar variações na amplitude do movimento vibracional ou rotacional. O momento de dipolo é determinado pela magnitude da diferença de carga e pela distância entre dois centros de carga. No entanto, não ocorre variação no momento de dipolo durante a vibração ou rotação de espécies homonucleares, tais como O₂, N₂ ou Cl₂, o que impede estas moléculas de absorver radiação no infravermelho (HOLLER, 2009).

As características de uma vibração atômica podem ser aproximadas por um modelo mecânico, o qual considera que a ligação entre duas massas (átomos) é feita por uma mola, em que a frequência de vibração da mola é descrita pela Lei de Hooke. Uma perturbação de uma dessas massas ao longo do eixo da mola resulta em uma vibração chamada movimento harmônico simples. As vibrações moleculares podem ser classificadas em dois tipos, vibração de deformação axial ou estiramento (*stretching*) e vibração de deformação angular (*bending*), podendo estas deformações ser simétricas ou assimétricas. As deformações de estiramento são oscilações radiais das distâncias entre os núcleos, e as deformações angulares envolvem mudanças dos ângulos entre o plano que contém a ligação e um plano de referência. As vibrações de deformação angular podem ser de quatro tipos: simétrica no plano (*scissoring*), assimétrica no plano (*rocking*), simétrica fora do plano (*wagging*) e assimétrica fora do plano (*twisting*) (FIGURA 6) (PAVIA *et al.*, 2010, OZAKI *et al.*, 2007).

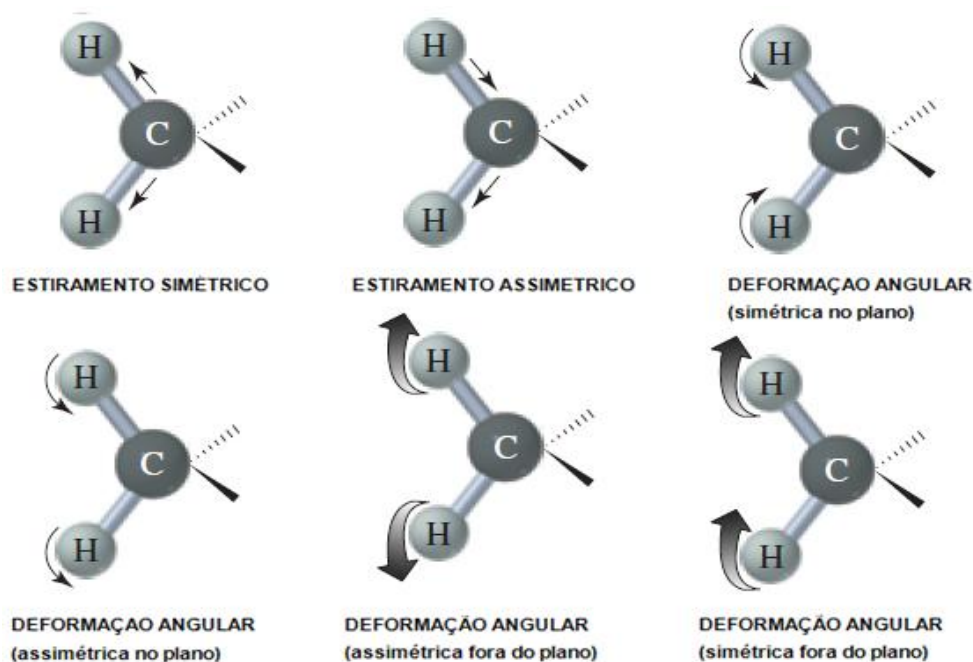


FIGURA 6 – DIFERENTES FORMAS DE VIBRAÇÕES MOLECULARES DE DEFORMAÇÃO
 FONTE: BRUCE (2006)

A espectroscopia de infravermelho próximo gera espectros que apresentam sobretons (*overtones*) de vibrações fundamentais e bandas de combinações de vibrações fundamentais. As bandas de combinação ocorrem quando dois modos vibracionais são excitados simultaneamente e as duas frequências vibracionais (ν_1 e ν_2) acoplam-se em uma molécula, originando uma nova frequência de vibração dentro da molécula (NAES, *et al.*, 2002; KELLNER *et al.*, 2004).

O espectro no infravermelho médio apresenta vibrações fundamentais, onde as bandas são intensas e bem definidas. O espectro infravermelho médio pode ser subdividido em duas regiões, sendo a região de 4.000 a 1.400 cm^{-1} , correspondente a bandas de absorção da maioria dos grupos funcionais. Já a região de 1.400 a 600 cm^{-1} é conhecida como região de impressão digital, onde cada substância apresenta um padrão específico de bandas de absorção (GIVENS *et al.*, 1997; HOLLER, 2009; PAVIA *et al.*, 2010).

Desta forma, as moléculas estão sujeitas ao desenvolvimento de vários efeitos inerentes as ondas eletromagnéticas. A interação da radiação eletromagnética com a matéria pode ocorrer, por exemplo, por meio de absorção, reflexão e espalhamento. Estas interações são características dos átomos presente nas moléculas, bem como da região do espectro envolvida, permitindo a identificação de compostos (Tabela 4) (SILVERSTEIN *et al.*, 2007).

TABELA 4 – REGIÕES ESPECTRAIS NO MID E NIR RELACIONADAS ÀS LIGAÇÕES QUÍMICAS DE DIVERSOS GRUPOS FUNCIONAIS

Número de onda (cm ⁻¹)	Características dos grupos funcionais	
Infravermelho próximo (NIR)		
≈ 8600 – 8150	C – H	2º sobretom – ácidos graxos
≈ 7400 – 7000	C – H	2º sobretom – ácidos graxos
≈ 5950 – 5600	C – H	1º sobretom – ácidos graxos
≈ 5500 – 5260	C = O e O – H	1º sobretom – ácidos graxos
≈ 4350 - 4150	C – H	Combinação – ácidos graxos
Infravermelho médio (MID)		
≈ 3300 – 2500	O – H	Ácido carboxílico
≈ 3040 – 2850	C – H	Estiramento; ácidos graxos
≈ 3005	C – H	Estiramento; ácidos graxos
≈ 2926	C – H	Lipídeos
≈ 2869	C – H	Lipídeos
≈ 1700 – 1765	C = O	Lipídeos
≈ 1780 – 1650	C = O	Lipídeos
≈ 1458 – 1375	C – H	Grupos alifáticos
≈ 1200 - 900	C – O e C – C	Região de impressão digital

FONTE: Adaptado de BRUICE (2006); FAGAN, O'DONNELL (2008); SHIROMA; RODRIGUEZ-SAONA (2009); XIABO *et al.* (2010)

Para a caracterização de compostos na região do infravermelho pode-se utilizar entre os mecanismos de medida, a absorvância, a transmitância e a reflectância, sendo que a técnica de reflectância tem sido utilizada com a finalidade de determinar quantitativamente proteínas, lipídeos e umidade, em produtos agrícolas como grãos e sementes oleaginosas, e em outras amostras sólidas (HOLLER, 2009).

A interação da radiação eletromagnética com a matéria pode sofrer diferentes tipos de reflexão, tais como: reflexão difusa, reflexão especular, reflexão interna e reflexão total atenuada. No caso da reflectância difusa, sabe-se que é um processo complexo, onde a radiação penetra na camada superficial das partículas e excita os modos vibracionais das moléculas do composto de interesse e a energia refletida emerge aleatoriamente e espalha-se em todas as direções (FIGURA 7). A energia refletida difusamente pela amostra carrega informação suficiente, para geração do espectro da amostra na região do infravermelho (MENDHAM, 2002). No caso da espectrometria no Infravermelho de Reflectância Difusa com Transformada de Fourier, obtêm-se diretamente espectros de amostras sólidas com um mínimo de preparo da amostra. Além de poupar tempo no preparo da amostra, esta técnica permite que dados espectrais de infravermelho sejam coletados com amostras que não foram alteradas em relação ao seu estado original (HOLLER, 2009).

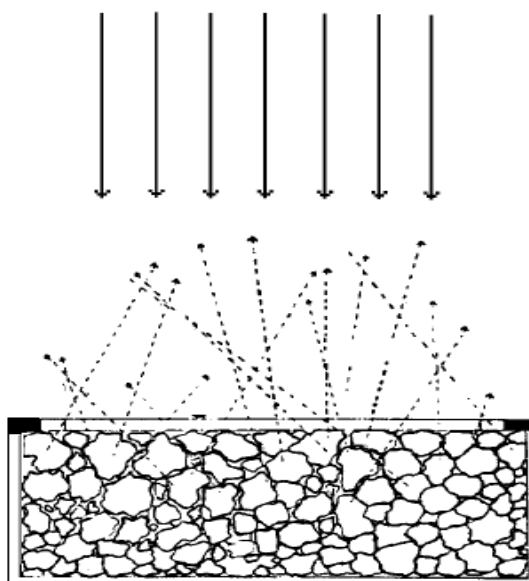


FIGURA 7 – REPRESENTAÇÃO ESQUEMÁTICA DE REFLECTÂNCIA DIFUSA NO INFRAVERMELHO EM AMOSTRA SÓLIDA
FONTE: NATALINO (2006)

As técnicas de espectroscopia no infravermelho próximo (NIR) e no infravermelho médio (MID) são cada vez mais utilizadas no controle de qualidade de processos e produtos industriais (BLANCO *et al.*, 1999), principalmente porque as medidas são rápidas, utilizam pequena quantidade de amostra e, envolvem preparação rápida e simples das amostras. Tanto técnicas de NIR como de MID permitem ainda a determinação simultânea de diferentes parâmetros (CAMPESTRINI, 2005; CEN; HE, 2007), além de serem métodos não destrutivos, limpos, sem uso de reagentes agressivos e nem geração de resíduos (HELFER *et al.*, 2006).

Com o intuito de prever a concentração de uma espécie química presente em uma amostra complexa e heterogênea, tanto o infravermelho próximo quanto o médio empregam o artifício de correlacionar resultados obtidos em laboratório, por meio de técnicas e métodos de referência, com os espectros obtidos nas regiões NIR e/ou MID, através da calibração multivariada.

2.4 MÉTODOS DE ANÁLISE MULTIVARIADA

2.4.1 Quimiometria

A Quimiometria é uma área da química que utiliza a aplicação de métodos estatísticos e matemáticos para análise de dados de origem química, com natureza multivariada, a fim de planejar e selecionar experimentos com o intuito de obter o máximo de informação química com a análise dos dados obtidos (OTTO, 2007). As diversas técnicas quimiométricas têm por finalidade desenvolver modelos empíricos ou semi-empíricos derivados dos dados, usados para estimar uma ou mais propriedades de um sistema, com base em medições de outras propriedades (WISE *et al.*, 2003).

Visto que muitas das análises químicas convencionais são caras e/ou demandam muito tempo de análise, é altamente desejável a utilização de métodos diretos para estimar rapidamente propriedades químicas de amostras. Dentre estes métodos podem ser citadas as análises realizadas na região do infravermelho próximo e médio, que necessitam da utilização de modelos de calibração multivariada para que se possam obter determinações quantitativas.

2.4.2 Análise Multivariada

Entre os métodos de análise multivariada tem-se a Análise de Componentes Principais (*Principal Component Analysis* – PCA) e a Regressão por Mínimos Quadrados Parciais (*Partial Least Squares* – PLSR). A PCA tem sido aplicada na área de alimentos em diferentes matrizes com a finalidade tanto de identificar diferentes variedades (KARÁSEK *et al.*, 2003) quanto de caracterizar oleaginosas (GARCÍA; TORRE; MARINA, 2000), utilizando apenas os dados espectrais obtidos nas regiões de infravermelho.

Enquanto que a PLSR tem sido aplicada para análises quantitativas, por exemplo, na determinação da composição química de sementes (FASSIO;

COZZOLINO, 2004) e na composição de ácidos graxos em óleo (SUNDARAM *et al.*, 2010), através da relação entre duas matrizes de dados químicos, quando há uma dependência entre as propriedades a serem analisadas (Nagata; Bueno; Peralta-Zamora, 2001).

2.4.2.1 Pré-processamento dos dados

Para melhorar a eficiência dos modelos multivariados, pode se aplicar algoritmos para minimizar fontes de variabilidade não informativas e condicionar o conjunto de dados a ser estudado, através de técnicas de pré-processamento, a exemplo do alisamento, centralização na média, cálculo de derivada, correção do espalhamento multiplicativo, variação de padrão normal, detrend e normalização.

O pré-processamento elimina variações devido à interação complexa entre os componentes da amostra, dispersão da luz em amostras sólidas ou líquidas turvas, baixa reprodutibilidade das medições, distorções espectrais, como deslocamento da linha de base e comprimentos de ondas, efeitos da não linearidade do detector, ruído no detector e variação no espectro devido à variações de temperatura no ambiente. O pré-processamento dos dados espectrais permite reduzir, eliminar ou padronizar o impacto destas fontes de variação nos espectros, sem alterar a informação contida neles (SIESLER *et al.*, 2002).

- Alisamento

O ruído aleatório que acompanha o sinal analítico (APARÍCIO, 2002), provém de variações imprevisíveis como vibrações, correntes de ar, movimento aleatório das moléculas e outros (O'HAVER, 2001). A aplicação da técnica de alisamento visa reduzir matematicamente o ruído instrumental, aproximadamente, pela raiz quadrada do número de pontos utilizados (BROAD *et al.*, 2001).

- Centralização na média

O cálculo da centralização na média previne que pontos mais distantes do centro dos dados tenham maior influência que os mais próximos, dando importância não a intensidade, mas sim a sua distância do valor médio. Desta forma, impede que a magnitude das variáveis afete os cálculos (FERREIRA *et al.*, 1999; THOMAS, 1994).

- Cálculo da derivada

O cálculo da primeira e segunda derivada, quando empregado a dados espectrais, gera mudanças na informação processada, podendo ter efeito positivo, ao evidenciar um sinal analítico de interesse, ou ter efeito negativo, ao acentuar ruídos instrumentais (SAUER, 2007). O cálculo da derivada permite corrigir deslocamentos da linha de base (CHAMINADE *et al.*, 1998) e melhorar a separação de sinais não totalmente sobrepostos.

- Correção do Espalhamento Multiplicativo

A correção do espalhamento multiplicativo ou correção do sinal multiplicativo (*Multiplicative Scatter Correction – MSC*) foi desenvolvida para corrigir irregularidades nos espectros de infravermelho de refletância difusa, relacionada ao espalhamento de luz provocado por irregularidades em amostras sólidas, a fim de linearizar os dados espectrais e minimizar fontes de variabilidade não informativas (HARRIS, 2008).

- Variação de Padrão Normal

A variação de padrão normal (*Standard Normal Variate – SNV*) é utilizada para minimizar interferências causadas pelo tamanho de partículas e diferenças de densidade das amostras (NAES *et al.*, 2002). Este pré-processamento tem características similares às descritas pelo MSC.

- Detrend

Detrend é aplicado no espectro para remover efeitos de deslocamento de linha de base. O método consiste em modelar a linha de base em função do comprimento de onda com um polinômio de segundo grau ou grau superior e sua função é então subtraída de cada espectro independentemente (LUYPAERT *et al.*, 2002).

- Normalização

A normalização é realizada a fim de regularizar a importância relativa de cada variável, assim todas as variáveis se tornam equivalentes em termos de magnitude. A escala do espectro pode ser normalizada dividindo-se cada intensidade espectral pela raiz quadrada da soma dos quadrados de todas as intensidades do espectro correspondente, de acordo com a equação 1 (FERRARINI, 2004).

$$X_n = \frac{X_{ij}}{\left(\sqrt{\sum_{j=i}^p X_{ij}^2} \right)} \quad (1)$$

2.4.3 Construção do modelo de regressão

Para construção do modelo de regressão, inicialmente obtém-se os espectros de um conjunto de amostras e a análise dessas amostras é realizada por meio de um método de referência a fim de obter os valores de referência correspondentes dos parâmetros de interesse. Logo, o desenvolvimento do modelo envolve a determinação das relações entre os espectros NIR e/ou MID e, os valores de referência dos parâmetros analisados, conforme diagrama na FIGURA 8.

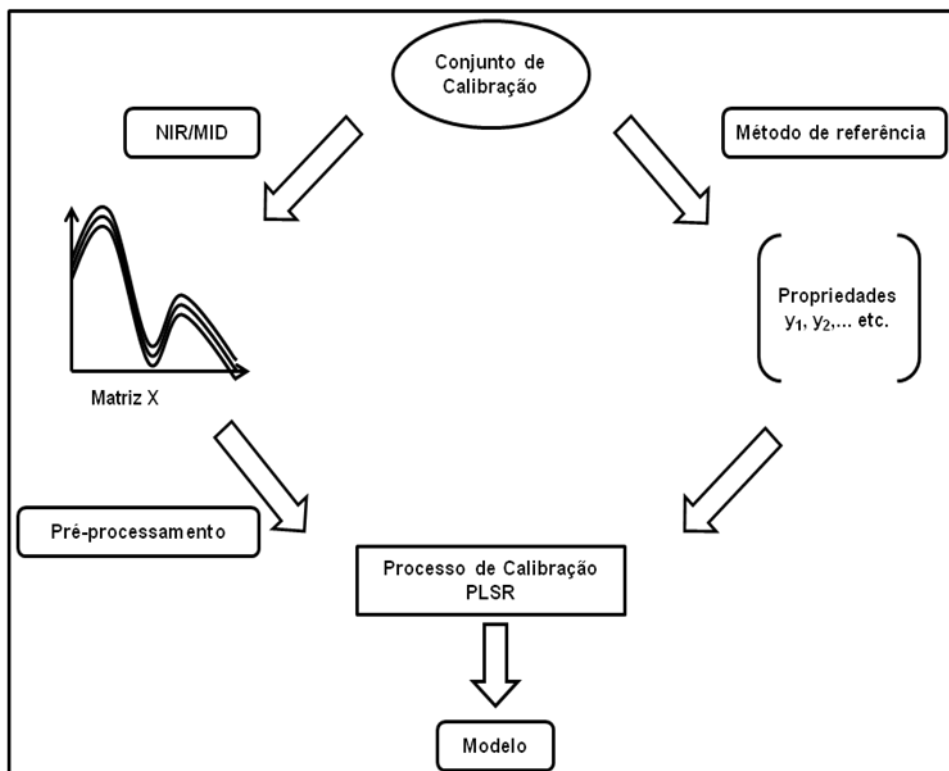


FIGURA 8 – DIAGRAMA DE CONSTRUÇÃO DO MODELO DE CALIBRAÇÃO
 FONTE: SOTELO (2006)

O modelo de calibração é uma relação entre X (espectros NIR ou MID) e Y (valores de referência). O conjunto de dados de Y representa a informação que se espera que o modelo represente e seja capaz de estimar quando amostras desconhecidas forem analisadas (SOTELO, 2006).

No processo de calibração, para um conjunto de amostras com composição conhecida são obtidos “ n ” espectros em “ p ” valores de comprimento de onda, formando uma matriz X_{cal} (X – matriz de calibração). Além disso, uma matriz Y_{cal} (Y – matriz de calibração), com os valores de concentrações pode ser formada, contendo “ n ” linhas correspondendo às diferentes amostras, e “ q ” colunas, indicando o número de diferentes espécies presentes nas amostras (FERREIRA *et al.*, 1999).

Os dados para a calibração multivariada podem ser organizados conforme mostrado na FIGURA 9, onde os valores de absorvância, transmitância ou reflectância dos espectros, a cada comprimento de onda, correspondem às variáveis independentes, e as concentrações das espécies de interesse nas amostras, as variáveis dependentes (TOZETTO, 2005).

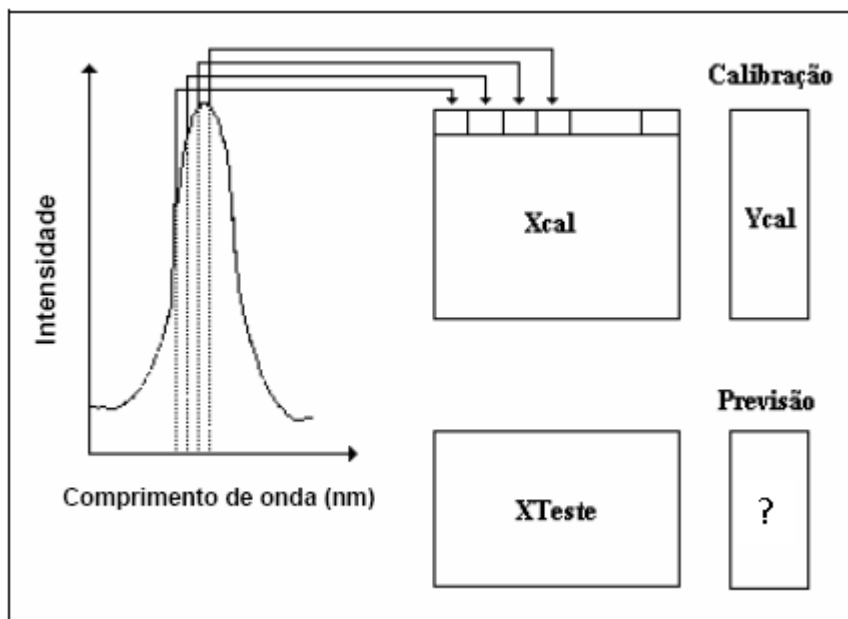


FIGURA 9 – ORGANIZAÇÃO DOS DADOS PARA CALIBRAÇÃO MULTIVARIADA
 FONTE: NAGATA (2001)

Resumidamente, o processo de calibração consiste de duas etapas, a modelagem, que estabelece uma relação matemática entre X e Y no conjunto de calibração, e a validação, que otimiza a relação no sentido de uma melhor descrição do analito de interesse.

2.4.4 Análise de Componentes Principais (PCA)

A análise de componentes principais (PCA) é um método multivariado aplicado na discriminação de dados com o objetivo de reduzir sua dimensionalidade minimizando a covariância entre as variáveis, sem perda de informações, permitindo a observação das semelhanças e diferenças entre as amostras, para fazer uma verificação inicial de sua homogeneidade e definir um número de componentes principais (SHIN *et al.*, 2010).

A transformação é feita de modo que o conjunto de dados possa ser representado por um número reduzido de novas variáveis chamadas componentes principais, que são combinações lineares das variáveis originais. A PCA decompõe a matriz de dados originais em uma soma de matrizes, em que a primeira componente principal (PC1) é definida na direção de máxima variância do conjunto de dados,

enquanto que a segunda componente principal (PC2) é definida na direção que descreve a máxima variância no espaço da PC1, de forma que cada componente principal é responsável pela fração sucessiva de variância de dados, consistindo em um sistema de coordenadas ortogonais entre si, conforme pode ser observado na FIGURA 10 (FERREIRA *et al.*, 1999).

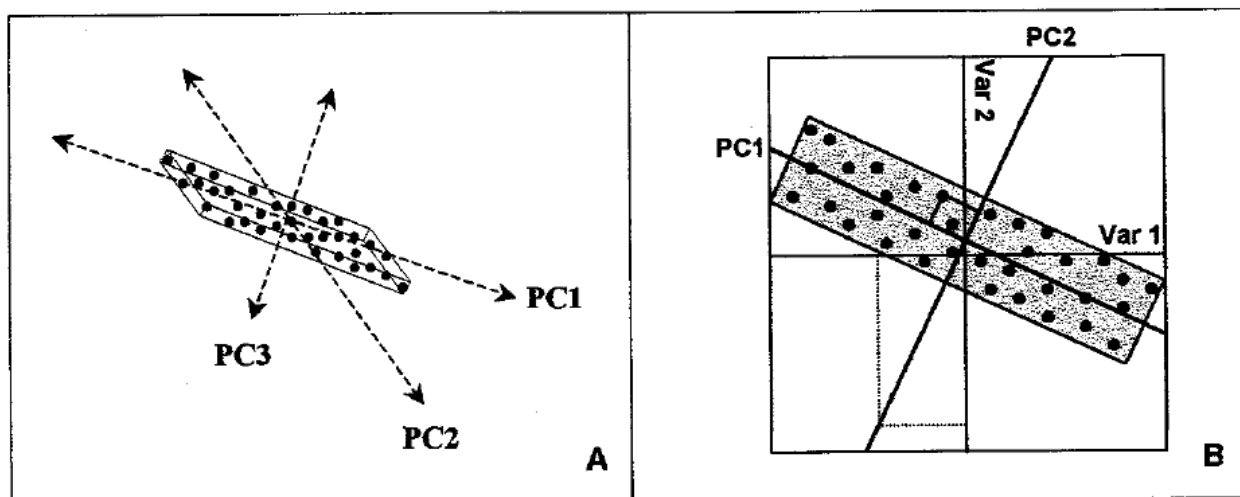


FIGURA 10 – A: GRÁFICO TRIDIMENSIONAL ILUSTRANDO OS EIXOS DAS COMPONENTES PRINCIPAIS e B: GRÁFICO BIDIMENSIONAL DA PC1 X PC2: ESCORES REPRESENTADOS POR (●) E LOADINGS POR (▭)
 FONTE: NAGATA (2001)

A PCA é aplicada em conjunto com dados gerados por diferentes técnicas analíticas nas mais diversas áreas, como por exemplo, na classificação de óleo de semente de abóbora, com base nas análises de NIR e UV-Vis (LANKMAYR *et al.*, 2004), na correlação entre o perfil de ácidos graxos e espectroscopia MID, a fim de classificar azeites de oliva em relação a diferentes safras e regiões geográficas (GURDENIZ; OZEN; TOKATLI, 2008), na avaliação das alterações durante o armazenamento de abacaxis fatiados usando métodos espectroscópicos como NIR e MID (DI EGIDIO *et al.*, 2009) e na identificação e caracterização de diferentes cultivares de oliveiras por meio da espectroscopia NIR (CASALE *et al.*, 2010).

2.4.5 Regressão por Mínimos Quadrados Parciais (PLSR)

O método multivariado aplicado quando se deseja realizar a previsão de uma variável dependente a partir de variáveis independentes é a regressão de mínimos quadrados parciais (PLSR), ou simplesmente, método dos mínimos quadrados parciais (PLS) (BASTIEN; ESPOSITO-VINZI; TENENHAUS, 2005).

A base do método PLSR está na Análise de Componentes Principais (PCA). A PCA decompõe a matriz de dados X , construindo-se um novo sistema de eixos, em termos da soma de várias matrizes M_i , que não podem mais ser expandidas, mais uma matriz de erros (que corresponde a parte não modelada de X). As matrizes M_i constituem as chamadas componentes principais, e são formadas pelo produto de dois vetores, t (*scores*) e p (*loadings*) (ABDI, 2003; NAGATA, BUENO, PERALTA-ZAMORA, 2001; FERREIRA *et al.*, 1999):

$$X = M_1 + M_2 + \dots + M_a + E \text{ ou} \quad (2)$$

$$X = t_1 p_1 + t_2 p_2 + \dots + t_a p_a + E \text{ ou} \quad (3)$$

$$X = TP' + E \quad (4)$$

Com a obtenção do novo sistema de eixos, o método PLSR pode ser descrito da seguinte forma: tanto a matriz das variáveis independentes X (Eq. 5) como a das variáveis dependentes Y (Eq. 6) sofrem decomposição matricial e são representadas pelos *scores* (T e U) e *loadings* (P e Q):

$$X = TP' + E \quad (5)$$

$$Y = UQ' + F \quad (6)$$

A relação entre as duas matrizes de *scores*, das variáveis dependentes e independentes, utilizando um modelo linear, é representada pela Equação 7:

$$U = b \cdot T \quad (7)$$

A determinação correta do vetor b depende da escolha do número de variáveis latentes suficientes para explicar toda a informação analítica necessária e da detecção de amostras anômalas com a sua subsequente remoção (NAGATA; BUENO; PERALTA-ZAMORA, 2001).

O método de regressão por mínimos quadrados parciais (PLSR) tem sido cada vez mais aplicado na área de alimentos, como por exemplo, no monitoramento da degradação hidrolítica e oxidativa do teor de ácidos graxos livres em óleo de peixe (COZZOLINO *et al.*, 2005), na determinação da composição de ácidos em sementes de canola (KIM *et al.*, 2007), na quantificação de ácido oleico em sementes de girassol por reflectância difusa (CANTARELLI *et al.*, 2009), bem como na determinação de ácidos graxos em grãos de soja (PATIL *et al.*, 2010), entre outros.

2.4.6 Validação

Após a construção dos modelos de calibração é necessário avaliar a qualidade dos mesmos através da etapa de validação. Esta etapa consiste em testar o modelo, utilizando-o para previsão do teor das amostras a fim de estabelecer se ele de fato irá refletir o comportamento do analito de interesse (FERREIRA *et al.*, 1999).

Pode-se validar um modelo por meio da validação cruzada (*cross validation*), método em que se emprega um conjunto de teste gerado a partir do mesmo conjunto de calibração, ou através da validação externa, onde se utiliza um conjunto de amostras separado e diferente do conjunto de calibração.

2.4.6.1 Validação cruzada

A validação cruzada é uma técnica de validação baseada nos dados de calibração, sendo usada para avaliar a capacidade de previsão dos modelos para o conjunto de amostras. Neste tipo de validação, pode-se aplicar diferentes métodos

de análise, contudo, para um conjunto com poucas amostras, é recomendável o uso do método “*leave one out*”, onde se eliminam sucessivamente amostras do próprio conjunto de calibração. Desta forma, a primeira amostra dos dados de calibração é eliminada. Em seguida, a calibração é efetuada usando as amostras restantes. Este procedimento é realizado para todas as amostras, sendo armazenado o modelo que revelou melhor desempenho, ou seja, menor erro (NAES *et al.*, 2002).

O erro calculado através desta técnica denomina-se RMSECV (Raiz Quadrada do Erro Médio de Validação Cruzada – *Root Mean Squared Error of Cross Validation*) e o mesmo é calculado com base no valor de referência (y_i) e no valor estimado (\hat{y}_i), onde N corresponde ao número de amostras do modelo (DAGNEW; CROWE; SCHOENAU, 2001).

$$\text{RMSECV} = \sqrt{\sum_{i=1}^N \frac{(\hat{y}_i - y_i)^2}{N - 1}} \quad (8)$$

A interpretação dos dados por validação cruzada precisa ser cautelosa porque os erros podem estar relacionados a distorções nos espectros de amostras puras (REEVES, 2003), pode haver inclusão de um número de variáveis latentes acima do ideal, levando a um super ajuste do modelo, ou mesmo falhar, ao não modelar a informação correspondente a pequenas variações nos dados (XU; LIANG, 2001).

2.4.6.2 Validação externa

A verificação da capacidade preditiva dos modelos no processo de validação externa é realizada a partir de um conjunto de amostras diferente do conjunto utilizado na calibração, em que calcula-se a Raiz Quadrada do Erro Médio de Previsão (RMSEP – *Root Mean Squared Error Prediction*) a partir dos valores de referência (y_i), previsão (\hat{y}_i) para a propriedade de interesse, e amostras de calibração (N_p) (PASQUINI, 2003).

$$\text{RMSEP} = \sqrt{\sum_{i=1}^N \frac{(\hat{y}_i - y_i)^2}{N_p}} \quad (9)$$

2.4.7 Amostras anômalas (*Outliers*)

Como mencionado anteriormente, para a construção de um modelo de calibração multivariada, geralmente usam-se dois conjuntos de amostras, um para a calibração e outro para a validação do modelo. Entre estas amostras, pode haver amostras anômalas (*“outliers”*), ou seja, amostras que apresentam comportamento diferente das demais. A presença de amostras anômalas entre as amostras de calibração pode gerar modelos com uma baixa capacidade de previsão. E, quando presentes entre as amostras de validação, estas amostras anômalas podem influenciar os resultados, podendo indicar que o modelo não é adequado. Podem-se utilizar grandezas na detecção de amostras anômalas, entre elas a *“Leverage”* e os Resíduos de *“Student”* (VANDEGINSTE *et al.*, 1998).

A medida que indica o grau de influência de uma amostra no modelo de regressão é o *“leverage”*. Caso uma amostra apresente um valor baixo de *“leverage”*, esta influencia pouco o modelo, contudo, se medidas experimentais de uma amostra forem muito diferentes das outras amostras, esta poderá influenciar negativamente o conjunto de calibração. O valor de limite de *“leverage”* pode ser obtido pela equação 10 (onde k representa o número de variáveis latentes e n o número de amostras) e representa o grau de afastamento da amostra em relação ao centro ou média dos dados (FERREIRA *et al.*, 1999).

$$\text{limite de leverage} = 3 \cdot k/n \quad (10)$$

Supondo-se que os Resíduos de *“Student”* estejam normalmente distribuídos, pode-se aplicar um teste t para avaliar se a amostra está ou não dentro da distribuição com nível de 95% de probabilidade. Como os Resíduos de *“Student”* são definidos em unidades de desvio padrão do valor médio, os valores além de $\pm 2,5$ são considerados altos sob as condições usuais da estatística. A análise gráfica

dos Resíduos de “*Student*” versus “*leverage*” para cada amostra é o método mais adequado para identificar amostras anômalas (FERREIRA *et al.*, 1999).

3 OBJETIVOS

3.1 OBJETIVO GERAL

Quantificar os ácidos graxos: oleico, linoleico e linolênico presentes nas sementes e farinhas de linhaça das variedades dourada e marrom por meio de cromatografia gasosa e desenvolver modelos de regressão com base na espectroscopia infravermelho médio (MID) e próximo (NIR).

3.2 OBJETIVOS ESPECÍFICOS

- a) Determinar o teor de lipídeos totais em sementes e em farinhas de linhaça das variedades dourada e marrom.
- b) Identificar os ácidos oleico, linoleico e linolênico presentes em sementes e nas farinhas de linhaça das variedades dourada e marrom através de cromatografia gasosa acoplada a espectrometria de massas.
- c) Quantificar os ácidos oleico, linoleico e linolênico presentes em sementes e farinhas de linhaça das variedades dourada e marrom por meio de cromatografia gasosa.
- d) Identificar potenciais diferenças entre as sementes e as farinhas das variedades de linhaça dourada e marrom, quanto ao teor de ácidos graxos, através de Análise de Componentes Principais (PCA).
- e) Desenvolver modelos de calibração multivariada, empregando o método PLSR, para a quantificação dos ácidos graxos presentes em sementes e farinhas de linhaça das variedades dourada e marrom utilizando espectroscopia no infravermelho por refletância difusa nas regiões do infravermelho próximo (NIR) e médio (MID).
- f) Validar os modelos de calibração multivariada desenvolvidos, por meio de validação externa.

- g) Comparar os modelos de NIR e MID no que diz respeito a sua eficiência para a previsão dos teores dos ácidos oleico, linoleico e linolênico presentes em sementes e farinhas de linhaça das variedades dourada e marrom.

4 MATERIAL E MÉTODOS

4.1 MATERIAL

As amostras de semente e farinha de linhaça das variedades dourada e marrom foram doadas pelas Empresas Jasmine Ltda. e Vitao Ltda. (Curitiba, PR). Todas as amostras foram mantidas em suas embalagens originais e acondicionadas em temperatura ambiente até o momento das análises.

Para a determinação do teor de ácidos graxos utilizou-se 78 amostras, separadas pelas denominações A e B, respectivamente, empresas Jasmine Ltda. e Vitao Ltda. subdivididas em sete grupos, conforme Tabela 5.

TABELA 5 – CODIFICAÇÃO DAS AMOSTRAS DE LINHAÇA DOURADA E MARROM

Quantidade de amostras	Codificação	Identificação
12	FLDA	Farinha de linhaça dourada estabilizada
14	FLDB	Farinha de linhaça dourada
14	FLMB	Farinha de linhaça marrom
5	SDOA	Semente de linhaça dourada orgânica
5	SLDA	Semente de linhaça dourada
14	SLDB	Semente de linhaça dourada
14	SLMB	Semente de linhaça marrom

4.2 MÉTODOS

4.2.1 Extração dos lipídeos totais

Para a extração dos lipídeos totais, foram pesadas 3 g das amostras (Tabela 5), colocadas em papel filtro qualitativo com gramatura 80 g.m⁻², espessura de 205 µm e poros de 14 µm, moídas em almofariz por 5 minutos e em seguida transferidas para cartuchos de extração. Os cartuchos foram colocados em extrator de Soxhlet

(Prodicil, Brasil) ao qual foi adaptado balão de vidro de 300 mL, onde foram colocados 200 mL de éter de petróleo P.A. (Carlo Erba, Itália). A extração foi realizada com aquecimento em banho-maria à 60°C por 2 horas.

Posteriormente, os cartuchos foram retirados do extrator e transferidos para um almofariz, para nova moagem por 5 minutos. Em seguida, os cartuchos contendo as amostras moídas foram submetidos a extração, conforme anteriormente descrito, por 3 horas. Finalmente, os balões contendo a fração lipídica foram resfriados e os extratos obtidos foram derivatizados para posterior quantificação, adaptado da *American Oil Chemistry Society (AOCS)* (2004).

4.2.2 Preparação de ésteres metílicos

Em tubos de vidro de 2 mL foram pesados 10 mg dos extratos lipídicos obtidos (item 4.2.1) e solubilizados em 1mL de n-hexano 95% (Mallinckrodt Chemicals, U.S.A.). Em seguida, para a derivatização adicionou-se diazometano (Sigma-Aldrich Co., Saint Louis, MO, U.S.A.) gota a gota até que a solução apresentasse coloração ligeiramente amarelada. As amostras foram mantidas em banho de gelo por cerca de 2 horas até que a reação terminasse, como indicado pela ausência de cor. Posteriormente, todas as amostras foram analisadas em cromatógrafo a gás (item 4.2.4) para quantificação.

4.2.3 Análise através de Cromatografia Gasosa associada à Espectrometria de Massas

Ésteres metílicos de ácidos graxos foram submetidos à análise em um Cromatógrafo a gás (Varian – 450, Agilent Technologies, Oxford, Reino Unido) acoplado a um Espectrômetro de Massas (Varian – 320). Para a cromatografia gasosa foi utilizada uma coluna capilar CP 8944 (J&W Scientific, Agilent Technologies, U.S.A) com 30m de comprimento, 0,25mm de diâmetro interno e 0,25µm de espessura. O gás de arraste foi o hélio (He), com vazão de 0,8 mL.min⁻¹.

com razão de *split* de 1:50 e volume de injeção, 1 μ L. A temperatura do forno da coluna foi programada a partir da temperatura inicial de 190 °C, sendo mantida nesta temperatura por 1 minuto. Em seguida, a temperatura foi elevada para 200 °C a uma taxa de 10 °C.min⁻¹. e permanecendo por 2 minutos à 200°C. A temperatura foi então elevada a 310°C a uma taxa de 3,5°C.min⁻¹, e permanecendo nesta temperatura por 1,57 minutos. O tempo total de análise foi de 37 minutos.

As temperaturas do injetor, do detector e da interface do detector do espectrômetro de massas foram de 315, 330 e 330°C, respectivamente. O detector de massas operou com ionização por impacto de elétrons de 70 eV e varredura de massas com base na razão massa/carga (*m/z*) no intervalo de 32 a 380 *m/z*.

Os ésteres metílicos dos ácidos graxos foram identificados pelos espectros de fragmentação de massas (Biblioteca *NIST – National Institute of Standards and Technology*) e a pela comparação dos tempos de retenção dos padrões dos ácidos oleico (*cis* 9-octadecenoico), linoleico (*cis* 9,12-octadecadienoico) e linolênico (*cis* 9,12,15-octadecatrienoico) (Sigma-Aldrich Chemical Co., Saint Louis, U.S.A.), grau de pureza \geq 99%.

4.2.4 Análise em Cromatografia Gasosa

Ésteres metílicos de ácidos graxos foram submetidos à análise em Cromatógrafo a gás GC 14B (Shimadzu, Co., Quioto, Japão), equipado com coluna capilar DB-23 (J&W Scientific, Agilent Technologies, U.S.A.) com 60m de comprimento, 0,25mm de diâmetro interno e 0,25 μ m de espessura (FIGURA 11). O gás de arraste foi o nitrogênio (N₂), com vazão de 0,71 mL.min⁻¹, velocidade linear de 14,5 m/s e pressão de 250 kPa. O volume de injeção foi de 1,0 μ L para ésteres metílicos obtidos a partir das amostras de farinha e de 2,0 μ L obtidos a partir das sementes, com razão de *split* de 1:63. As temperaturas do injetor, da coluna e do detector foram de 235, 220 e 260°C, respectivamente, sendo que a análise ocorreu em isoterma de 220°C por 11 minutos. A quantificação dos ésteres metílicos dos ácidos graxos foi realizada por normalização de área, em que os percentuais dos ácidos graxos foram determinados com base na área relativa de cada pico, em função da área total dos picos identificados.



FIGURA 11 – CROMATÓGRAFO A GÁS (SHIMADZU GC 14B)
FONTE: O autor (2012)

4.2.5 Análise de espectroscopia no infravermelho próximo e médio

Os espectros na região do infravermelho foram obtidos em espectrofotômetro BRUKER – TENSOR 37 (Bruker Optics, Ettlingen, Germany), um espectrofotômetro com Transformada de Fourier e esfera de integração para análise de reflectância difusa (FIGURA 12), em conjunto com o programa de aquisição espectral OPUS versão 6.0 (Bruker Optics, Ettlingen, Germany). Na região do infravermelho próximo os espectros obtidos têm número de onda entre 10.000 e 4.000 cm^{-1} , e os espectros na região do infravermelho médio foram adquiridos entre 10.000 e 400 cm^{-1} . No entanto, para o desenvolvimento dos modelos de calibração foram eliminadas as regiões que apresentavam pouca informação analítica relevante, já que o ruído era predominante. Desta forma, para a região do infravermelho próximo foi utilizada a região entre 9.000 e 4.000 cm^{-1} e para a região do infravermelho médio foram utilizados espectros entre 4.000 e 750 cm^{-1} , ambos com resolução de 4 cm^{-1} , acumulando-se 128 varreduras.



FIGURA 12 – ESPECTROFOTOMETRO FTNIR/MID
FONTE: O autor (2011)

4.2.6 Desenvolvimento dos modelos de regressão

Para a elaboração das matrizes de dados espectrais e de teor dos ácidos graxos de interesse utilizou-se o software Origin Pro 8.0 (Northampton, MA 01060, U.S.A.) enquanto que para os modelos de previsão foi empregado o pacote PLSR - Toolbox 3.0 (Eigenvector Research Inc., Manson, WA), que opera em ambiente Matlab 7.0 (Mathworks Inc., Natick, MA, USA).

5 RESULTADOS E DISCUSSÃO

5.1 PERFIL LIPÍDICO DAS FARINHAS E SEMENTES DE LINHAÇA

Inicialmente determinou-se o teor de lipídeos totais para as 78 amostras de farinhas e sementes de linhaça.

As 26 amostras de farinhas de linhaça dourada analisadas apresentaram valores para lipídeos totais entre 13,1 e 39,6 g.100 g⁻¹, com coeficiente de variação de 46,78%, indicando uma elevada heterogeneidade entre as amostras em relação a este parâmetro. Dentre as 14 amostras de farinha de linhaça marrom, o teor de lipídeos totais variou entre 13,58 e 16,15 g.100 g⁻¹, com coeficiente de variação de 4,53%, indicando assim baixa variabilidade entre as amostras, sendo este grupo de amostras considerado como um grupo homogêneo.

Os teores de lipídeos totais para as amostras de semente de linhaça dourada e semente de linhaça marrom apresentaram baixos coeficientes de variação, 10,81% e 8,32% respectivamente. Para semente dourada, o teor de lipídeos totais variou entre 23,41 e 36,64 g.100 g⁻¹. Enquanto que os dados referentes às sementes de linhaça marrom apresentaram valores entre 22,67 e 29,47 g.100 g⁻¹ para o teor de lipídeos totais.

A composição química da semente e, conseqüentemente da farinha de linhaça é influenciada principalmente pelas características genéticas da planta, condições ambientais de cultivo e métodos de análise (DAUN *et al.*, 2003). Tendo em vista que foi empregado um único método para análise dos teores de lipídeos totais em todas as amostras utilizadas neste estudo, variações genéticas e diferenças nas condições de cultivo, como por exemplo, a temperatura, possivelmente, são responsáveis pelas diferenças observadas entre as 78 amostras de linhaça, quanto o teor de lipídeos totais.

5.1.1 Identificação dos Ésteres Metílicos de Ácidos Graxos

Com intuito de identificar e quantificar os principais ácidos graxos presentes nas amostras de farinha e semente de linhaça das variedades dourada e marrom utilizou-se a cromatografia gasosa acoplada à espectrometria de massas (CG-EM), e a cromatografia gasosa com detector de ionização de chama (CG-DIC), respectivamente.

A identificação dos ésteres metílicos de ácidos graxos presentes nas amostras foi realizada através da comparação do tempo de retenção dos picos presentes no cromatograma com o tempo de retenção dos padrões dos ácidos oleico, linoleico e linolênico. Além da comparação do tempo de retenção, também foram comparados os espectros de massas das amostras com os espectros de massas dos padrões e com os espectros de fragmentação de massas do banco de dados do sistema *NIST (National Institute of Standards and Technology)*, conforme FIGURA 13.

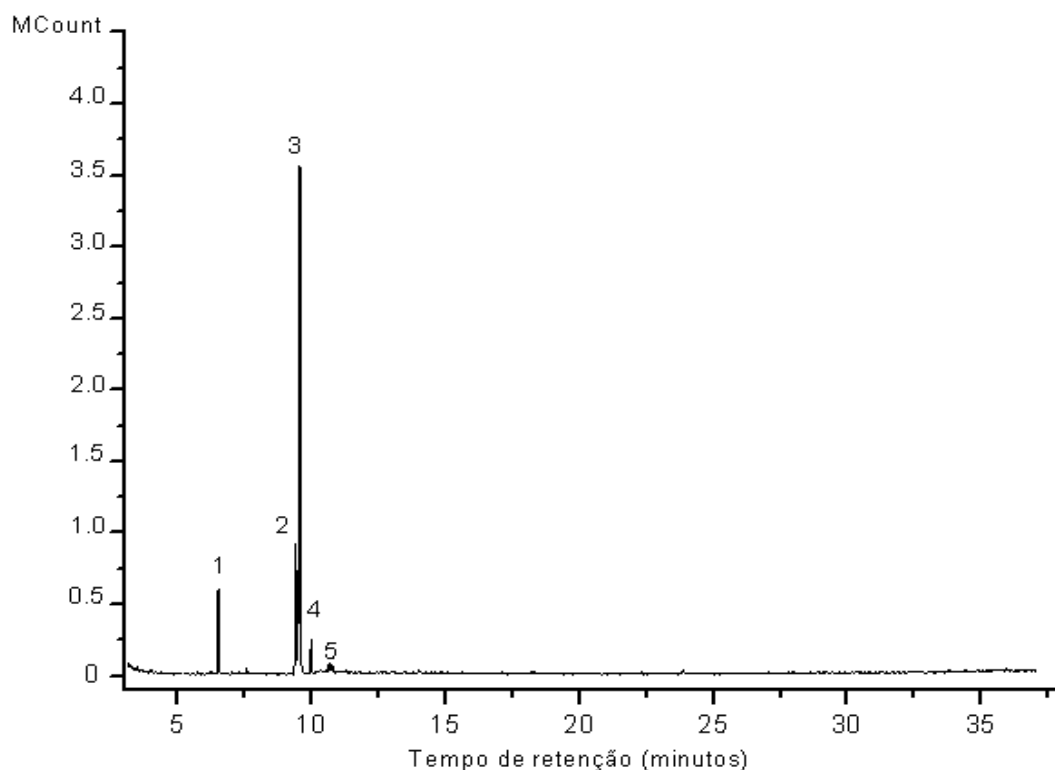


FIGURA 13 – CROMATOGRAMA PARA IDENTIFICAÇÃO DOS PRINCIPAIS ÉSTERES METÍLICOS DE ÁCIDOS GRAXOS OBTIDOS DE AMOSTRA DE FARINHA DE LINHAÇA DOURADA (FLDA 21) POR CROMATOGRAFIA GASOSA ACOPLADA À ESPECTROMETRIA DE MASSAS
NOTA: Compostos identificados: picos 1. C16:0 (Ácido palmítico); 2. C18:2 (Ácido linoleico); 3. C18:3 (Ácido linolênico); 4. C18:0 (Ácido Esteárico); 5. C18:1 (Ácido oleico)

Para a quantificação dos ésteres metílicos de ácidos graxos foi realizada normalização de área e os resultados expressos em porcentagem (%) do total de ácidos graxos presentes, de acordo com trabalhos descritos por Velasco e Becker (1998). Foram detectados cinco ácidos graxos principais, havendo a predominância dos ácidos graxos insaturados: ácido oleico (C18:1), linoleico (C18:2) e linolênico (C18:3), perfazendo em média 88% do total. Com relação aos ácidos graxos saturados, observou-se a presença dos ácidos palmítico (16:0) e esteárico (18:0) (Tabela 6). A FIGURA 14 mostra o cromatograma obtido para ésteres metílicos de ácidos graxos obtidos a partir da farinha de linhaça dourada (FLDA 21). Cromatogramas semelhantes foram encontrados para farinhas de linhaça marrom, assim como para sementes de linhaça dourada e marrom.

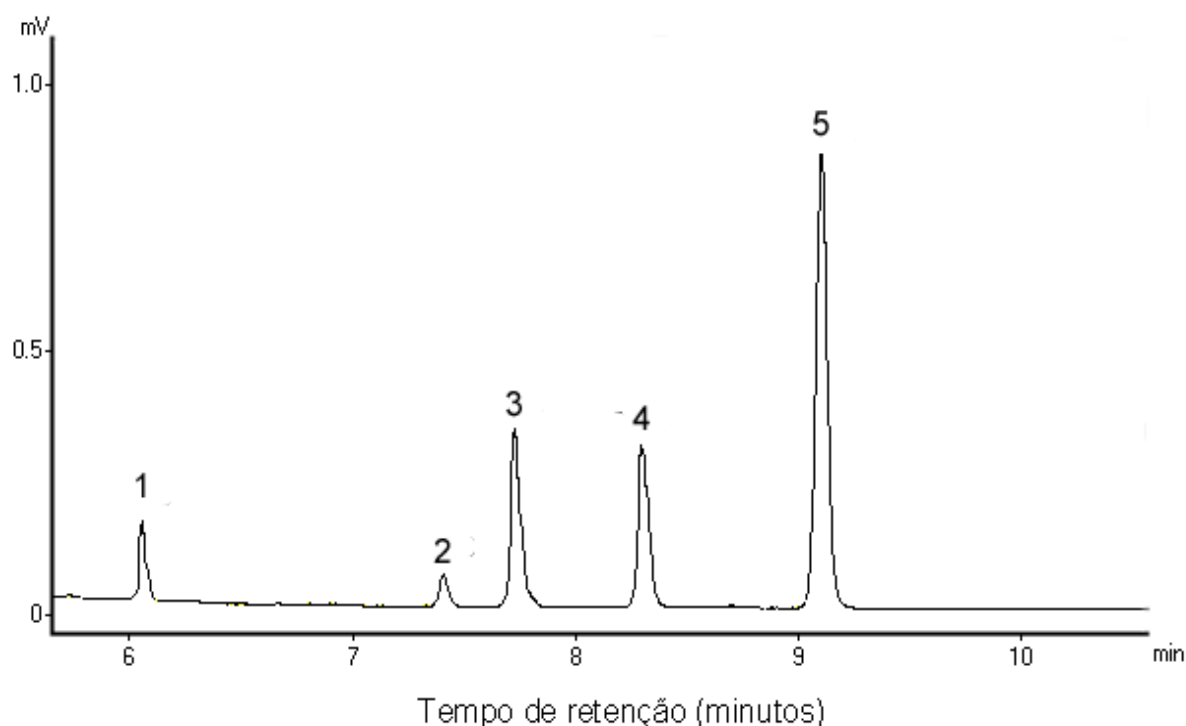


FIGURA 14 – CROMATOGRAMA PARA QUANTIFICAÇÃO DOS PRINCIPAIS ÉSTERES METÍlicos DE ÁCIDOS GRAXOS OBTIDOS DE AMOSTRA DE FARINHA DE LINHAÇA DOURADA (FLDA 21) POR CROMATOGRAFIA GASOSA COM DETECTOR DE IONIZAÇÃO DE CHAMA
NOTA: Compostos identificados: picos 1. C16:0 (ácido palmítico); 2. C18:0 (ácido esteárico); 3. C18:1 (ácido oleico); 4. C18:2 (ácido linoleico); 5. C18:3 (ácido linolênico)

TABELA 6 – TEORES DE ÁCIDO OLEICO, ÁCIDO LINOLEICO E ÁCIDO LINOLÊNICO OBTIDOS A PARTIR DAS AMOSTRAS DE LINHAÇA DOURADA E MARROM (continua)

Amostra	Código	Ácido Oleico (%)	Ácido Linoleico (%)	Ácido Linolênico (%)
1	FLDA 11	18,66	18,24	54,04
2	FLDA 12	18,41	18,11	54,93
3	FLDA 13	18,55	18,28	54,33
4	FLDA 21	18,34	18,15	54,1
5	FLDA 22	18,5	18,07	54,09
6	FLDA 23	18,17	18,02	54,14
7	FLDA 31	18,42	17,93	55,13
8	FLDA 32	18,76	18,24	55,62
9	FLDA 33	19,03	18,95	55,31
10	FLDA 41	18,6	18,42	54,05
11	FLDA 42	18,05	18,43	54,59
12	FLDA 43	18,1	18,61	54,41
13	FLDB 1	22,98	23,04	43,36
14	FLDB 2	23,42	22,82	42,52
15	FLDB 31	23,5	22,91	42,67
16	FLDB 32	23,38	22,94	43,27
17	FLDB 33	23,35	23,01	43,22
18	FLDB 4	23,35	22,79	42,89
19	FLDB 5	23,99	23,57	44,01
20	FLDB 6	23,58	23,49	43,52
21	FLDB 7	23,29	23,12	44,18
22	FLDB 8	23,61	22,32	44,31
23	FLDB 9	23,67	23,09	43,35
24	FLDB10	23,3	23,22	43,18
25	FLDB 11	23,67	23,26	43,48
26	FLDB 12	23,38	23,17	43,76
27	FLMB 1	22,8	19,99	43,59
28	FLMB 2	22,84	20,14	43,07
29	FLMB 31	23,44	19,85	43,46
30	FLMB 32	23,13	19,72	42,81
31	FLMB 33	23,25	19,82	43,39
32	FLMB 4	23,26	19,82	43,56
33	FLMB 5	23,26	19,79	43,94
34	FLMB 6	23,15	19,92	44,09
35	FLMB 7	23,17	19,69	43,7
36	FLMB 8	23,28	19,82	43,42
37	FLMB 9	23,38	19,04	44,92
38	FLMB 10	23,69	19,39	44,07
39	FLMB 11	23,22	19,73	43,81
40	FLMB 12	23,35	19,71	44,29

TABELA 6 – TEORES DE ÁCIDO OLEICO, ÁCIDO LINOLEICO E ÁCIDO LINOLÊNICO OBTIDOS A PARTIR DAS AMOSTRAS DE LINHAÇA DOURADA E MARROM (conclusão)

Amostra	Código	Ácido Oleico (%)	Ácido Linoleico (%)	Ácido Linolênico (%)
41	SDOA 1	20,21	20,06	48,29
42	SDOA 2	20,78	20,17	48,7
43	SDOA 3	20,62	20,14	47,98
44	SDOA 4	21,07	20,43	48,57
45	SDOA 5	20,5	20,29	48,19
46	SLDA1	21,3	20,95	48,23
47	SLDA 2	21,29	20,74	48,28
48	SLDA 3	22,49	20,4	48,17
49	SLDA 4	21,06	20,15	48,28
50	SLDA 5	21,14	20,93	48,17
51	SLDB 1	22,17	18,06	47,63
52	SLDB 2	21,96	17,39	47,49
53	SLDB 31	22,58	17,26	47,26
54	SLDB 32	22,23	17,46	47,41
55	SLDB 33	21,96	17,79	47,28
56	SLDB 4	21,99	17,49	47,81
57	SLDB 5	22,34	18,75	47,05
58	SLDB 6	22,21	17,85	47,16
59	SLDB 7	22,46	18,14	46,9
60	SLDB 8	22,3	18,01	48,36
61	SLDB 9	22,16	18,21	46,92
62	SLDB 10	22,47	17,66	46,99
63	SLDB 11	22,44	18,19	47,02
64	SLDB 12	22,53	18,05	47,1
65	SLMB 1	21,3	15,78	48,36
66	SLMB 2	21,59	16,61	48,24
67	SLMB 31	21,63	15,56	48,15
68	SLMB 32	21,58	15,88	47,36
69	SLMB 33	21,37	15,47	47,8
70	SLMB 4	21,97	15,99	47,23
71	SLMB 5	21,85	15,39	48,16
72	SLMB 6	21,69	16,01	48,06
73	SLMB 7	21,98	16,83	48,24
74	SLMB 8	21,32	15,9	48,51
75	SLMB 9	21,44	16,17	47,26
76	SLMB 10	22,08	16,16	47,08
77	SLMB 11	21,46	16,33	47,53
78	SLMB 12	21,67	16,81	47,76

Os teores de ácido oleico para todas as amostras, tanto de farinhas como de sementes, apresentaram baixa variabilidade. Para as farinhas de linhaça dourada, o teor mínimo de ácido oleico foi de 18,05% e o máximo de 24%, enquanto que para as farinhas de linhaça marrom, os teores deste ácido variaram entre 22,8 a 23,70%. Teores semelhantes também foram observados para as sementes, tanto de linhaça dourada como de marrom, que apresentaram teores de ácido oleico entre 20,21 a 22,58% e 21,30 a 22,80%, respectivamente. O coeficiente de variação na

determinação do teor de ácido oleico para as farinhas de linhaça dourada e marrom foi de 10,42%, enquanto que para as sementes de linhaça foi de 2,81%. Em função desta baixa variabilidade, considerou-se o conjunto de amostras como sendo um conjunto homogêneo em relação ao teor de ácido oleico.

Os teores de ácido linoleico para farinhas de linhaça dourada variaram entre 17,93 e 23,57%, enquanto que para as farinhas de linhaça marrom houve variação de 19,04 a 20,14%. O coeficiente de variação para as farinhas de linhaça dourada e marrom foi de 9,92%. As sementes de linhaça dourada apresentaram teores de ácido linoleico similares aos das farinhas de linhaça dourada, variando de 17,26 a 20,95%. Nas sementes de linhaça marrom foram observados teores de ácido linoleico entre 15,39 e 16,83%. O coeficiente de variação para o teor de ácido linoleico nas sementes de linhaça foi de 9,94%. Os resultados encontrados para ácido linoleico nas amostras de farinha e semente de linhaça dourada são condizentes com os obtidos por Pitat e Zadernowski (2010), que descreveram valores entre 22,73 e 21,62% para este parâmetro em semente e farinha de linhaça dourada, respectivamente.

Para os teores de ácido linolênico, as farinhas de linhaça dourada apresentaram elevada variabilidade, visto que os teores apresentavam-se entre 42,52 e 55,62%. Lazzari e Chiantore (1999) também obtiveram valores elevados de ácido linolênico entre 48 e 60% em amostras de linhaça. Assim como Pitat e Zadernowski (2010), que relataram valores entre 50 e 53,17% em semente e farinha de linhaça dourada, respectivamente. Os teores de ácido linolênico para as sementes douradas variaram entre 46,90% e 48,7%, enquanto que os teores deste ácido para as farinhas e sementes de linhaça marrom variaram de 42,81 a 44,92% e 47,08 a 48,51%, respectivamente. Estas pequenas variabilidades em teor de ácido linolênico indicam que estas amostras são homogêneas em relação ao teor de ácido linolênico.

Os limites de detecção (LD) e quantificação (LQ) para os ácidos oleico, linoleico e linolênico foram realizados em CG/DIC, onde para o ácido oleico, os valores calculados de LD e LQ foram de 2,08 mg.L⁻¹ e 4 mg.L⁻¹, respectivamente, enquanto que o ácido linoleico apresentou valores de LD e LQ foram de 0,5 mg.L⁻¹ e 1 mg.L⁻¹, respectivamente, e para o ácido linolênico foram obtidos LD e LQ de 2 mg.L⁻¹ e 4 mg.L⁻¹, respectivamente.

5.2 ANÁLISE DAS COMPONENTES PRINCIPAIS (PCA)

A análise das componentes principais (PCA) é uma técnica utilizada para demonstrar variações nos dados multivariados. Ela foi empregada para agrupar amostras de acordo com suas similaridades e correlacionar as variações espectrais das diferentes amostras de linhaça, na região do infravermelho próximo (NIR) e médio (MID).

5.2.1 Análise de Componentes Principais de Espectros no Infravermelho Próximo

As 78 amostras de linhaça foram avaliadas com base nos espectros NIR na região espectral entre 9.000 e 4.000 cm^{-1} (FIGURA 15). Para análise da PCA, os dados espectrais foram pré-processados aplicando-se a correção do espalhamento multiplicativo (MSC), dados centrados na média (DCM) e alisamento com 1ª derivada (algoritmo Savitzky – Golay) com tamanho da janela móvel de 11 pontos.

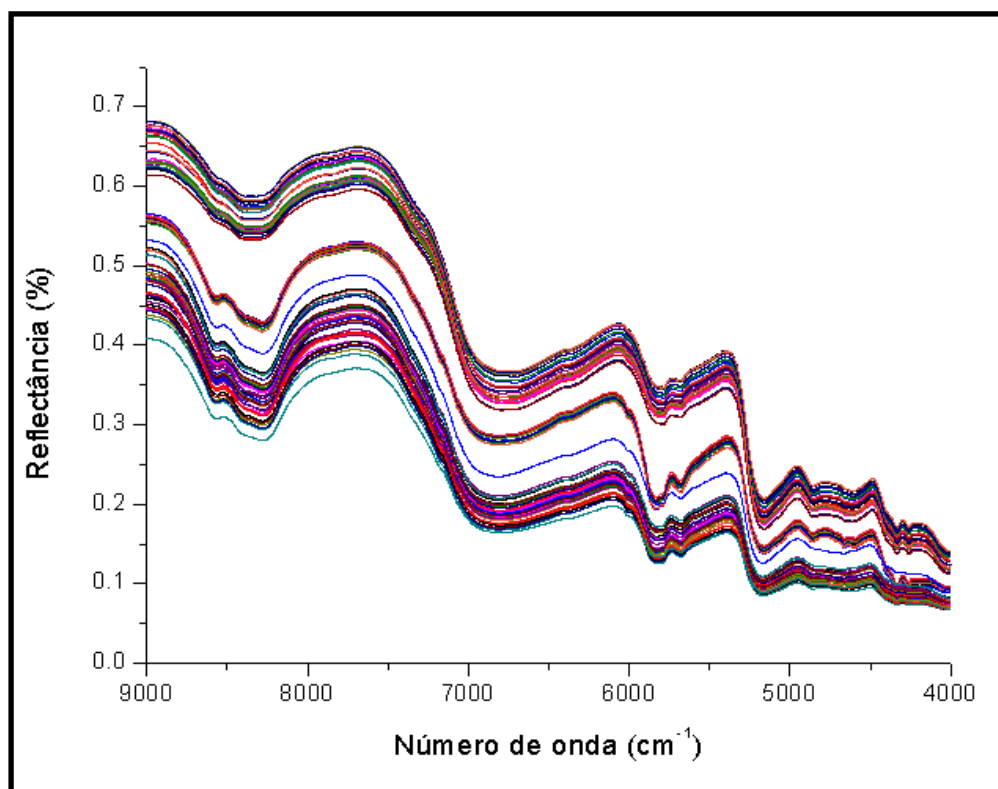


FIGURA 15 – ESPECTROS NIR PARA 78 AMOSTRAS DE LINHAÇA

As três primeiras componentes principais (PC) foram utilizadas para análise exploratória dos dados, tendo as mesmas explicado a maior parte (98,96%) da variância dos dados espectrais (FIGURA 16).

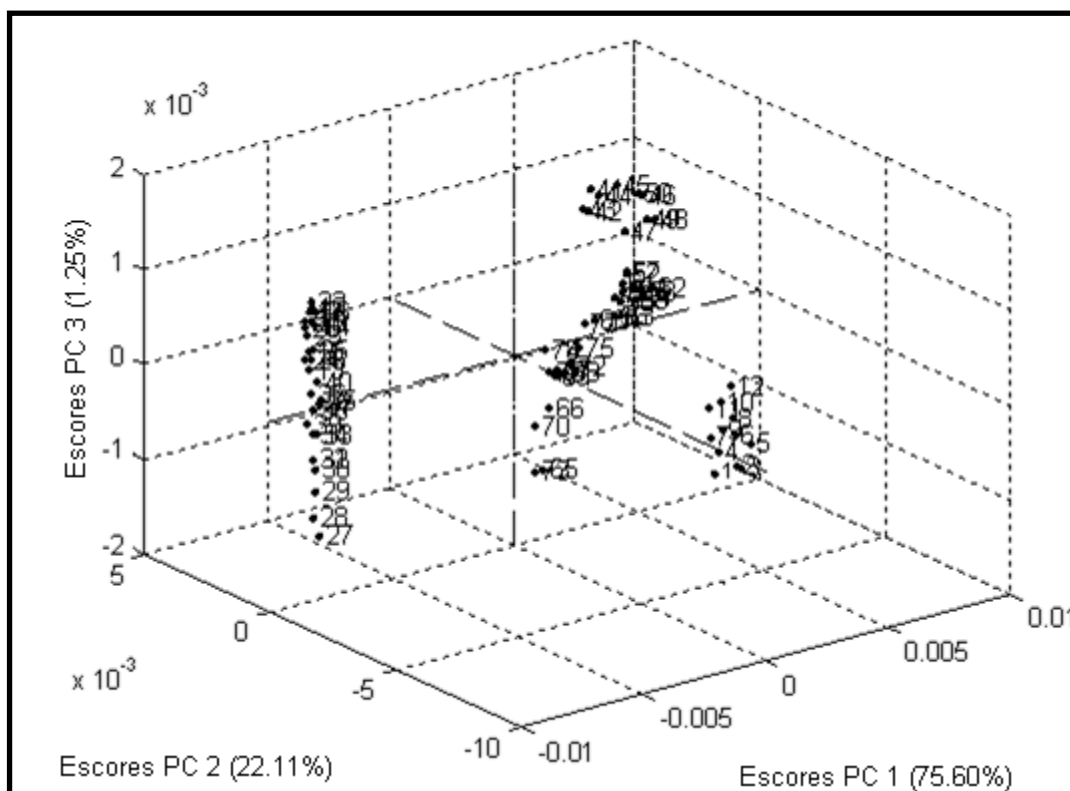


FIGURA 16 – GRÁFICO TRIDIMENSIONAL DOS ESCORES DAS COMPONENTES PRINCIPAIS PC1 x PC2 x PC3 NA REGIÃO DO NIR

No entanto, somente a partir da análise do gráfico de escores da PC1 x PC3 (FIGURA 17) foi possível identificar semelhanças e diferenças entre as amostras. Estas duas componentes principais explicaram 76,85% da variância observada para os dados espectrais, sendo que a primeira componente principal (PC1) explicou 75,60% enquanto que a terceira componente principal (PC3) explicou 1,25% da variância. Pode-se observar a separação das 78 amostras de linhaça em sete grupos (A, B, C, D, E, F e G). As amostras com maiores teores de lipídeos e, conseqüentemente de ácidos graxos, estão localizadas nos valores positivos de escores na PC1, enquanto que as amostras com os menores teores encontram-se nos escores negativos.

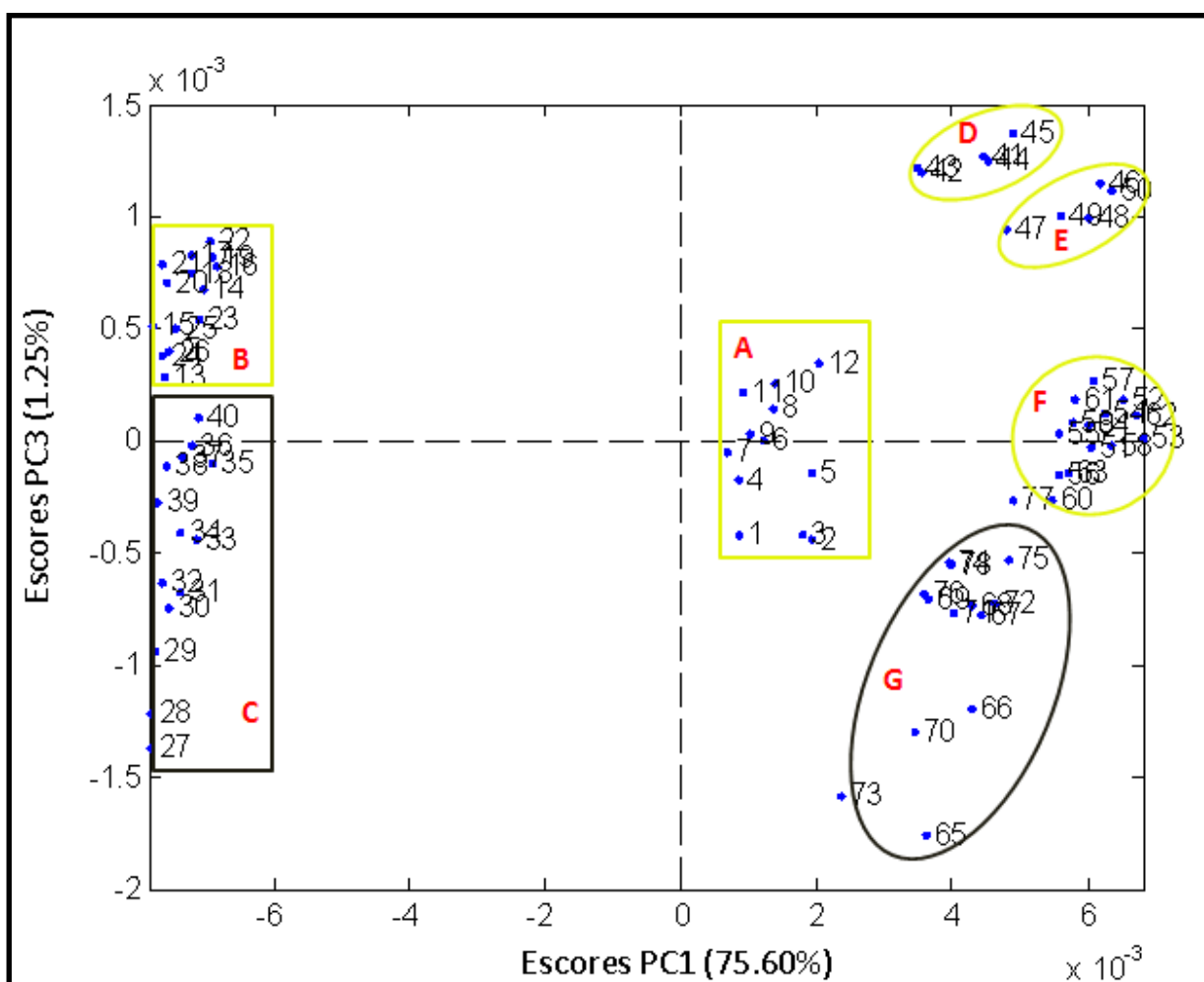


FIGURA 17 – GRÁFICO DE ESCORES DA PC1 x PC3 NA REGIÃO NIR

NOTA: Grupos identificados: A. Farinha de linhaça dourada (Empresa A); B. Farinha de linhaça dourada (Empresa B); C. Farinha de linhaça marrom (Empresa B); D. Semente de linhaça dourada orgânica (Empresa A); E. Semente de linhaça dourada (Empresa A); F. Semente de linhaça dourada (Empresa B); G. Semente de linhaça marrom (Empresa B)

Para evidenciar as regiões espectrais do NIR responsáveis por apontar as diferenças entre amostras em relação aos teores de lipídeos e ácidos graxos utilizou-se o gráfico de *loadings* (FIGURA 18), que mostra as regiões espectrais de maior representatividade para as tendências observadas anteriormente com base na PC1.

A análise do gráfico de escores juntamente com o gráfico de *loadings* indicou que o grupo de amostras com maiores teores de lipídeos e ácidos graxos foram caracterizadas por regiões do espectro com *loadings* positivos, onde se encontram os sinais com maior amplitude, conforme pode ser observado na Figura 18. As regiões em destaque representam as principais ligações químicas orgânicas consideradas pela análise de infravermelho, sendo a região 1 correspondente aos

números de onda entre 8.270 e 7.730 cm^{-1} com sinais característicos de ligações C – H, a região 2 com números de onda entre 5.415 a 5.145 cm^{-1} característicos de bandas de C = O e O – H e a região 3 com sinais de menor intensidade caracterizados por ligações C – H. Os menores teores de lipídeos e ácidos graxos estão na região de escores negativos, sendo representados por *loadings* negativos e caracterizados por sinais de baixa intensidade, sendo a região 4 correspondente aos números de onda entre 9.000 e 8.270 cm^{-1} cujas ligações são predominantemente de C – H (CEN; HE, 2007; XIABO *et al.*, 2010).

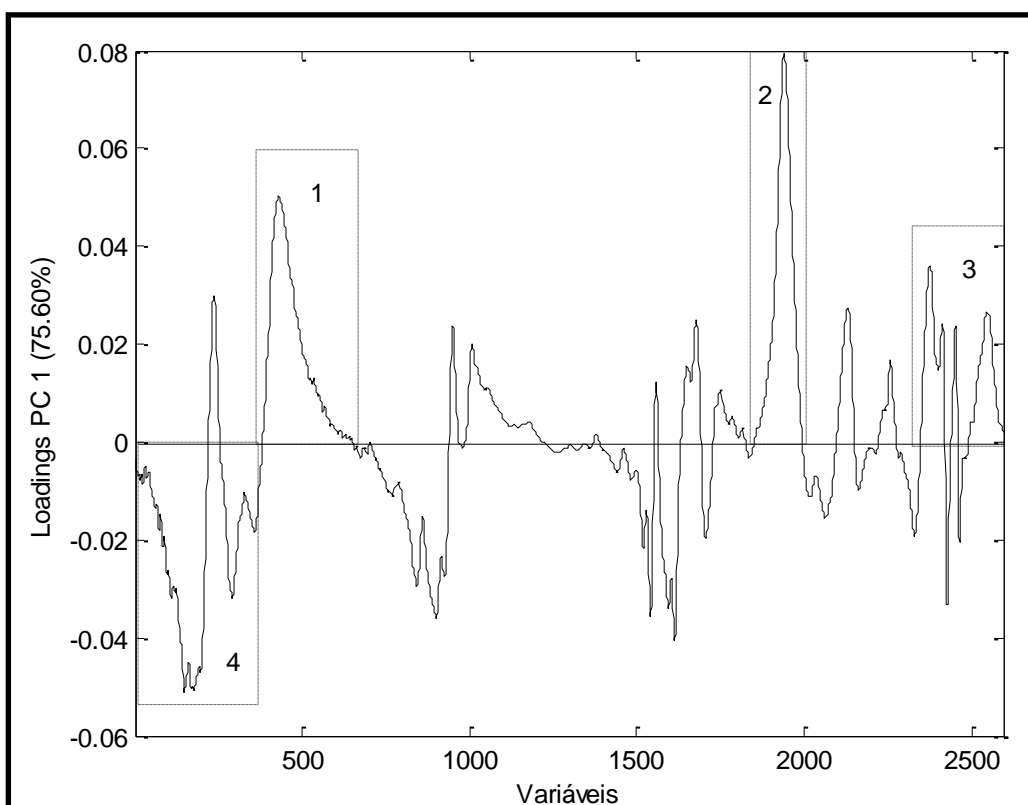


FIGURA 18 – GRÁFICO DE *LOADINGS* PARA A PC1 NA REGIÃO NIR

A PC3 é responsável pela discriminação das amostras em relação à variedade, pois dois grupos foram formados: um grupo formado pela variedade dourada, situado na parte positiva do gráfico de escores, e outro grupo representado pela variedade marrom, situado na parte negativa do gráfico de escores. Os sinais característicos destes grupos são evidenciados na FIGURA 19, com destaque para a região espectral 1 com número de onda entre 7.435 e 6.670 cm^{-1} com ligações orgânicas características de C – H, C = O e N – H, e a região espectral 2 com número de onda entre 5.455 a 5.145 cm^{-1} , onde encontram-se bandas de absorção das ligações C = O e O – H (CEN; HE, 2007; XIABO *et al.*, 2010). Tais vibrações são

relatadas na literatura como sendo características de bandas de ácidos carboxílicos e proteínas (SHIROMA; RODRIGUEZ-SAONA, 2009). Tendo em vista que a linhaça pode apresentar teor de proteínas entre 20 – 25% (COSKUNER; KARABABA, 2007), pode ser que a PC3 esteja considerando este parâmetro para a separação das variedades.

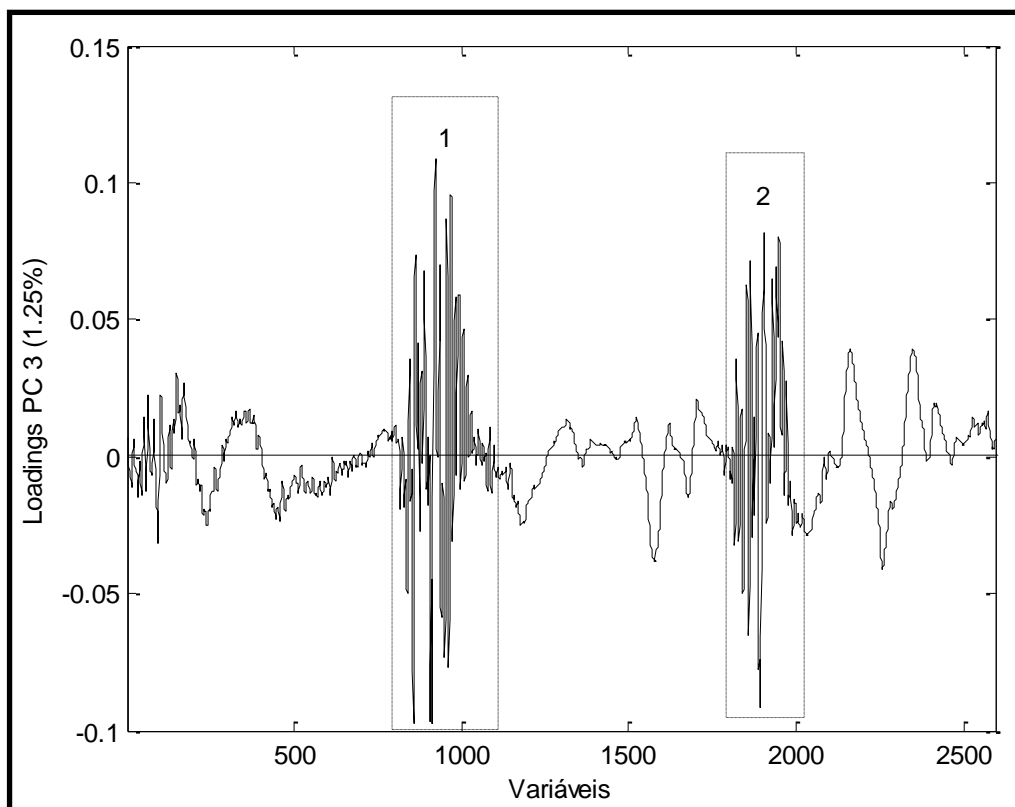


FIGURA 19 – GRÁFICO DE *LOADINGS* PARA A PC 3 NA REGIÃO NIR

5.2.2 Análise de Componentes Principais no Infravermelho Médio

As amostras de linhaça também foram avaliadas através dos espectros MID na região espectral entre 4.000 e 750 cm^{-1} (FIGURA 20). Os dados espectrais foram pré-processados, aplicando-se a transformação padrão de variação SNV (*Standard Normal Variate*) que corrige o espalhamento da luz decorrente da reflectância difusa e Detrend que permite a eliminação de tendências lineares com intuito de realizar análise por PCA.

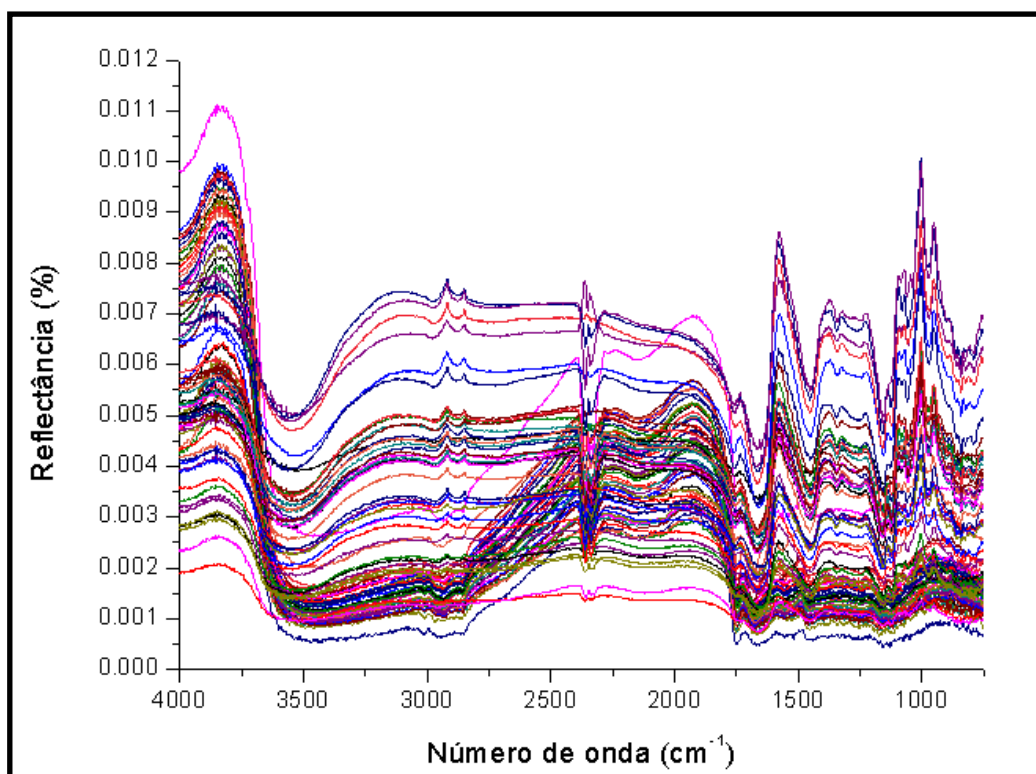


FIGURA 20 – ESPECTROS MID PARA 78 AMOSTRAS DE LINHAÇA

Para análise exploratória dos dados, inicialmente utilizou-se as três primeiras componentes principais (PC) que explicaram 99,45% da variância dos dados espectrais (FIGURA 21). Apesar da primeira componente principal (PC1) explicar a maior parte da variância (78,63%) dos dados, ela não diferencia as amostras, todas as 78 amostras apresentaram-se similares em relação a PC1. Portanto, fez-se análise apenas entre a PC2 x PC3.

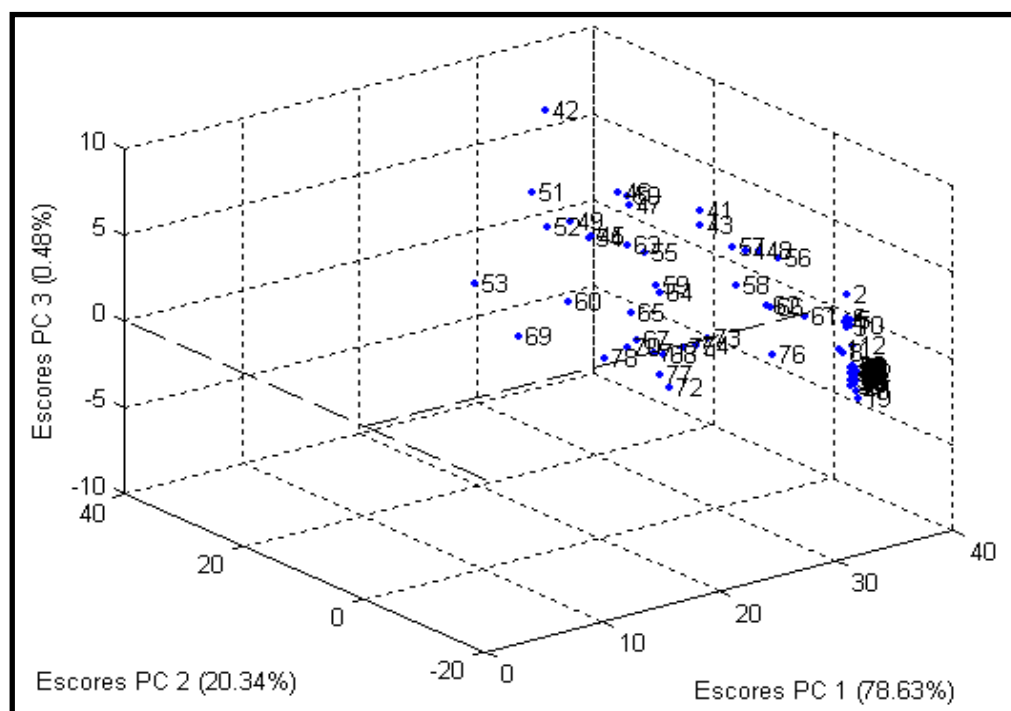


FIGURA 21 – GRÁFICO TRIDIMENSIONAL DOS ESCORES DAS COMPONENTES PRINCIPAIS PC1 x PC2 x PC3 NA REGIÃO DO MID

A análise de escores da PC2 x PC3 (FIGURA 22) possibilitou determinar apenas algumas diferenças entre as amostras, uma vez que estas duas componentes principais explicaram somente 20,82% da variância dos dados, sendo que a PC2 explicou 20,34% e a PC3 0,48%. Esta análise possibilitou a separação das 78 amostras de linhaça em quatro grupos (A, B, C e D), com destaque para o grupo B e C. Para o grupo B as farinhas de linhaça dourada e marrom (Empresa B) não apresentaram diferenças com relação ao teor de ácidos graxos. Enquanto que para o grupo C, a maioria das amostras de semente de linhaça dourada apresentaram-se dispersas, não sendo possível diferenciá-las quanto à procedência.

A segunda componente principal (PC2) possibilitou a discriminação entre as amostras de farinhas e sementes, onde as farinhas ficaram localizadas nos valores negativos dos escores, enquanto que as sementes ficaram localizadas nos valores positivos dos escores. A terceira componente principal (PC3) possibilitou a separação das amostras por variedades, onde a maioria das amostras da variedade dourada (grupos A e C) ficaram localizadas nos valores positivos de escores, enquanto que todas as amostras da variedade marrom (grupo D e parte do grupo B) localizaram-se nos valores negativos.

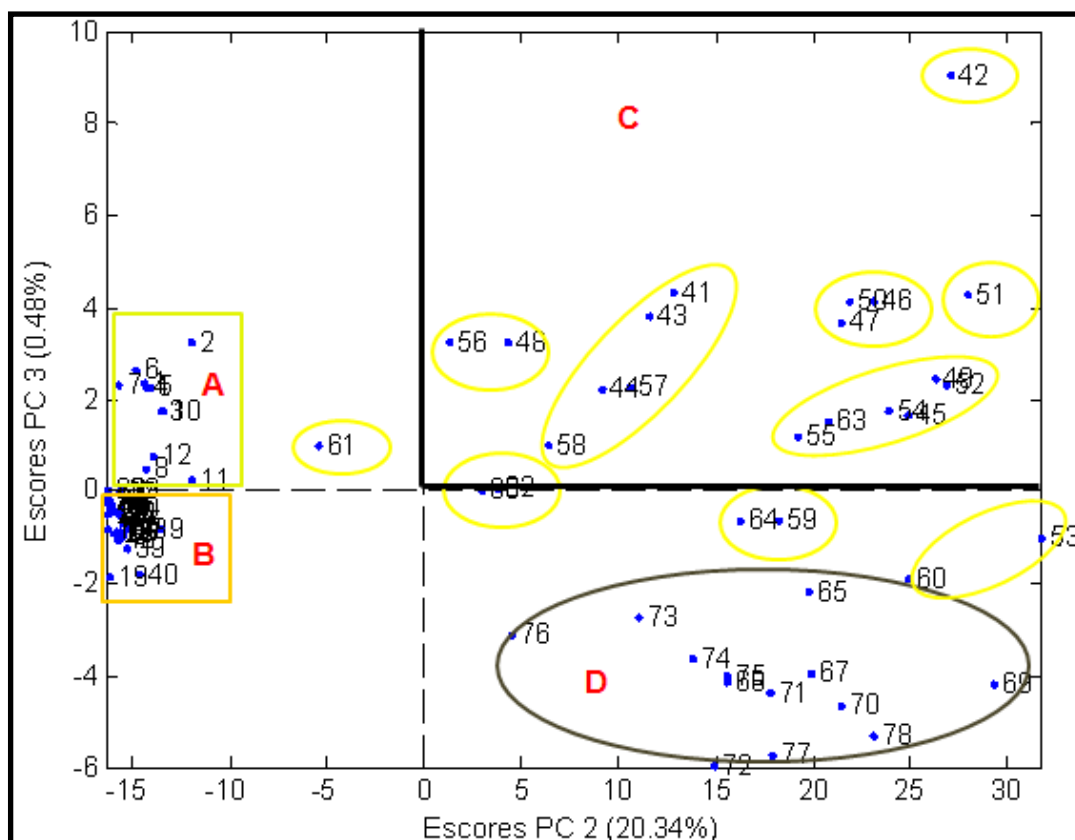


FIGURA 22 – GRÁFICO DE ESCORES DA PC2 x PC3 NA REGIÃO MID

NOTA: Grupos identificados: A. Farinha de linhaça dourada (Empresa A); B. Farinha de linhaça dourada e marrom (Empresa B); C. Semente de linhaça dourada (Empresa A e B); D. Semente de linhaça marrom (Empresa B)

As regiões espectrais do MID responsáveis por apontar as diferenças das amostras, em relação aos teores de lipídeos e ácidos graxos, estão evidenciadas nos gráficos de *loadings* (FIGURAS 23 e 24).

Analisando o gráfico de escores e *loadings* para a PC2, observou-se que em valores positivos de escores encontram-se as amostras de sementes de linhaça das variedades dourada e marrom, as quais apresentaram valores de ácidos graxos muito similares. As regiões do espectro com *loadings* positivos foram caracterizadas por sinais mais expressivos na região 1 correspondente a banda de absorção de O – H (ácido carboxílico) com números de onda entre 3.385 e 2.400 cm^{-1} e, subsequentemente por sinais característicos das ligações C = C, C – H e C – O presentes na região 2 com números de onda entre 2.370 e 890 cm^{-1} . As regiões do espectro com *loadings* negativos que representam as amostras de farinha, estão relacionados à região 3 com bandas de O – H, N – H, cujos sinais estão entre 4.000 e 3.385 cm^{-1} , e a região 4 com ligação C = O, correspondente aos números de onda

entre 2.370 e 1.600 cm^{-1} (FIGURA 23) (LAZZARI; CHIANTORE, 1999; PAVIA *et al.*, 2010).

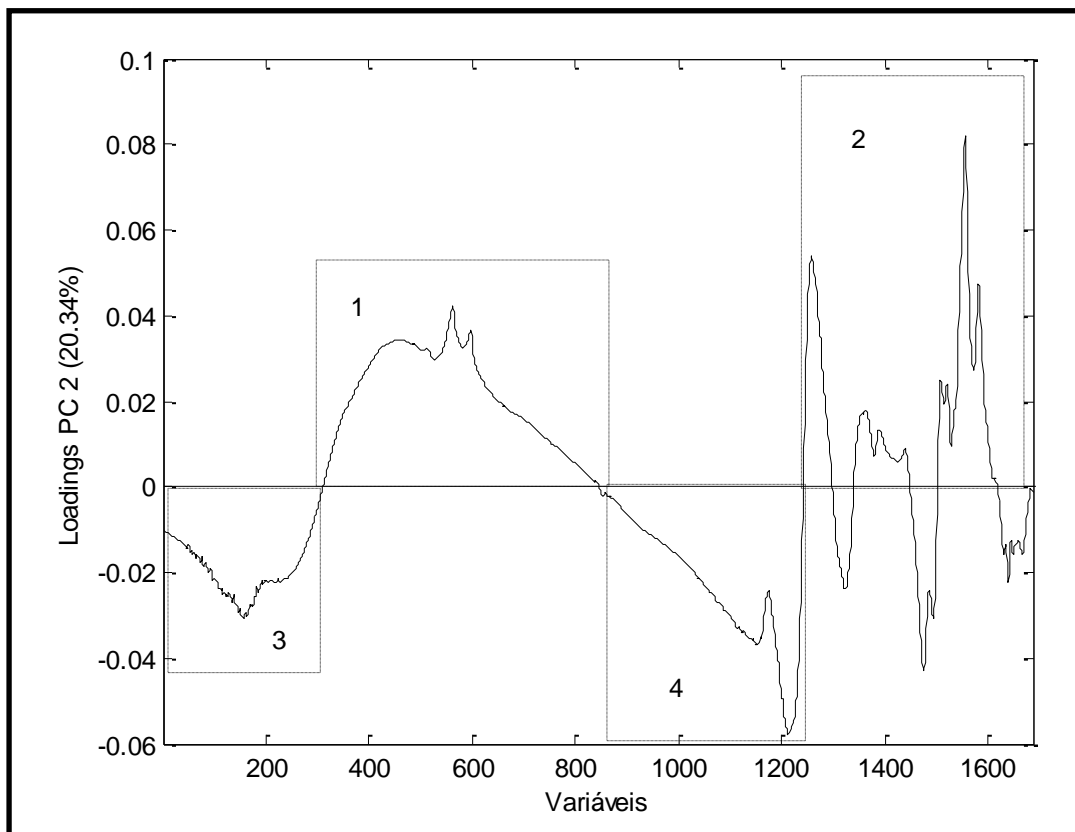


FIGURA 23 – GRÁFICO DE *LOADINGS* PARA A PC2 NA REGIÃO MID

Como mencionado anteriormente, a terceira componente principal possibilitou a separação das amostras quanto a variedade. Um grupo formado pela maior parte das amostras da variedade dourada ficou situado na parte positiva do gráfico de escores, enquanto que outro grupo, representado pela variedade marrom, ficou situado na parte negativa do gráfico de escores. Os sinais característicos destes grupos estão destacados na FIGURA 24. As principais regiões espectrais relacionadas a PC3 correspondem a região 1 com sinais de 1.850 a 1.580 cm^{-1} onde encontraram-se bandas de $\text{C} = \text{O}$ e $\text{C} = \text{C}$, a região 2 com ligações de $\text{C} - \text{H}$ e $\text{C} = \text{C}$ com números de onda entre 1.100 e 810 cm^{-1} , a região 3 com números de onda entre 3.740 a 3.500 cm^{-1} correspondentes a ligações de $\text{O} - \text{H}$ e $\text{N} - \text{H}$, e a região 4 com vários sinais entre 1.580 e 1.100 cm^{-1} relacionados as bandas de $\text{C} - \text{H}$, $\text{C} = \text{C}$

e C – O (LAZZARI; CHIANTORE, 1999; SILVERSTEIN *et al.*, 2007; PAVIA *et al.*, 2010).

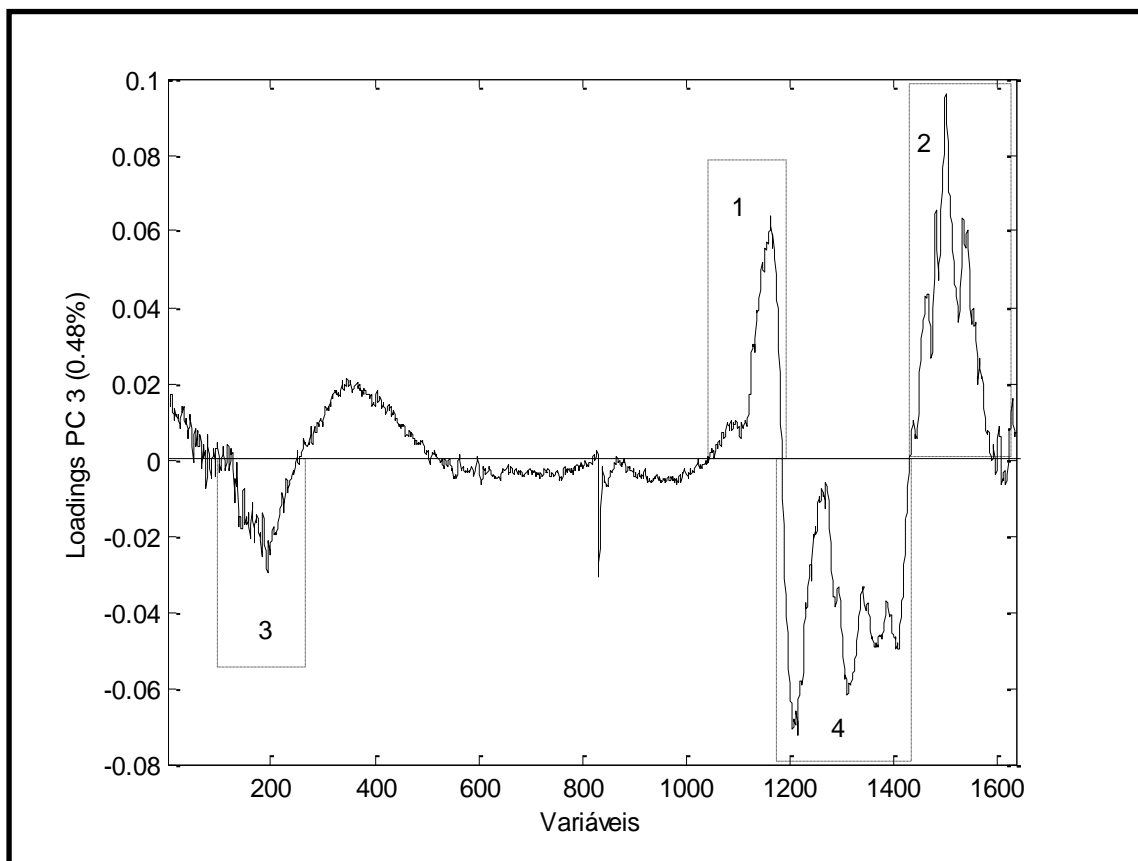


FIGURA 24 – GRÁFICO DE *LOADINGS* PARA A PC3 NA REGIÃO MID

5.2.3 Comparação entre Análise de Componentes Principais (PCA) no NIR e MID

Os espectros obtidos no infravermelho próximo (NIR) apresentaram-se com menos ruídos, quando comparados aos espectros obtidos no infravermelho médio e, quando aplicado o pré-processamento de alisamento, os ruídos foram minimizados. Em consequência, foram utilizadas apenas as regiões espectrais relevantes, com variância explicada de 76,85%. Os pré-processamentos dos espectros obtidos na região do MID foram realizados a fim de minimizar os ruídos e analisar apenas as regiões espectrais mais importantes, o que resultou em variância capturada de apenas 20,82%.

A PCA baseada nos dados espectrais do NIR apresentou melhor separação das amostras de farinhas e sementes de linhaça quando comparada a PCA baseada nos dados espectrais do MID. No MID, houve apenas uma distinção entre farinhas e sementes, e com relação aos espectros das farinhas de linhaça dourada e marrom, ambos tiveram um comportamento muito similar, indicando que todas seriam iguais, em relação ao teor de ácidos graxos.

As ligações orgânicas com sinais mais evidentes foram C – H, O – H, C = O, as quais são características de ácidos carboxílicos. Estas ligações foram encontradas tanto nos espectros NIR quanto nos espectros MID. Também foi possível observar diferenças entre as farinhas e sementes de linhaça quanto ao teor de lipídeos e ácidos graxos, em ambas as regiões, NIR e MID. Com base nestas informações, foi decidido desenvolver modelos de calibração separadamente para farinhas e sementes de linhaça.

5.3 DESENVOLVIMENTO DE MODELOS DE CALIBRAÇÃO MULTIVARIADA PARA PREVISÃO DO TEOR DE ÁCIDOS GRAXOS EM FARINHAS DE LINHAÇA

Os espectros no infravermelho do conjunto de calibração para farinhas de linhaça estão apresentados na FIGURA 25. A faixa espectral utilizada para o infravermelho próximo (NIR) foi de 9.000 a 4.000 cm^{-1} (2.594 pontos), enquanto que para o infravermelho médio (MID) foi de 4.000 a 750 cm^{-1} (1.686 pontos). Para o desenvolvimento dos modelos de calibração foram utilizadas 40 amostras de farinhas de linhaça, dentre as quais 34 amostras (85%) foram aplicadas na etapa de calibração e 6 amostras (15%) para validação externa, as quais foram selecionadas com base na PCA.

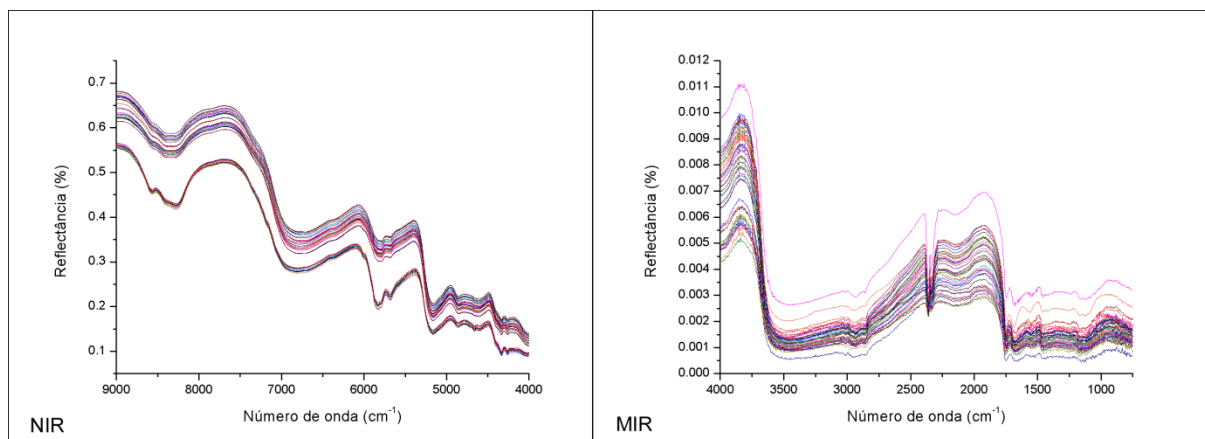


FIGURA 25 – ESPECTROS NO INFRAVERMELHO NA REGIÃO NIR E MID PARA FARINHAS DE LINHAÇA

As amostras de farinhas de linhaça apresentam em sua constituição uma mistura complexa de componentes. Assim, existem em seus espectros informações de diversos grupos funcionais, cujos sinais analíticos podem estar sobrepostos, dificultando à atribuição de uma determinada banda a uma única função química. Desta forma, procurou-se obter modelos otimizados para cada ácido graxo (ácidos oleico, linoleico e linolênico) e, em seguida verificar se os números de onda considerados relevantes para a construção do modelo estavam relacionados às atribuições de absorção de radiação das ligações químicas e grupos funcionais dos respectivos componentes estudados.

O desenvolvimento dos modelos multivariados foi realizado com base na metodologia de Regressão de Mínimos Quadrados Parciais (PLSR), que possibilita otimizar um modelo para cada componente de interesse. A aplicação desta técnica requer maior investimento de tempo para a elaboração dos modelos. Contudo, os modelos obtidos tendem a apresentar uma melhor capacidade de previsão. O algoritmo empregado foi o NIPLS (*Nonlinear Iterative Partial Least Squares*), por ser mais robusto.

Com o intuito de minimizar os efeitos causados por variações aleatórias ou sistemáticas indesejáveis, foram empregados como pré-processamentos aos dados espectrais a correção do espalhamento da luz (MSC), normalização, SNV, detrend, alisamento e primeira derivada antes da construção dos modelos.

5.3.1 Modelos de calibração multivariada para previsão do teor de ácido oleico para farinha de linhaça

Durante a etapa de calibração, o conjunto de 34 espectros no infravermelho das amostras de farinha de linhaça dourada e marrom foi correlacionado com o respectivo teor de ácido oleico presente nas amostras. Vários modelos PLSR foram construídos a fim de determinar a região espectral (NIR ou MID) produziria o menor erro de previsão (Erro padrão relativo das amostras de validação externa), RMSEC (Raiz quadrada da soma dos erros de calibração) e RMSECV (Raiz quadrada da soma dos erros de validação cruzada), em função do menor número de Variáveis Latentes (VLs) obtidas pelo método de validação interna (*leave one out*). Além disso, avaliaram-se os coeficientes de determinação dos conjuntos de calibração (R^2_{cal}) e validação externa (R^2_{val}) e o gráfico do coeficiente de regressão. A Tabela 7 mostra um resumo dos melhores resultados obtidos pelos modelos multivariados, na região do NIR e MID.

TABELA 7 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO OLEICO NO NIR E MID EM FARINHAS DE LINHAÇA

Região	Modelo	Pré-processamento	Nº de VLs	RMSEC	RMSECV	R^2_{cal}	R^2_{val}	Erro Padrão Relativo (%)
NIR	1	MSC/Alisamento/1ª derivada	3	0,38	0,43	0,985	0,981	0,90
NIR	2	MSC/Alisamento/1ª derivada	4	0,29	0,37	0,991	0,986	0,91
NIR	3	Alisamento/1ª derivada	2	0,36	0,38	0,987	0,986	1,08
NIR	4	MSC/Normalização	2	0,31	0,33	0,990	0,989	1,02
MID	5	MSC/Alisamento	4	0,30	0,54	0,991	0,971	1,23
MID	6	MSC/Alisamento/1ª derivada	4	0,28	0,63	0,992	0,965	1,85
MID	7	SNV/Detrend	4	0,29	0,46	0,991	0,979	1,24
MID	8	Alisamento/Normalização	4	0,33	0,53	0,989	0,972	1,38

VLs = Variáveis latentes

RMSEC/ RMSECV= Raiz quadrada da soma dos erros de calibração/validação cruzada

$R^2_{cal/val}$ = Coeficiente de determinação de calibração/validação

Ao avaliar a capacidade de previsão do teor de ácido oleico nas amostras de farinha de linhaça pelos modelos desenvolvidos na região do infravermelho próximo (NIR), observou-se que o modelo 1 apresentou o menor valor de erro padrão relativo

(0,90%), no entanto, considerando como parâmetros o número de variáveis latentes e RMSECV, foi observado que entre os modelos obtidos no NIR, o modelo 4 apresentou o menor valor de RMSECV (0,33) e o menor número de variáveis latentes, apenas 2 variáveis, com maior coeficiente de determinação para validação (R^2_{val}) de 0,989. Assim, o modelo 4 foi considerado o melhor modelo para farinhas de linhaça, para região no NIR.

Com relação aos modelos de regressão para farinhas, para o infravermelho médio (MID), os modelos 5 e 7 apresentaram valores de erro padrão relativo muito próximos, todavia os valores obtidos para RMSEC e RMSECV para o modelo 5 foram maiores dos que os obtidos para o modelo 7. Considerando que ambos os modelos apresentaram o mesmo número de variáveis latentes, optou-se pelo modelo que apresentasse os menores valores de RMSECV e RMSEC e, os maiores de coeficiente de determinação (R^2) para calibração e validação. Assim, o modelo 7 foi considerado como o melhor modelo ao empregar como pré-processamentos SNV e Detrend.

A FIGURA 26 apresenta a evolução dos valores de RMSEC e RMSECV, em função do número de variáveis latentes, para os modelos no NIR (modelo 1) e MID (modelo 7).

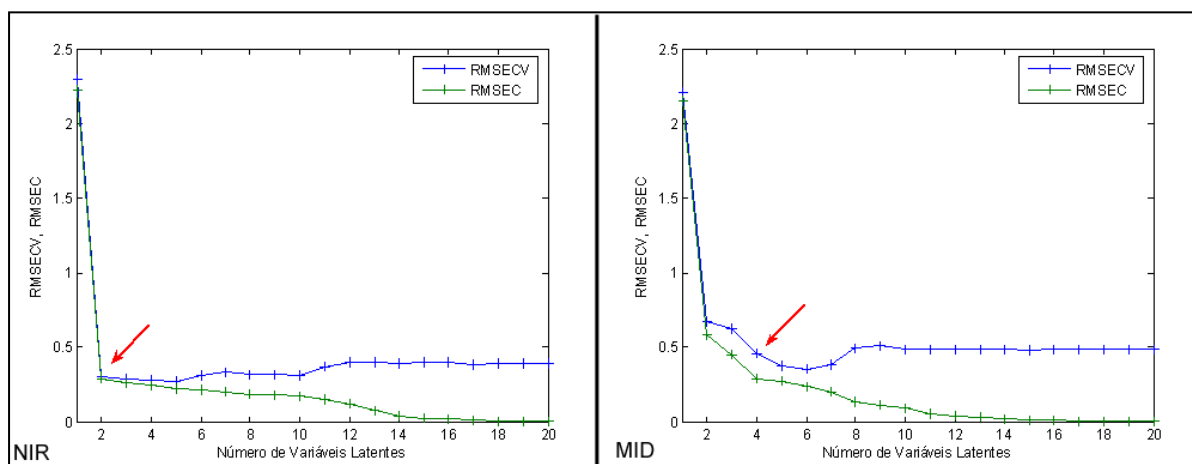


FIGURA 26 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Com base na FIGURA 26, percebeu-se que o menor valor de RMSECV para NIR foi obtido com 2 variáveis latentes. Enquanto que para o MID, o valor de RMSECV diminui a partir de 4 variáveis latentes, sendo que para as variáveis 5 e 6 o

valor de RMSECV não se altera significativamente, por isso optou-se pelo modelo que apresentava o menor número de variáveis latentes, uma vez que o uso de um número elevado de variáveis latentes pode gerar modelos superestimados e comprometer a previsão, devido à agregação de ruídos.

As duas variáveis latentes do modelo 4 no NIR são responsáveis por explicar 100% da variabilidade espectral (matriz X) e 99,98% da variabilidade do teor de ácido oleico (matriz Y). Com relação ao modelo no MID, as 4 variáveis latentes do modelo 7, descrevem 99,93% da variância espectral e 99,98% do teor de ácido oleico.

Para verificar se nos modelos de calibração para a região no NIR e no MID havia amostras anômalas (*outliers*) avaliaram-se os parâmetros *Leverage* e Resíduos de *Student* (FIGURA 27). No modelo para NIR, não houve indícios de amostras anômalas no limite de *leverage* de 0,18, considerando 2 variáveis latentes e Resíduos de *Student* de $\pm 2,5$ (95% de confiança). Enquanto que para o MID, tanto as amostras 02 e 07 apresentaram um valor de *leverage* maior que o limite de 0,35, considerando-se 4 variáveis latentes. Para efeitos de comparação, é possível observar que as replicatas da amostra 02 (amostras 01 e 03) e da amostra 07 (amostras 05 e 06) apresentaram valores de *leverage* dentro do limite aceitável de 0,35. A amostra 17 apresentou elevado valor de Resíduos de *Student*, no entanto ao analisar suas replicatas (amostras 15 e 16), percebeu-se que ambas estavam dentro de limite recomendado de $\pm 2,5$. As amostras 02, 07 e 17 foram retiradas do conjunto de calibração e o modelo foi refeito, no entanto as características do modelo não foram alteradas significativamente, não houve redução no número de variáveis latentes e nos valores de RMSEC, RMSECV e Erro Padrão Relativo. Portanto, as amostras 02, 07 e 17 foram mantidas no modelo.

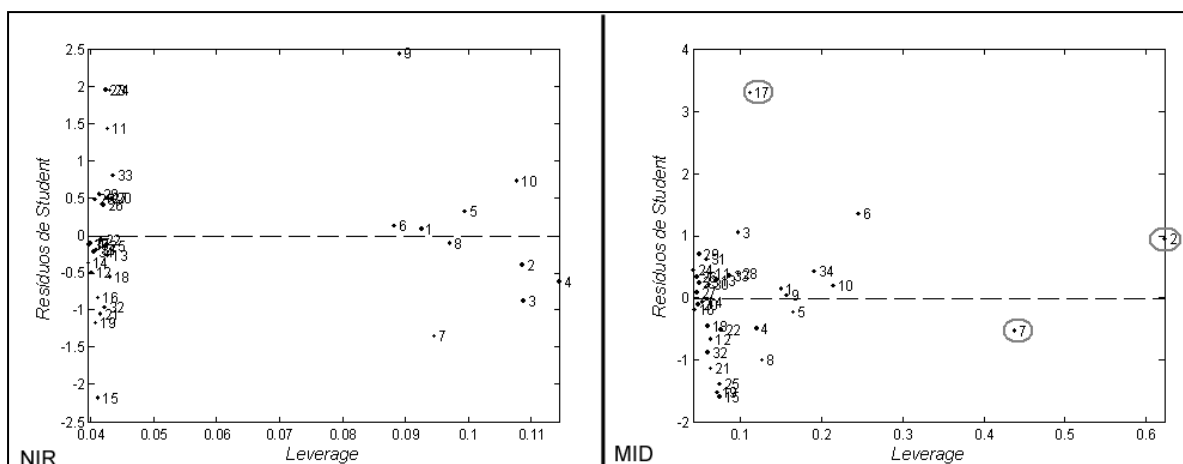


FIGURA 27 – GRÁFICOS DE LEVERAGE x RESÍDUO DE STUDENT OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Tanto os modelos no NIR quanto no MID têm em seus conjuntos de amostras uma elevada variação entre os dados, apresentando um grupo de amostras muito concentrado e outro pouco concentrado, em relação ao teor de ácido oleico. Esta relação pode ser visualizada na FIGURA 28. Entretanto, os modelos apresentaram elevados coeficientes de determinação. Para a região do NIR, os coeficientes de determinação de calibração (R^2_{cal}) e validação (R^2_{val}) foram de 0,990 e 0,989, respectivamente, já para o MID obteve-se 0,991 e 0,979, respectivamente para o R^2_{cal} e R^2_{val} .

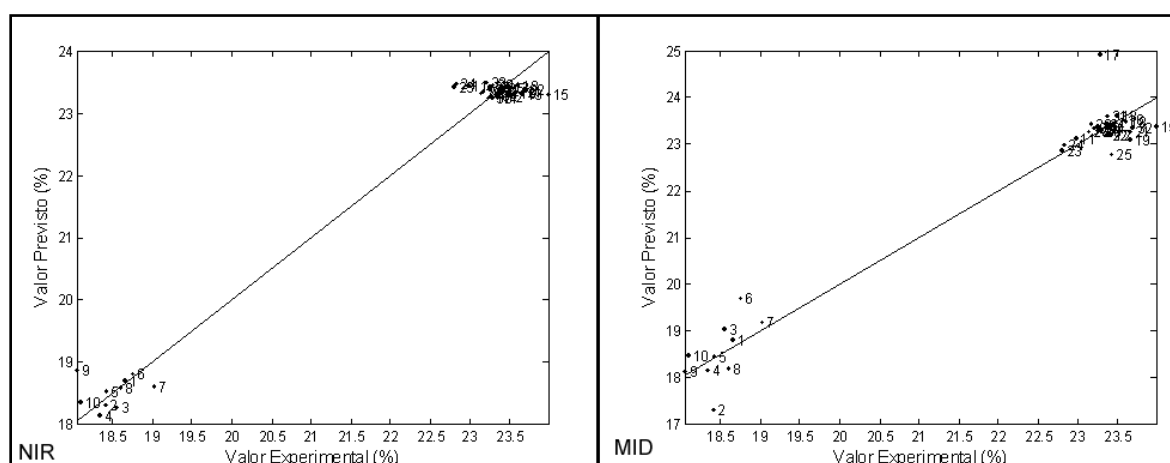


FIGURA 28 – GRÁFICO DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Os gráficos dos coeficientes de regressão evidenciam as principais regiões do espectro que apresentam informações analíticas relevantes para a determinação do ácido oleico (FIGURA 29). Estas regiões incluem para o NIR, os números de onda entre 8.761 a 8.037 cm^{-1} (referente à ligação C – H), 7.362 a 6.976 cm^{-1} (referente às ligações C – H, O – H) e 6.012 a 5.240 cm^{-1} (referente às ligações C – H, C = O, O – H). Para o MID, as principais regiões possuem números de onda entre 4.000 e 3.615 cm^{-1} (referente à ligação O – H), 2.787 e 2.310 cm^{-1} (referente à ligação C – H) e 1.975 a 1.685 cm^{-1} (referente à ligação C = O) (LAZZARI; CHIANTORE, 1999; SILVERSTEIN *et al.*, 2007; CEN; HE, 2007; PAVIA *et al.*, 2010; XIABO *et al.*, 2010).

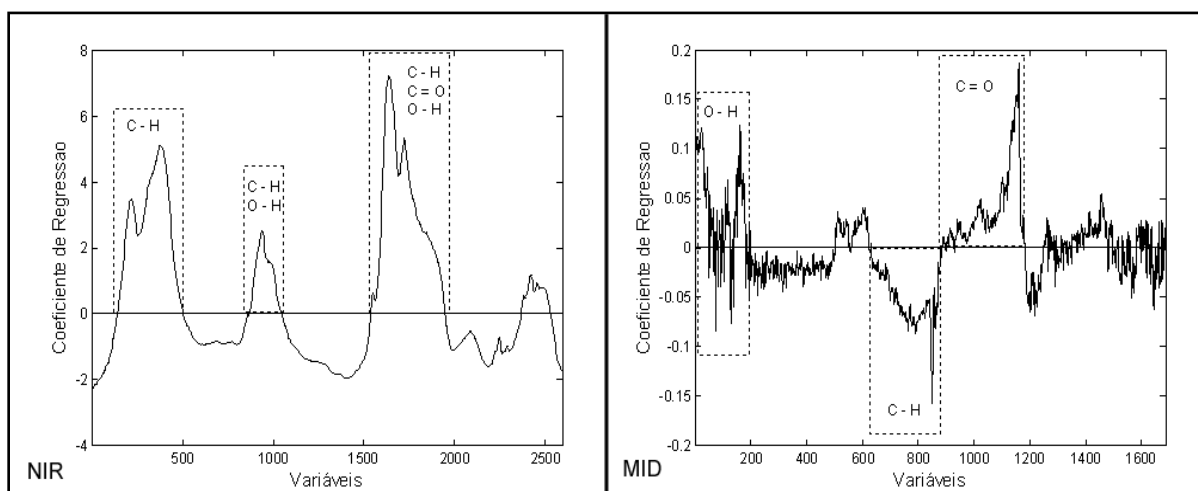


FIGURA 29 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Os valores previstos para o teor de ácido oleico nas seis amostras do conjunto de validação externa estão representados na Tabela 8, de acordo com os melhores modelos obtidos nas regiões do NIR (modelo 4) e MID (modelo 7), respectivamente. Analisando os resultados na Tabela 8 com relação aos valores previstos pelos modelos e com base no erro padrão relativo, o melhor modelo de previsão do teor de ácido oleico foi obtido no NIR.

TABELA 8 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO OLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA FARINHA DE LINHAÇA

Amostras	Valor Experimental (%)	Valores Previstos (%)		Erro Padrão Relativo (%)	
		NIR	MID	NIR	MID
4	18,5	18,11	18,47	2,13	0,16
8	18,17	18,46	19,29	1,60	6,17
21	23,38	23,23	23,42	0,65	0,15
22	23,35	23,32	23,21	0,14	0,60
30	23,13	23,37	23,16	1,05	0,13
36	23,25	23,37	23,19	0,53	0,27
Erro relativo médio (%)				1,02	1,24

5.3.2 Modelos de calibração multivariada para previsão do teor de ácido linoleico para farinha de linhaça

Para a fase de calibração, utilizaram-se 34 espectros no infravermelho das amostras de farinha de linhaça dourada e marrom, os quais foram correlacionados com o respectivo teor de ácido linoleico presente nas amostras. A faixa espectral utilizada para o infravermelho próximo (NIR) foi entre 9.000 e 4.000 cm^{-1} (2.594 pontos), e infravermelho médio (MID), entre 4.000 e 750 cm^{-1} (1.686 pontos). Diversos modelos PLSR foram construídos a fim de verificar a região espectral (NIR ou MID) e os tipos de pré-processamento de dados que produziram os melhores resultados, conforme pode ser visto na Tabela 9.

TABELA 9 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLEICO NO NIR E MID EM FARINHAS DE LINHAÇA

Região	Modelo	Pré-processamento	Nº de VLS	RMSEC	RMSECV	R^2_{cal}	R^2_{val}	Erro Padrão Relativo (%)
NIR	1	MSC/Normalização	4	0,32	0,37	0,987	0,983	1,18
NIR	2	MSC/Alisamento/1ª derivada	4	0,33	0,43	0,986	0,977	1,45
NIR	3	SNV/Detrend	4	0,36	0,41	0,984	0,979	2,36
NIR	4	Alisamento/Normalização	4	0,38	0,43	0,982	0,977	1,71
MID	5	MSC/Alisamento/Normalização	3	0,84	0,95	0,91	0,883	0,76
MID	6	Alisamento/Normalização	4	0,71	0,82	0,936	0,914	2,21
MID	7	SNV/Detrend	4	0,69	0,83	0,939	0,912	1,85
MID	8	SNV/Detrend	6	0,39	0,54	0,982	0,963	2,64

VLS = Variáveis latentes

RMSEC/ RMSECV= Raiz quadrada da soma dos erros de calibração/validação cruzada

$R^2_{\text{cal/val}}$ = Coeficiente de determinação de calibração/validação

Analisando a Tabela 9, pode-se perceber que os melhores modelos desenvolvidos na região do infravermelho próximo (NIR) empregaram 4 variáveis latentes. No entanto, o menor erro padrão relativo (1,18%) foi obtido através do modelo 1, com MSC e Normalização como pré-processamentos. Além disto, o modelo 1 também apresentou os menores valores de RMSECV e RMSEC, quando comparado aos demais modelos no NIR.

Com relação aos modelos de regressão no infravermelho médio (MID), os pré-processamentos MSC, Alisamento e Normalização, geraram melhores resultados, com erro padrão relativo de 0,76%, empregando-se 3 variáveis latentes para explicar a variabilidade dos dados.

A variação dos valores de RMSEC e RMSECV, em função do número de variáveis latentes, na região do NIR e MID pode ser visualizada na FIGURA 30. O menor valor de RMSECV para NIR foi obtido com 4 variáveis latentes. Já para o MID, o valor de RMSECV diminuiu a partir de 3 variáveis latentes, tendo sido o valor de RMSECV mínimo encontrado para 6 variáveis latentes. No entanto, a porcentagem de variância capturada entre as variáveis latentes 3, 4, 5 e 6, manteve-se constante. Então, optou-se pelo modelo 5 que apresentava o menor número de variáveis latentes, pois o uso de um número elevado de variáveis latentes pode conduzir a superestimação durante a previsão do teor de ácido linoleico.

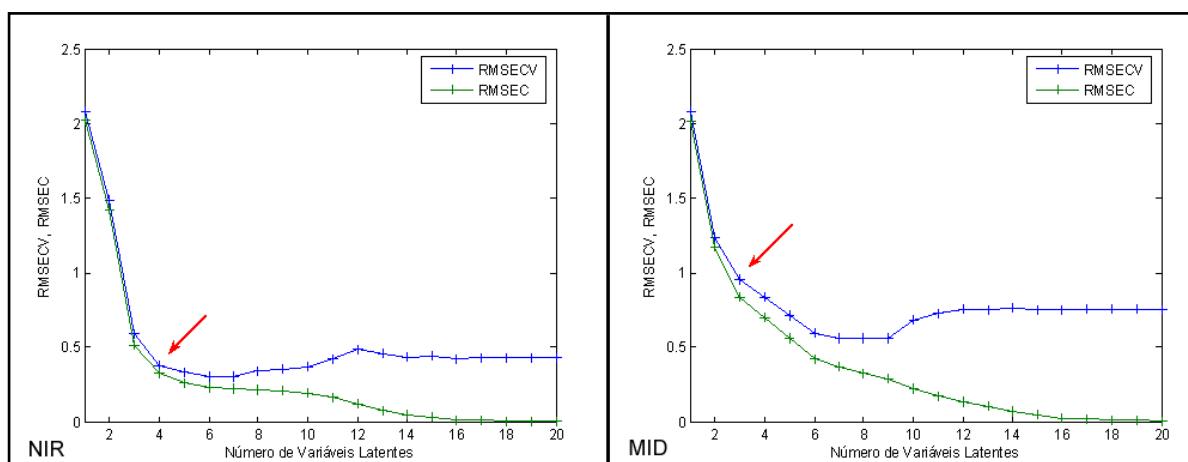


FIGURA 30 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

O modelo 1 obtido no NIR, com 4 variáveis latentes, descreveu 100% da variabilidade dos dados espectrais (matriz X) e 99,98% da variabilidade do teor de

ácido linoleico (matriz Y). Já o modelo 5 no MID, utilizando 3 variáveis latentes, explica 99,97% da variância espectral e 99,83% do teor de ácido linoleico.

No modelo 1 para NIR, não houve indícios de amostras anômalas no limite de *leverage* de 0,35 e, para Resíduos de *Student* de $\pm 2,5$ (95% de confiança) (FIGURA 31). Já para o MID, apenas a amostra 02 apresentou um valor de *leverage* maior que o recomendado, de 0,26. Para efeitos de comparação, foi possível observar que as replicatas desta amostra (amostras 01 e 03) apresentaram valores de *leverage* dentro do limite de 0,26. Quando se retirou a amostra 02 do conjunto de calibração não houve alteração significativa nas características do modelo (não houve redução no número de variáveis latentes e nos valores de RMSEC, RMSECV e Erro Padrão Relativo), por isto a amostra 02 foi mantida no modelo.

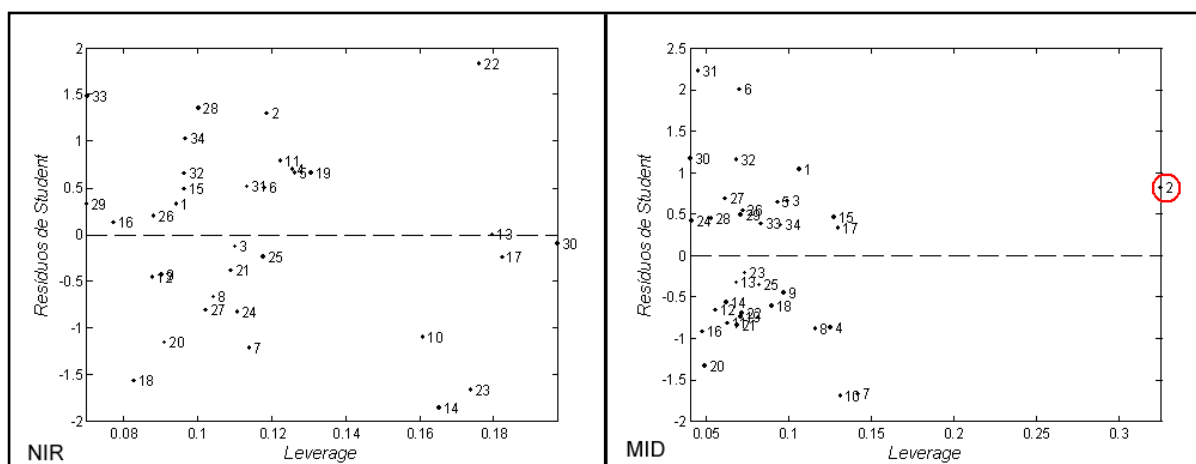


FIGURA 31 – GRÁFICOS DE *LEVERAGE* x RESÍDUO DE *STUDENT* OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Os conjuntos de amostras, em relação ao teor de ácido linoleico, apresentaram variação entre os dados. Pode-se observar que há amostras mais concentradas e menos concentradas (FIGURA 32). O modelo para a região do NIR apresentou maiores coeficientes de determinação quando comparado ao modelo desenvolvido para a região do MID. Para o NIR, os coeficientes de determinação de calibração (R^2_{cal}) e validação (R^2_{val}) foram 0,985 e 0,983, respectivamente. Já para o MID, obteve-se os valores de 0,950 e 0,883, respectivamente para R^2_{cal} e R^2_{val} .

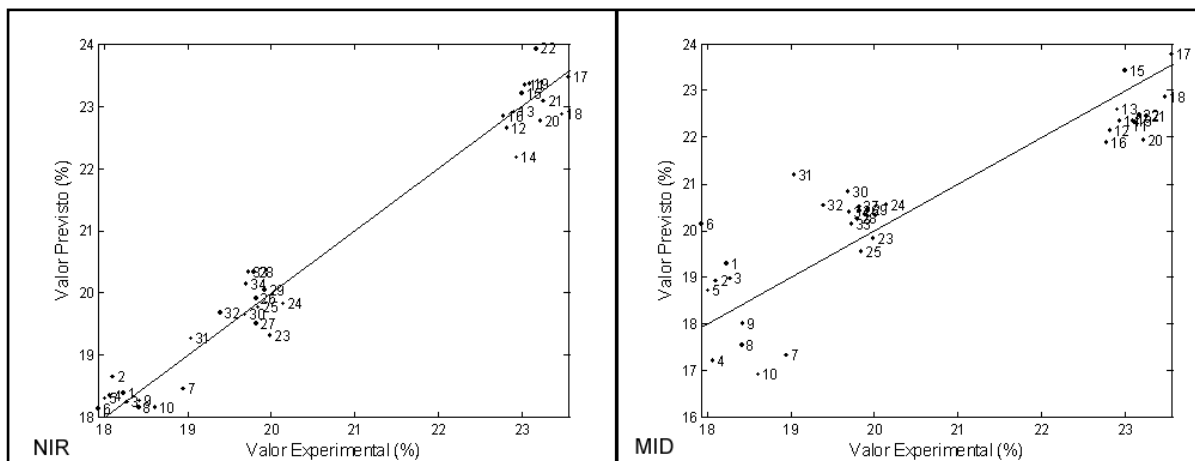


FIGURA 32 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Nos gráficos dos coeficientes de regressão foram evidenciadas as regiões com informações analíticas relevantes para a determinação de ácido linoleico. Entre as principais regiões do NIR estão às correspondentes aos números de onda entre 8.040 e 7.073 cm^{-1} (referente às ligações $\text{C} = \text{O}$, $\text{C} - \text{H}$), 7.073 e 6.568 cm^{-1} (referente às ligações $\text{C} - \text{H}$, $\text{C} = \text{O}$, $\text{N} - \text{H}$) e 6.016 a 4.961 cm^{-1} (referente às ligações $\text{C} - \text{H}$, $\text{C} = \text{O}$, $\text{O} - \text{H}$). Para o MID, as principais regiões possuem números de onda entre 3.726 e 3.560 cm^{-1} (referente à ligação $\text{O} - \text{H}$), 2.404 e 2.295 cm^{-1} (referente às ligações $\text{C} = \text{N}$, $\text{O} - \text{H}$), 2.294 e 1.688 cm^{-1} (referente às ligações $\text{C} = \text{C}$, $\text{C} = \text{O}$) e 1.110 a 918 cm^{-1} (referente às ligações $\text{C} - \text{O}$ e $\text{C} - \text{H}$) (LAZZARI; CHIANTORE, 1999; SILVERSTEIN *et al.*, 2007; CEN; HE, 2007; PAVIA *et al.*, 2010; XIABO *et al.*, 2010).

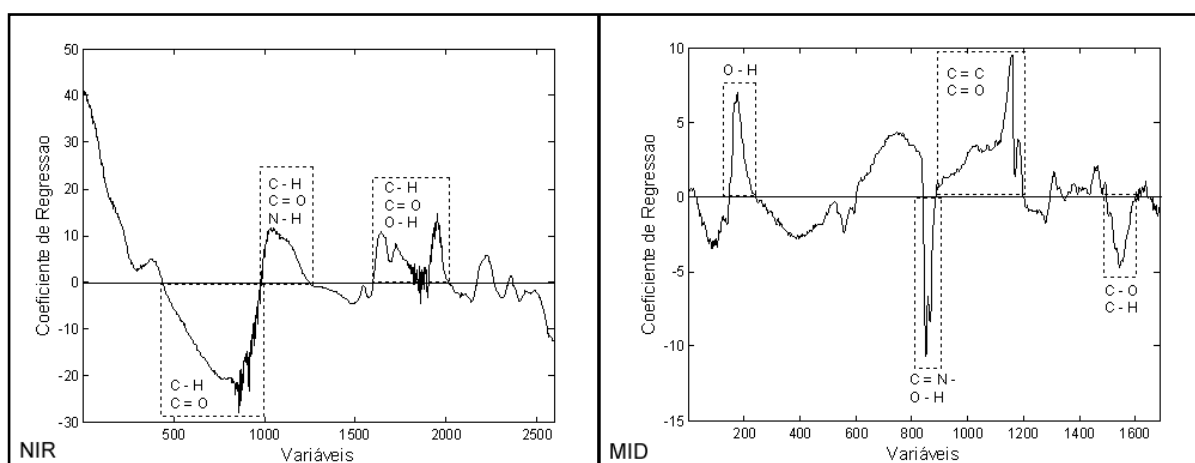


FIGURA 33 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

A capacidade preditiva do teor de ácido linoleico das amostras do conjunto de validação externa está representada na Tabela 10, com os valores experimental e previsto pelos modelos no infravermelho próximo e médio e seus respectivos valores de erro padrão relativo. Com base nos resultados obtidos na Tabela 10, verificou-se que o menor erro padrão relativo foi obtido pelo modelo no MID e com aplicação do menor número de variáveis latentes, somente 3 variáveis. Assim, concluiu-se que o melhor modelo de calibração para a previsão do teor de ácido linoleico foi obtido no MID.

TABELA 10 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA FARINHA DE LINHAÇA

Amostras	Valor Experimental (%)	Valores Previstos (%)		Erro Padrão Relativo (%)	
		NIR	MID	NIR	MID
4	18,15	18,05	18,01	0,56	0,78
8	18,24	18,30	17,87	0,34	2,02
21	23,12	23,18	23,17	0,28	0,22
22	22,32	23,19	22,47	3,88	0,66
30	19,72	19,62	19,88	0,49	0,80
36	19,82	19,51	19,83	1,56	0,05
Erro relativo médio (%)				1,18	0,76

5.3.3 Modelos de calibração multivariada para previsão do teor de ácido linolênico para farinha de linhaça

Os resultados dos modelos de calibração PLSR para determinação do teor de ácido linolênico, com base na faixa espectral para NIR e MID, estão apresentados na Tabela 11.

A melhor capacidade de previsão na região do NIR para farinhas de linhaça foi apresentada pelo o modelo 2 que empregou duas variáveis latentes, com erro padrão relativo de 1,21% e pré-processamento MSC e Normalização. Para a região do MID, entre os modelos gerados, o modelo 5 apresentou o menor erro padrão relativo (2,00%). No entanto, o modelo 8, com pré-processamento SNV e Detrend, apresentou menores valores de RMSEC e RMSECV. Assim, assumiu-se o modelo 8 como o melhor modelo para previsão do teor de ácido linolênico para a região MID.

TABELA 11 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLÊNICO NO NIR E MID EM FARINHA DE LINHAÇA

Região	Modelo	Pré-processamento	Nº de VLS	RMSEC	RMSECV	R ² _{cal}	R ² _{val}	Erro Padrão Relativo (%)
NIR	1	MSC/Alisamento/1ª derivada	3	0,54	0,61	0,994	0,993	1,56
NIR	2	MSC/Normalização	2	0,59	0,64	0,993	0,992	1,21
NIR	3	SNV/Detrend	2	0,67	0,71	0,991	0,990	1,43
NIR	4	Alisamento/ Normalização	3	0,59	0,66	0,993	0,991	2,38
MID	5	MSC/Alisamento/ Normalização	4	0,85	1,44	0,986	0,959	2,00
MID	6	Alisamento/ Normalização	3	1,35	1,70	0,963	0,941	3,29
MID	7	Alisamento/ Normalização	5	0,82	1,48	0,987	0,956	2,29
MID	8	SNV/Detrend	4	0,77	1,26	0,988	0,968	2,28

VLS = Variáveis latentes

RMSEC/ RMSECV= Raiz quadrada da soma dos erros de calibração/validação cruzada

R²_{cal/val} = Coeficiente de determinação de calibração/validação

Os valores de RMSEC e RMSECV em função do número de variáveis latentes na região do NIR e MID podem ser visualizados na FIGURA 34. O menor valor de RMSECV para NIR foi obtido com o uso de 2 variáveis latentes, enquanto que para o MID, o menor valor de RMSECV foi encontrado com o uso de 5 variáveis latentes, no entanto entre 4 e 5 variáveis latentes não houve mudança significativa nos valores de RMSECV, portanto optou-se pelo menor número de variáveis latentes (4).

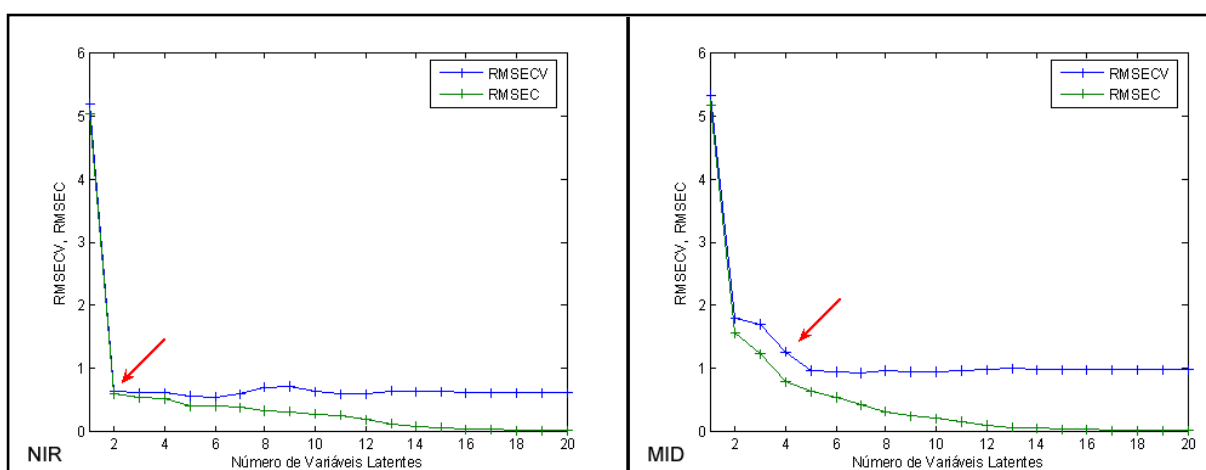


FIGURA 34 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Com 2 variáveis latentes, as porcentagens de variância capturada pelo modelo no NIR foram de 100% para os dados espectrais e 99,98% para os dados do teor de ácido linolênico. No MID, para 4 variáveis latentes estas porcentagens foram 99,93% para variância espectral, e 99,97% para a matriz de teor de ácido linolênico.

Não houve indícios de amostras anômalas no limite de *leverage* de 0,18 e Resíduos de *Student* de $\pm 2,5$ (95% de confiança) para o modelo no NIR. Entretanto, para o modelo no MID, duas amostras (02 e 07) apresentaram comportamento anômalo, com um valor de *leverage* maior que o recomendado (0,35). No entanto, quando ambas as amostras foram retiradas e o modelo refeito, observou-se que novas amostras tornaram-se *outliers* e não houve redução no número de variáveis latentes e nos valores de RMSEC, RMSECV e Erro Padrão Relativo. Analisando o comportamento das replicatas tanto para a amostra 02 (amostras 01 e 03), quanto para a amostra 07 (amostras 05 e 06), observou-se que todas aquelas amostras apresentaram valores de *leverage* dentro do limite. Assim, as amostras 02 e 07 foram mantidas no modelo (FIGURA 35).

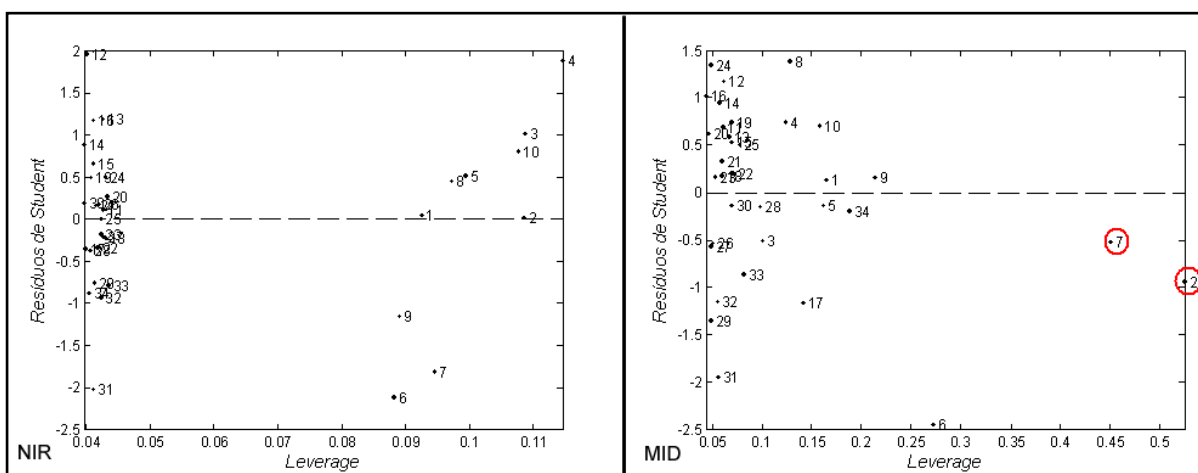


FIGURA 35 – GRÁFICOS DE *LEVERAGE* x RESÍDUO DE *STUDENT* OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

Os modelos no NIR e no MID apresentaram em seus conjuntos de amostras uma elevada variação entre os dados. Na FIGURA 36, é evidente a diferença de teor entre as amostras de 1 a 10, e as demais. As amostras de 1 a 10 apresentaram os maiores teores, com média de 54,64% para o ácido linolênico, contra média de 43,56% das outras amostras. Mesmo apresentando diferenças significativas para os teores de ácido linolênico, os modelos apresentaram bons coeficientes de

determinação para a região do NIR, com R^2_{cal} e R^2_{val} de 0,993 e 0,992, respectivamente. Para o MID, R^2_{cal} e R^2_{val} foram iguais a 0,988 e 0,968, respectivamente.

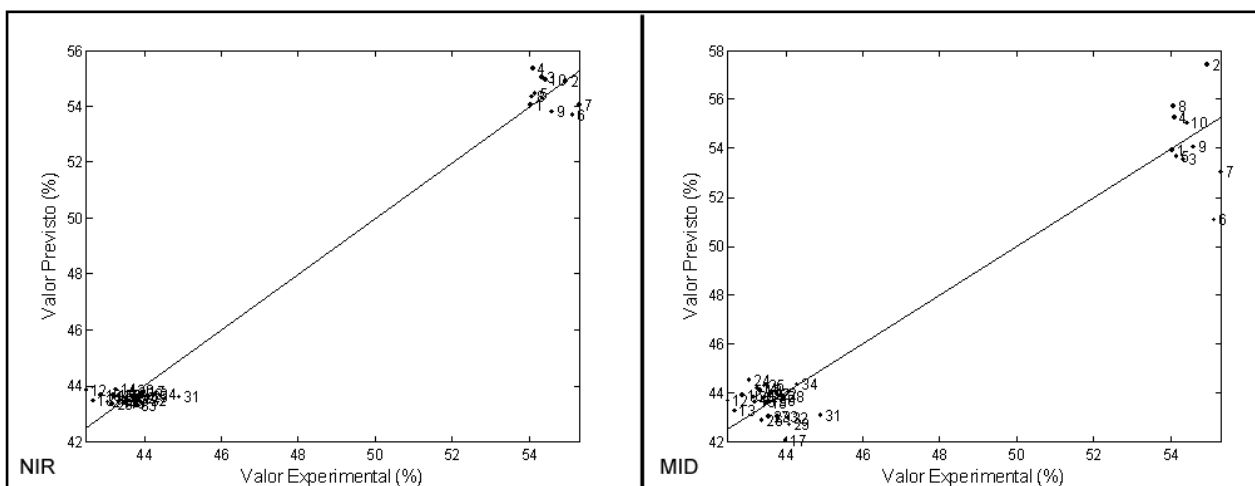


FIGURA 36 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHAS DE LINHAÇA

Os gráficos dos coeficientes de regressão (FIGURA 37) evidenciaram as principais regiões do espectro que apresentaram informações analíticas relevantes para a determinação do ácido linolênico. Estas regiões incluem para o NIR, números de onda entre 8.640 e 8.128 cm^{-1} (referente à ligação C – H) e 6.016 a 5.237 cm^{-1} (referente às ligações C – H, C = O, O – H). Para o infravermelho médio (MID), as principais regiões possuem números de onda entre 2.749 e 2.330 cm^{-1} (referente à ligação O – H) e 1.880 a 1.687 cm^{-1} (referente à ligação C = O) (LAZZARI; CHIANTORE, 1999; SILVERSTEIN *et al.*, 2007; CEN; HE, 2007; PAVIA *et al.*, 2010; XIABO *et al.*, 2010).

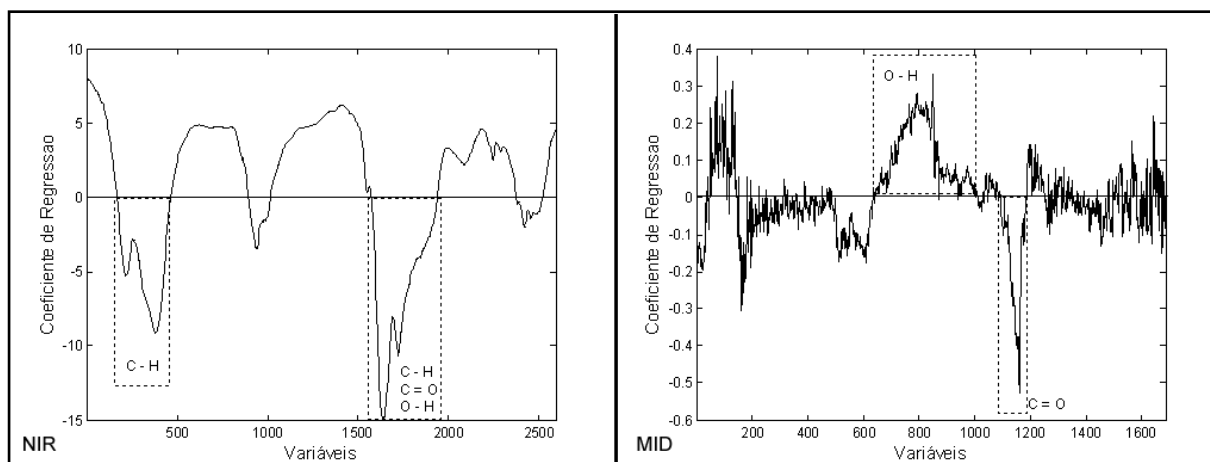


FIGURA 37 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA FARINHA DE LINHAÇA

A capacidade preditiva do teor de ácido linolênico das amostras do conjunto de validação externa está representada na Tabela 12, com os valores experimental e previsto pelos modelos no infravermelho próximo e médio e seus respectivos valores de erro padrão relativo. Com base nos resultados obtidos na Tabela 12, verificou-se que o menor erro padrão relativo foi obtido pelo modelo no NIR e com aplicação do menor número de variáveis latentes, somente 2 variáveis. Assim, concluiu-se que o melhor modelo de calibração para a previsão do teor de ácido linolênico foi o modelo 2, obtido no NIR.

TABELA 12 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLÊNICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA FARINHA DE LINHAÇA

Amostras	Valor Experimental (%)	Valores Previstos (%)		Erro Padrão Relativo (%)	
		NIR	MID	NIR	MID
4	54,1	53,98	53,50	0,21	1,12
8	55,62	54,53	51,83	1,96	6,82
21	44,18	43,36	43,97	1,86	0,48
22	44,31	43,73	45,09	1,31	1,76
30	42,81	43,49	43,70	1,59	2,07
36	43,42	43,57	42,80	0,34	1,42
Erro relativo médio (%)				1,21	2,28

5.4 MODELOS DE CALIBRAÇÃO MULTIVARIADA PARA SEMENTES DE LINHAÇA

Os espectros no infravermelho do conjunto de calibração para sementes de linhaça na região NIR e MID estão apresentados na FIGURA 38. A faixa espectral utilizada para o infravermelho próximo (NIR) foi entre 9.000 e 4.000 cm^{-1} (2.594 pontos) e infravermelho médio (MID), entre 4.000 e 750 cm^{-1} (1.634 pontos), com o objetivo de determinar qual região apresenta melhor capacidade de previsão para o teor de cada ácido graxo de interesse nas sementes de linhaça. Foram utilizadas 38 amostras de sementes, sendo 33 amostras (85%) aplicadas na etapa de calibração e 5 amostras (15%) usadas para validação externa, as quais foram selecionadas com base na PCA.

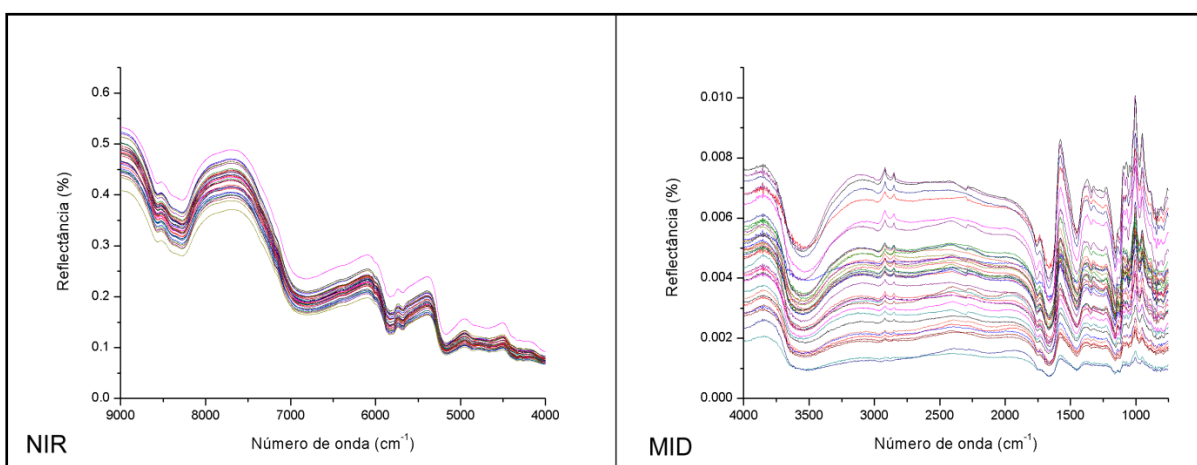


FIGURA 38 – ESPECTROS NO INFRAVERMELHO NA REGIÃO NIR E MID PARA SEMENTES DE LINHAÇA

Assim como as farinhas de linhaça as sementes têm em sua constituição uma mistura complexa de componentes, havendo em seus espectros de infravermelho informações a respeito de diversos grupos funcionais. Desta forma, procurou-se obter modelos otimizados para cada ácido graxo (ácido oleico, linoleico e linolênico), em função da absorção de ligações químicas e grupos funcionais dos respectivos componentes estudados.

O desenvolvimento dos modelos multivariados foi baseado na metodologia de Regressão de Mínimos Quadrados Parciais (PLSR), empregando o algoritmo NIPLS, por ser mais robusto.

Os pré-processamentos, correção do espalhamento da luz (MSC), normalização, SNV, detrend, alisamento e primeira derivada foram aplicados aos dados espectrais a fim de minimizar os efeitos causados por variações aleatórias ou sistemáticas indesejáveis, antes da construção dos modelos, para que estas variações não influenciassem os resultados finais.

Os parâmetros avaliados para indicar o melhor modelo foram os menores erros de previsão (Erro padrão relativo das amostras de validação externa), RMSEC (Raiz quadrada da soma dos erros de calibração) e RMSECV (Raiz quadrada da soma dos erros de validação cruzada), em função do menor número de Variáveis Latentes (VLs) obtidas pelo método de validação interna (*leave one out*). Além destes parâmetros, também foram avaliados os coeficientes de determinação dos conjuntos de calibração (R^2_{cal}) e validação (R^2_{val}), e o coeficiente de regressão, obtidos nas regiões do NIR e MID.

5.4.1.1 Modelos de calibração multivariada para previsão do teor de ácido oleico para semente de linhaça

Durante a etapa de calibração, o conjunto de 33 espectros no infravermelho das amostras de sementes de linhaça dourada e marrom foi correlacionado com o respectivo teor de ácido oleico presente nas amostras. Desenvolveram-se modelos PLSR a fim de verificar a região espectral (NIR ou MID) e os tipos de pré-processamento de dados que produziram os melhores resultados (Tabela 13).

TABELA 13 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO OLEICO NO NIR E MID EM SEMENTES DE LINHAÇA

Região	Modelo	Pré-processamento	Nº de VLs	RMSEC	RMSECV	R^2_{cal}	R^2_{val}	Erro Padrão Relativo (%)
NIR	1	MSC/Alisamento/1ª derivada	4	0,28	0,36	0,878	0,788	0,77
NIR	2	MSC/Normalização	5	0,27	0,35	0,884	0,80	0,79
NIR	3	Normalização	4	0,28	0,34	0,871	0,81	0,80
NIR	4	SNV/Detrend	4	0,31	0,36	0,842	0,792	0,96
MID	5	MSC/Normalização	5	0,32	0,51	0,830	0,526	2,14
MID	6	MSC/Normalização	8	0,16	0,48	0,961	0,596	1,74
MID	7	Alisamento/Normalização	4	0,42	0,52	0,681	0,469	1,02
MID	8	Normalização	4	0,41	0,52	0,696	0,474	1,05

VLs = Variáveis latentes

RMSEC/ RMSECV= Raiz quadrada da soma dos erros de calibração/validação cruzada

$R^2_{cal/val}$ = Coeficiente de determinação de calibração/validação

A melhor capacidade preditiva para o teor de ácido oleico nas amostras de semente de linhaça pelos modelos desenvolvidos na região do NIR foi obtida pelo modelo 3 que apresentou menores valores de RMSECV (0,34) e número de variáveis latentes (4).

Com relação aos modelos de regressão no MID, o menor valor de RMSECV foi encontrado para o modelo 6. Contudo, este modelo utiliza 8 variáveis latentes, o que pode tornar o modelo superestimado e comprometer a previsão da concentração do teor de ácido oleico devido a inclusão de ruídos. Considerando, o menor erro padrão relativo e o menor número de variáveis latentes, o modelo 7, empregando-se como pré-processamentos Alisamento e Normalização, se mostrou como o modelo com a melhor capacidade preditiva.

Os valores de RMSEC e RMSECV, em função do número de variáveis latentes, para os modelos no NIR e MID podem ser visualizados na FIGURA 39. O menor valor de RMSECV para NIR foi obtido com 4 variáveis latentes, que são responsáveis por explicar 100% da variabilidade espectral e 99,98% da variabilidade no teor de ácido oleico. Já para o MID, optou-se pelo modelo 7, que apresentava 4 variáveis latentes, as quais explicaram 99,99% da variância espectral e 99,96% do teor de ácido oleico.

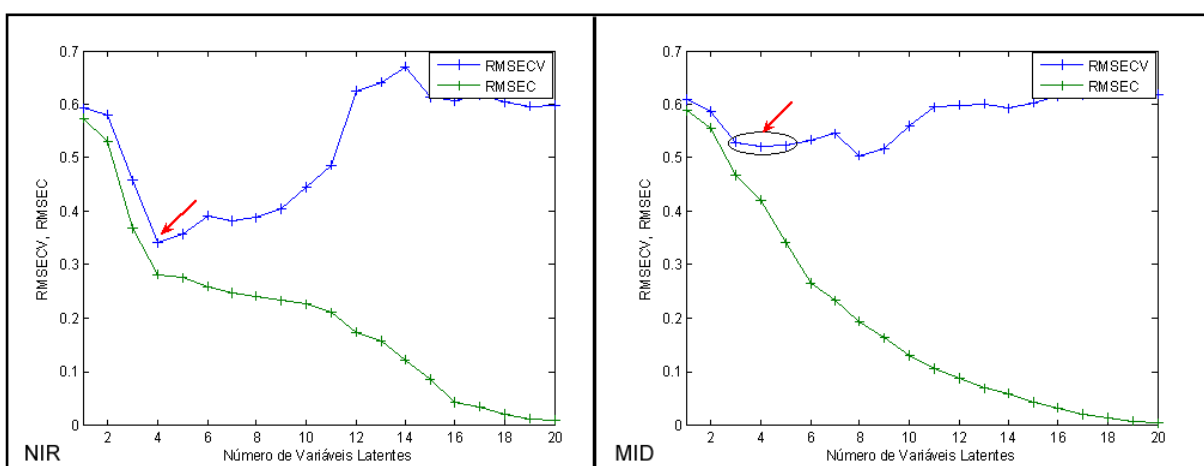


FIGURA 39 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Para avaliar se haviam amostras anômalas (*outliers*) nos modelos de calibração no NIR e MID avaliaram-se os parâmetros *Leverage* e Resíduos de *Student* (FIGURA 40). No modelo para NIR houve duas amostras anômalas

(amostras 17 e 29), no limite de *leverage* de 0,36. Todavia, a classificação destas duas amostras como anômalas pode ter sido causada por problemas na aquisição dos espectros, pois as replicatas das amostras 17 e 29 apresentaram-se dentro do limite. Para Resíduos de *Student* $\pm 2,5$ (95% de confiança) houve apenas uma amostra anômala (amostra 06), a amostra mais concentrada no conjunto de sementes de linhaça dourada.

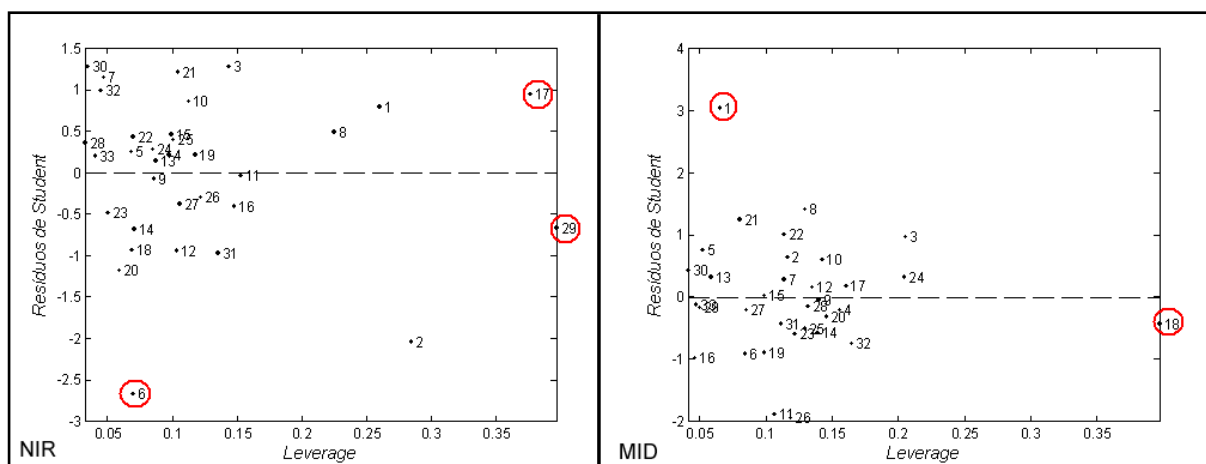


FIGURA 40 – GRÁFICOS DE *LEVERAGE* x RESÍDUO DE *STUDENT* OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

No modelo para o MID, duas amostras apresentaram comportamento anômalo, as amostras 01 e 18. A amostra 01 apresentou maior valor de Resíduos de *Student*, valor este atribuído ao fato da amostra apresentar baixa concentração de ácido oleico. Embora a amostra 18 tenha apresentado um valor de *leverage* maior que o recomendado, de 0,36, a mesma não apresenta nenhuma particularidade, assim, sua classificação como anômala foi atribuída a problemas relacionados a aquisição dos espectros. Deste modo, estas amostras permaneceram no conjunto de calibração, pois quando retiradas não alteraram significativamente as características do modelo.

Os valores determinados pelo método de referência e aqueles previstos pelos modelos multivariados resultaram em coeficientes de determinação maiores para o NIR do que para o MID, sendo que R^2_{cal} e R^2_{val} foram de 0,87 e 0,81, respectivamente, para a região do NIR, e de 0,681 e 0,469, respectivamente, para a região MID (TABELA 14).

Ao analisar a FIGURA 41, percebeu-se que o gráfico para MID apresentou elevada variação entre os valores previstos e os valores experimentais. Estas variações podem estar relacionadas a problemas na aquisição dos espectros.

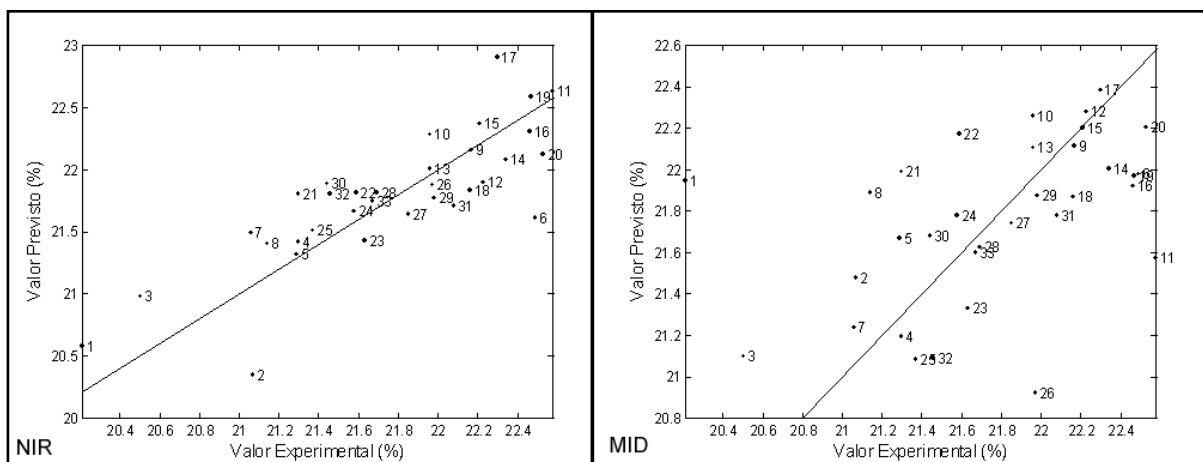


FIGURA 41 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Os gráficos dos coeficientes de regressão (FIGURA 42) evidenciaram as principais regiões do espectro que apresentaram informações analíticas relevantes para a determinação do ácido oleico. Estas regiões incluem para o NIR, números de onda entre 9.000 e 8.037 cm^{-1} (referente à ligação C – H), 8.037 e 5.933 cm^{-1} (referente às ligações C – H, C = O), 5.933 e 5.788 cm^{-1} (referente às ligações C – H) e 5.407 a 4.442 cm^{-1} (referente às ligações C – H, C = O, O – H). Para o MID, as principais regiões possuem números de onda entre 1.877 e 1.221 cm^{-1} (referente às ligações C = C, C = O, C – O) (LAZZARI; CHIANTORE, 1999; SILVERSTEIN *et al.*, 2007; CEN; HE, 2007; PAVIA *et al.*, 2010; XIABO *et al.*, 2010).

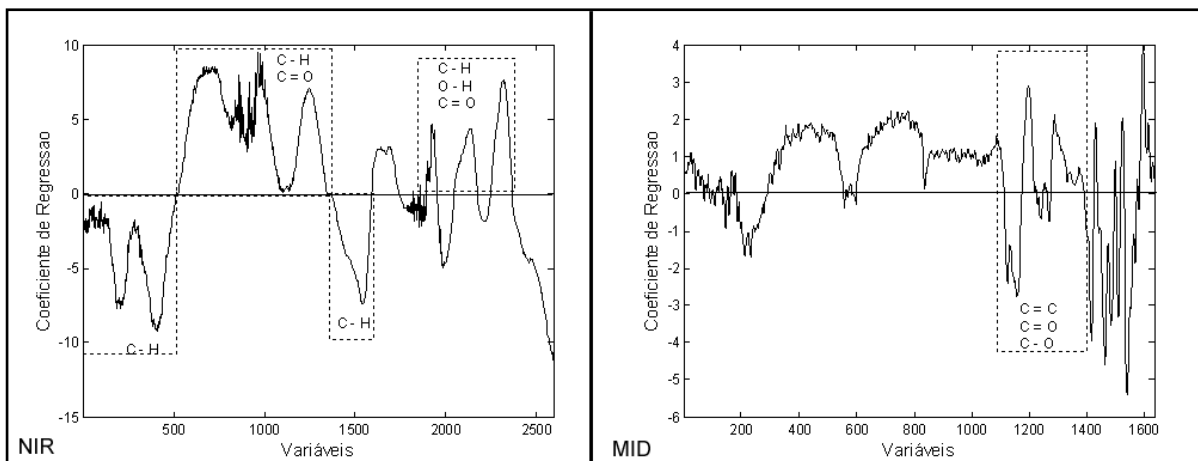


FIGURA 42 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO OLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Os valores previstos a partir dos modelos obtidos no NIR e MID para o teor de ácido oleico nas amostras do conjunto de validação externa encontram-se na Tabela 14.

TABELA 14 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO OLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA SEMENTE DE LINHAÇA

Amostras	Valor Experimental (%)	Valores Previstos (%)		Erro Padrão Relativo (%)	
		NIR	MID	NIR	MID
2	20,78	20,65	20,93	0,61	0,71
3	20,62	20,54	20,70	0,40	0,38
16	21,99	22,22	22,38	1,05	1,79
23	22,44	22,20	22,24	1,06	0,87
34	21,32	21,51	21,61	0,88	1,35
Erro relativo médio (%)				0,80	1,02

O modelo obtido no NIR apresentou menor erro padrão relativo (0,80) e as melhores características para a previsão do teor de ácido oleico para sementes de linhaça, ou seja, menores valores de RMSEC e RMSECV, maiores porcentagens de variância capturada para descrever os dados espectrais, além de maiores coeficientes de determinação (calibração e validação).

5.4.1.2 Modelos de calibração multivariada para previsão do teor de ácido linoleico para semente de linhaça

Para o desenvolvimento dos modelos para prever o teor de ácido linoleico em sementes de linhaça, foram utilizados 33 espectros no infravermelho em faixas espectrais descritas anteriormente. Os modelos PLSR e os pré-processamentos dos dados que produziram os melhores resultados estão descritos na Tabela 15. O modelo 3 apresentou o menor valor de RMSECV, usando 9 variáveis latentes na região do NIR. Todavia, o uso deste número elevado de variáveis latentes pode superestimar os modelos e influenciar na capacidade de previsão. Então, considerou-se o modelo 1 como melhor modelo, em função do seu menor erro padrão relativo de 1,61%, além de empregar um número menor de variáveis latentes (3), utilizando MSC, Alisamento e 1ª derivada como pré-processamentos.

TABELA 15 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLEICO NO NIR E MID

Região	Modelo	Pré-processamento	Nº de VLS	RMSEC	RMSECV	R ² _{cal}	R ² _{val}	Erro Padrão Relativo (%)
NIR	1	MSC/ Alisamento/ 1ª derivada	3	0,68	0,76	0,921	0,901	1,61
NIR	2	Normalização	4	0,54	0,66	0,950	0,926	4,33
NIR	3	SNV/Detrend	9	0,33	0,56	0,982	0,948	2,45
NIR	4	Alisamento/1ª derivada	4	0,86	1,14	0,881	0,796	2,76
MID	5	MSC/Normalização	5	0,82	1,16	0,884	0,757	6,57
MID	6	MSC/ Alisamento/ 1ª derivada	4	0,63	1,01	0,933	0,823	3,61
MID	7	SNV/Detrend	4	0,88	1,13	0,867	0,778	5,21

VLS = Variáveis latentes

RMSEC/ RMSECV= Raiz quadrada da soma dos erros de calibração/validação cruzada

R²_{cal/val} = Coeficiente de determinação de calibração/validação

Assim como no NIR, os pré-processamentos MSC, Alisamento e 1ª derivada, geraram o melhor modelo de regressão no infravermelho médio (MID), o modelo 6. Este modelo empregou 4 variáveis latentes, além de ter apresentado menor valor de RMSECV (1,01) e menor erro padrão relativo (3,61%) quando comparado ao erro padrão relativo obtido pelos outros modelos.

Na FIGURA 43, pode-se avaliar o efeito do número de variáveis latentes na região do NIR e do MID sobre os valores de RMSEC e RMSECV. O menor valor de RMSECV para o NIR foi obtido com 3 variáveis latentes que explicaram 99,96 % da

variância espectral e 99,85% da variabilidade do teor de ácido linoleico. Já para o MID, o menor valor de RMSECV foi obtido com 4 variáveis latentes, que descreveram 98,53% da variabilidade dos dados espectrais e 99,87% do teor de ácido linoleico nas amostras de semente de linhaça.

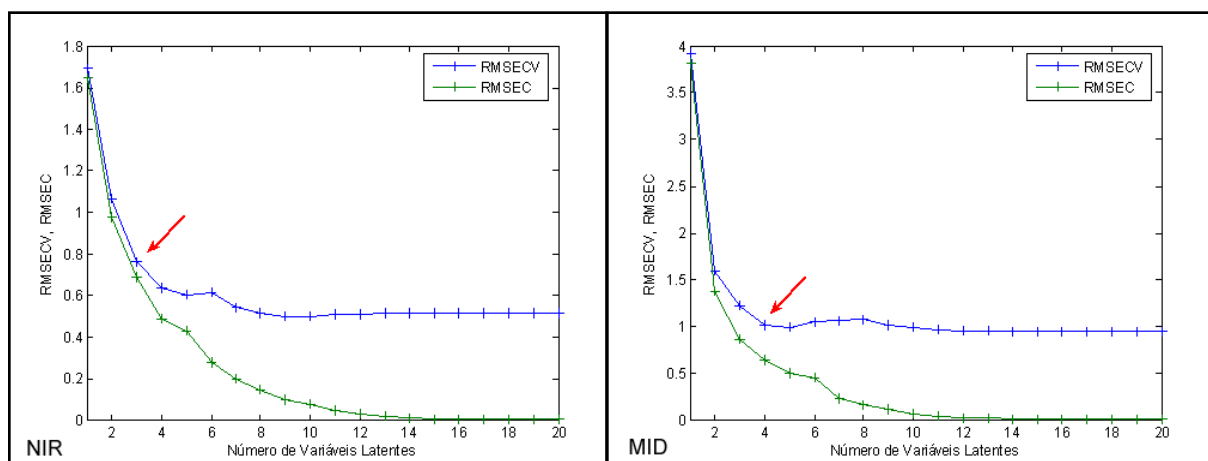


FIGURA 43 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Com base na FIGURA 44, analisou-se à presença de amostras anômalas para os limites de *leverage* de 0,27 para NIR e 0,36 para MID e, Resíduos de *Student* de $\pm 2,5$ (95% de confiança). Não houve indícios de amostras anômalas para os modelos referentes às regiões do NIR e MID.

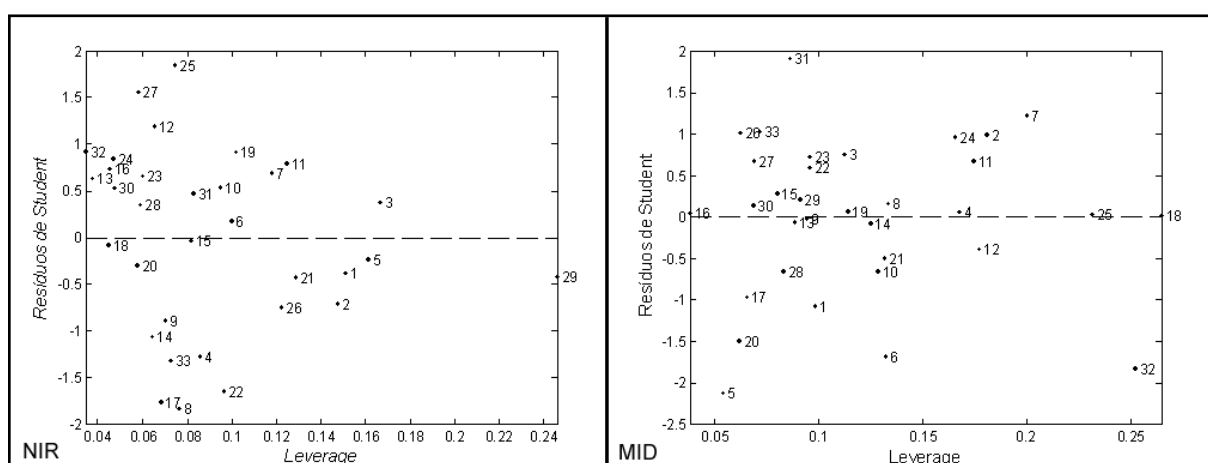


FIGURA 44 – GRÁFICOS DE LEVERAGE x RESÍDUO DE STUDENT OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Os coeficientes de determinação obtidos para os dois modelos (NIR e MID) revelaram a boa capacidade preditiva para o teor de ácido linoleico. Para o NIR, obteve-se coeficientes de determinação para amostras do conjunto de calibração e validação externa iguais a 0,921 e 0,901, respectivamente. Enquanto que para o MID, estes valores foram de 0,933 e 0,823. Os gráficos para os valores experimentais e valores previstos pelos modelos estão apresentados na FIGURA 45, nos quais é possível observar, para ambos os modelos, um grupo de amostras mais concentradas na extremidade superior. Para o modelo no MID, estas amostras apresentaram-se mais dispersas quando comparadas ao modelo no NIR.

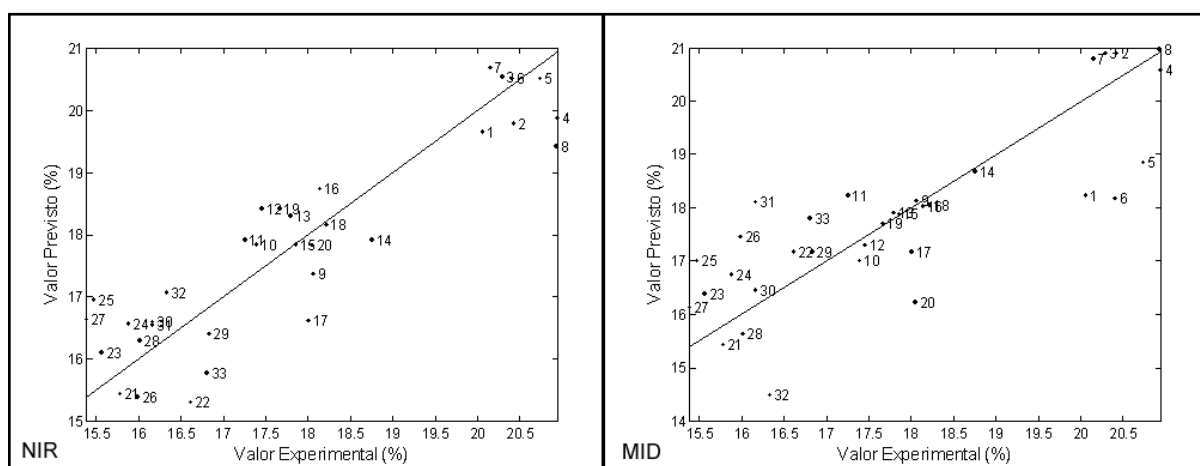


FIGURA 45 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

As regiões em destaque nos gráficos dos coeficientes de regressão evidenciaram as informações analíticas relevantes para a determinação de ácido linoleico nas regiões NIR e MID (FIGURA 46). Estas regiões incluem para o NIR, os números de onda entre 7.555 e 6.853 cm^{-1} (referente às ligações C – H, C = O, N – H) e de 5.713 a 5.010 cm^{-1} (referente às ligações C = O, C – H, N – H e O – H). Para o MID, as principais regiões possuem números de onda entre 3.892 e 3.588 cm^{-1} (referente à ligação O – H), 2.569 e 2.166 cm^{-1} (referente à ligação C = O) e de 1.202 a 900 cm^{-1} (referente às ligações C – H, C – O) (LAZZARI; CHIANTORE, 1999; SILVERSTEIN *et al.*, 2007; CEN; HE, 2007; PAVIA *et al.*, 2010; XIABO *et al.*, 2010).

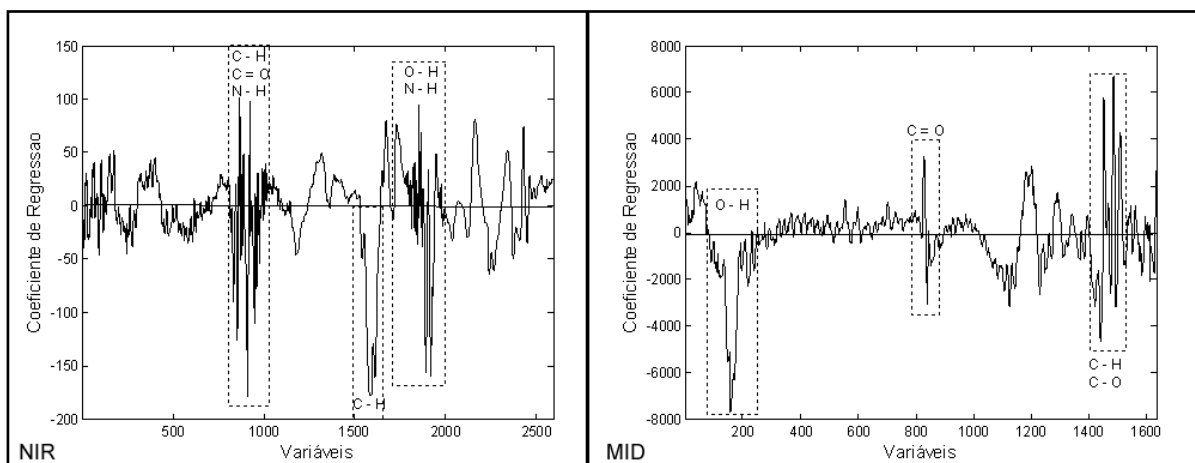


FIGURA 46 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLEICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Os resultados da previsão do teor de ácido linoleico nas amostras do conjunto de validação externa encontram-se na Tabela 16, com os valores experimentais, valores previstos e erro padrão relativo, gerados a partir dos modelos obtidos no NIR e MID. O modelo obtido no NIR apresentou as melhores características para a previsão do teor de ácido linoleico, pois empregou um menor número de variáveis latentes, além de ter apresentado menores erros de previsão (erro padrão relativo e RMSEC), maiores porcentagens de variância capturada para descrever os dados espectrais e maiores coeficientes de determinação (calibração e validação).

TABELA 16 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLEICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA SEMENTE DE LINHAÇA

Amostras	Valor Experimental (%)	Valores Previstos (%)		Erro Padrão Relativo (%)	
		NIR	MID	NIR	MID
2	20,17	20,57	21,04	1,97	4,32
3	20,14	20,35	19,56	1,06	2,89
16	17,49	17,78	18,27	1,69	4,48
23	18,19	17,88	18,64	1,69	2,48
34	15,9	16,16	16,51	1,63	3,87
Erro relativo médio (%)				1,61	3,61

5.4.1.3 Modelos de calibração multivariada para previsão do teor de ácido linolênico para semente de linhaça

Os modelos desenvolvidos para prever o teor de ácido linolênico em sementes de linhaça utilizaram faixas espectrais descritas anteriormente. Os resultados dos modelos PLSR e os tipos de pré-processamento aplicados aos dados estão apresentados na Tabela 17. O modelo 2, referente ao NIR apresentou menor valor de erro padrão relativo (0,63%) e menor valor de RMSECV, quando aplicado os pré-processamentos MSC, Alisamento e 1ª derivada. Para a região do MID, o menor erro padrão relativo (3,73%) foi obtido pelo modelo 4, que também apresentou os menores valores de RMSEC e RMSECV, após os pré-processamentos MSC e Alisamento.

TABELA 17 – MODELOS DE CALIBRAÇÃO PARA DETERMINAÇÃO DE ÁCIDO LINOLÊNICO NO NIR E MID

Região	Modelo	Pré-processamento	Nº de VLS	RMSEC	RMSECV	R ² _{cal}	R ² _{val}	Erro Padrão Relativo (%)
NIR	1	MSC/Normalização	6	0,22	0,30	0,907	0,819	0,61
NIR	2	MSC/Alisamento/1ª derivada	4	0,22	0,26	0,907	0,863	0,63
NIR	3	SNV/Detrend	4	0,26	0,30	0,863	0,814	0,79
MID	4	MSC/Alisamento	2	1,28	1,58	0,267	0,249	3,73
MID	5	Alisamento/1ª derivada	3	8,01	9,95	0,211	0,180	16,34
MID	6	SNV/Detrend	2	1,89	2,10	0,218	0,210	3,90

VLS = Variáveis latentes

RMSEC/ RMSECV= Raiz quadrada da soma dos erros de calibração/validação cruzada

R²_{cal/val} = Coeficiente de determinação de calibração/validação

O efeito do número de variáveis latentes na região do NIR e MID sobre os valores de RMSEC e RMSECV pode ser visualizado na FIGURA 47. O menor valor de RMSECV, para NIR, foi observado com 4 variáveis latentes e variância espectral de 99,97% e explicação de 100% para o teor de ácido linolênico. Já para o MID, o valor mínimo de RMSECV foi observado com 2 variáveis latentes, que explicaram 97,93% da variabilidade espectral e 99,93% do teor de ácido linolênico.

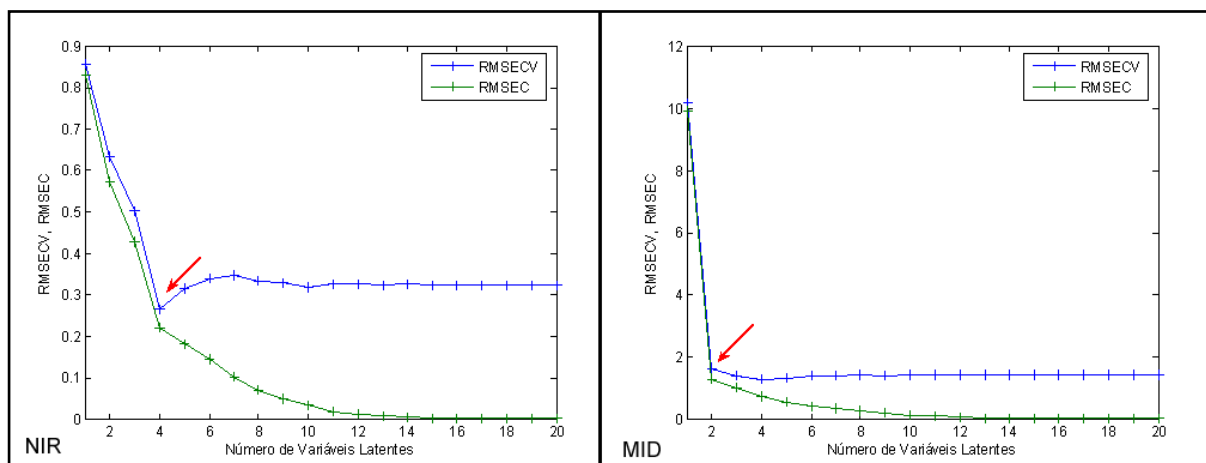


FIGURA 47 – GRÁFICOS DE RMSEC E RMSECV x NÚMERO DE VARIÁVEIS LATENTES OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

A presença de amostras anômalas pode ser analisada na FIGURA 48 através dos limites de *leverage* de 0,36 para NIR e 0,18 para MID, e Resíduos de *Student* de $\pm 2,5$ (95% de confiança). Tanto o modelo no NIR quanto no MID apresentou amostras anômalas em relação aos limites de *leverage*. A classificação das amostras 18 e 29 como anômalas pode ser atribuída a problemas associados a aquisição dos espectros, no entanto, quando estas amostras foram retiradas do conjunto de calibração, não houve alteração significativa nas características do modelo quanto à redução no número de variáveis latentes, valores de RMSEC, RMSECV e Erro Padrão Relativo, por isso as amostras foram mantidas.

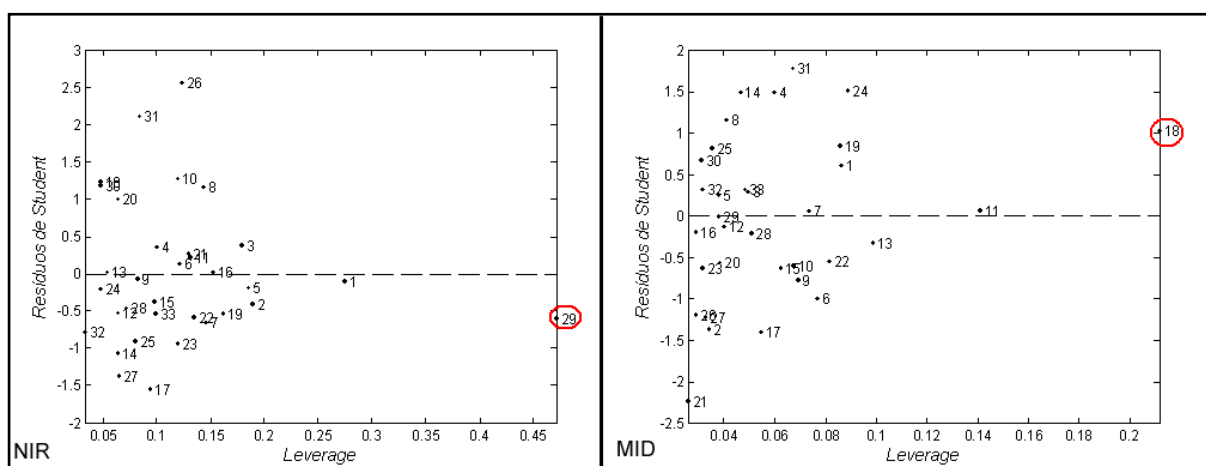


FIGURA 48 – GRÁFICOS DE *LEVERAGE* x RESÍDUO DE *STUDENT* OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

A capacidade preditiva dos modelos no infravermelho próximo e médio foi avaliada pelos valores dos coeficientes de determinação (FIGURA 49). O modelo com melhor desempenho para prever o teor de ácido linolênico foi obtido na região do NIR, o qual forneceu R^2_{cal} e validação R^2_{val} de 0,907 e 0,863, respectivamente. Já o melhor modelo no MID não apresentou boa correlação entre os dados experimentais e valores previstos, uma vez que foram obtidos os valores de 0,267 para o conjunto de calibração e de 0,249 para o conjunto de validação.

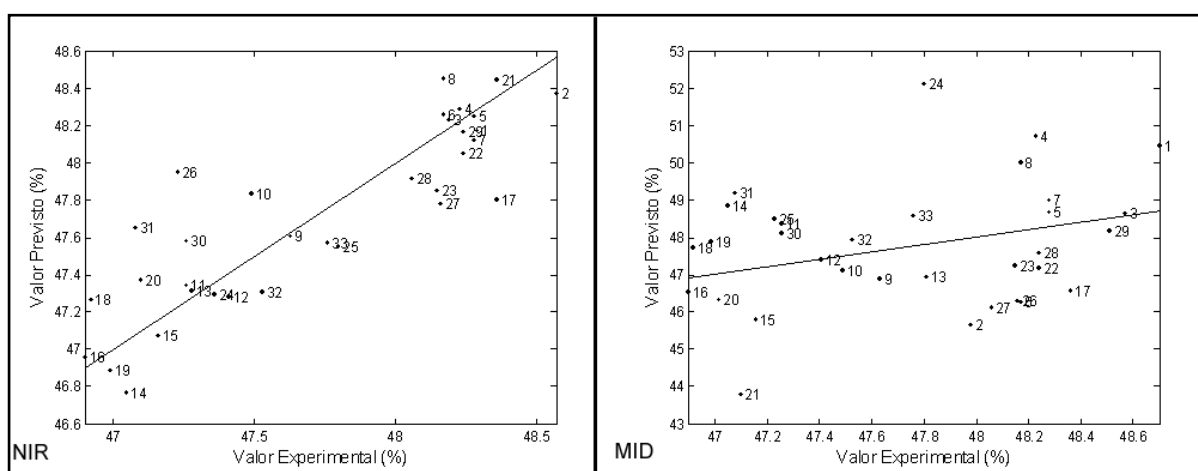


FIGURA 49 – GRÁFICOS DE VALOR PREVISTO x VALOR EXPERIMENTAL OBTIDOS PELOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Os gráficos dos coeficientes de regressão (FIGURA 50) evidenciaram as principais regiões do espectro que apresentaram informações analíticas relevantes para a previsão do teor de ácido linolênico. Estas regiões incluem para o NIR, os números de onda entre 7.686 e 6.635 cm^{-1} (referente às ligações C – H, C = O) e 5.452 a 5.145 cm^{-1} (referente às ligações C = O, O – H). Para o MID, as principais regiões possuem números de onda entre 3.892 e 3.480 cm^{-1} (referente às ligações O – H) e 1.974 a 1.670 cm^{-1} (referente à ligação C = O) (LAZZARI; CHIANTORE, 1999; SILVERSTEIN *et al.*, 2007; CEN; HE, 2007; PAVIA *et al.*, 2010; XIABO *et al.*, 2010).

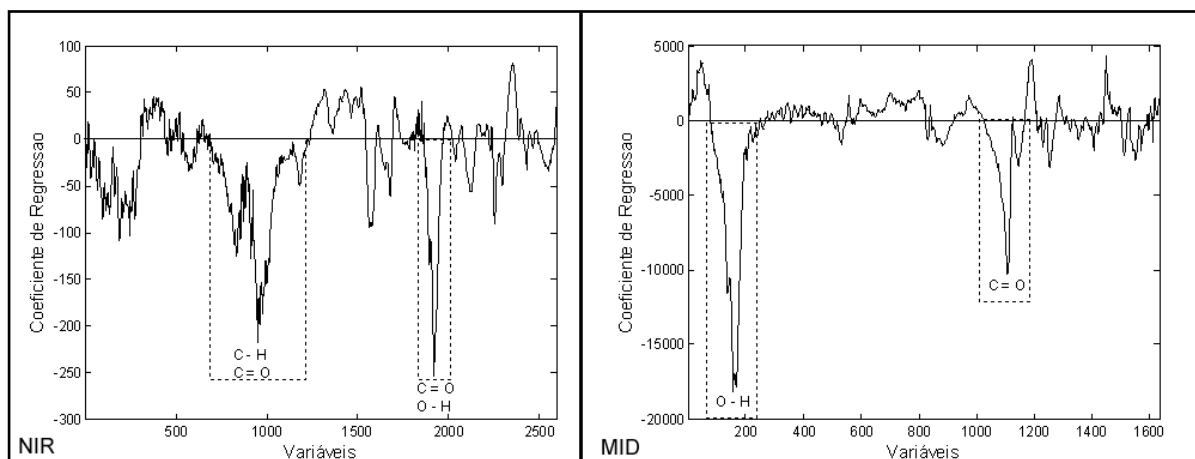


FIGURA 50 – GRÁFICO DOS COEFICIENTES DE REGRESSÃO DOS MODELOS DE CALIBRAÇÃO PARA ÁCIDO LINOLÊNICO NO INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID) PARA SEMENTES DE LINHAÇA

Os valores experimentais, valores previstos e erro padrão relativo para os modelos no NIR e no MID para a determinação do teor de ácido linoleico das amostras do conjunto de validação externa encontram-se na Tabela 18. Apesar do modelo 2 no NIR ter um número maior de variáveis latentes, em relação ao MID, o mesmo apresentou menores valores de erros de previsão (erro padrão relativo e RMSEC), maiores porcentagens de variância capturada para descrever os dados espectrais e maiores coeficientes de determinação (calibração e validação). Sendo assim, ficou evidenciado que a região no NIR forneceu o melhor modelo para determinação do teor de ácido linolênico.

TABELA 18 – RESULTADOS DA PREVISÃO DO TEOR DE ÁCIDO LINOLÊNICO PARA AS AMOSTRAS DO CONJUNTO DE VALIDAÇÃO EXTERNA, OBTIDOS PELOS MODELOS DESENVOLVIDOS PARA O INFRAVERMELHO PRÓXIMO (NIR) E MÉDIO (MID), PARA SEMENTE DE LINHAÇA

Amostras	Valor Experimental (%)	Valores Previstos (%)		Erro Padrão Relativo (%)	
		NIR	MID	NIR	MID
2	48,7	48,52	46,13	0,38	4,47
3	47,98	48,14	46,84	0,32	2,81
16	47,81	47,66	47,34	0,31	0,13
23	47,02	47,36	44,06	0,73	8,89
34	48,51	47,82	48,48	1,43	2,36
Erro relativo médio (%)				0,63	3,73

5.5 COMPARAÇÃO DOS MODELOS DE CALIBRAÇÃO MULTIVARIADA PARA NIR E MID

Para a previsão do teor de ácido oleico, observou-se que o melhor modelo para ambos, farinhas e sementes, foi obtido no NIR. Para farinhas de linhaça, o erro padrão relativo foi de 1,02%, o RMSECV de 0,33 e R^2_{val} de 0,989, com uso de 2 variáveis latentes. Já para as sementes, o melhor modelo apresentou erro padrão relativo de 0,80%, RMSECV de 0,34 e R^2_{val} de 0,81, com uso de 4 variáveis latentes. Os valores obtidos de modelos para semente de linhaça neste trabalho foram próximos aos valores determinados por modelos obtidos por Patil *et al.* (2010), que encontraram para o teor de ácido oleico em sementes de soja um coeficiente de validação igual 0,886 e RMSECV de 0,90.

Na determinação da composição de ácidos graxos em sementes de girassol através de modelos de calibração, Cantarelli *et al.* (2009) encontraram para ácido oleico um valor satisfatório de R^2_{val} igual a 0,995, maior do que aquele encontrado neste trabalho para sementes de linhaça, no entanto, o valor RMSECV (4,72) obtido por Cantarelli *et al.* (2009) é muito mais elevado quando comparado ao valor de RMSECV (0,34) determinado para sementes de linhaça neste trabalho.

Para a determinação do teor de ácido linoleico em farinhas de linhaça, o melhor modelo foi obtido no MID, com erro padrão relativo de 0,76, RMSECV de 0,95, e R^2_{val} de 0,883, com uso de 4 variáveis latentes. No entanto, quando se desenvolveu modelos para sementes de linhaça, observaram-se melhores resultados na região do NIR com erro padrão relativo de 1,61%, RMSECV de 0,76 e R^2_{val} de 0,901, com uso de 3 variáveis latentes. Os resultados obtidos neste trabalho apresentam melhor capacidade preditiva, quando comparado, por exemplo, a resultados encontrados por Kim *et al.* (2007), que avaliaram a composição de ácidos graxos em sementes de canola, e estimaram para o ácido linoleico, coeficiente de validação de 0,847 e RMSECV de 1,296.

Com relação à determinação do teor de ácido linolênico, os melhores modelos tanto para farinhas quanto para sementes de linhaça foram obtidos na NIR, nos quais 2 variáveis latentes foram responsáveis pelo erro padrão relativo de 1,21% RMSECV de 0,64 e R^2_{val} de 0,992, para farinhas. Para as sementes, o RMSECV foi 0,26, o erro relativo 0,63% e R^2_{val} de 0,863, com 4 variáveis latentes. Os resultados

obtidos para modelos desenvolvidos neste trabalho apresentam melhor capacidade preditiva, quando comparado, por exemplo, a resultados de modelos desenvolvidos por Patil *et al.* (2010), que avaliaram a composição de ácidos graxos em sementes de soja, e estimaram para o ácido linolênico valores de R^2_{val} de 0,776 e RMSECV de 0,29.

As diferenças que foram observadas entre os modelos desenvolvidos neste trabalho podem ser atribuídas a heterogeneidade e complexidade das amostras de farinha e semente de linhaça. No entanto, em todos os modelos desenvolvidos, tanto no NIR quanto no MID, observou-se que os coeficientes de regressão evidenciaram principalmente ligações C – H, O – H e C = O, características de ácidos carboxílicos. Desta forma, pode-se considerar que os modelos apresentaram boa capacidade de previsão e especificidade para os teores de ácidos graxos majoritários em farinhas e sementes de linhaça (oleico, linoleico e linolênico).

6 CONCLUSÕES

A análise das componentes principais permitiu agrupar as amostras de farinha e semente de linhaça de acordo com suas diferenças e similaridades, com base nos dados espectrais gerados pelo infravermelho, tendo sido a região do NIR mais eficiente no agrupamento das amostras.

A aplicação da espectroscopia no infravermelho próximo e médio, por reflectância difusa, aliada ao método de Regressão por Mínimos Quadrados Parciais (PLSR), possibilitou o desenvolvimento de modelos de regressão multivariados com elevados coeficientes de regressão e baixos erros de previsão. Estes modelos se mostraram eficientes para previsão dos teores de ácidos oleico, linoleico e linolênico.

Os modelos para os ácidos oleico, linoleico e linolênico referentes às sementes de linhaça, foram melhores na região do NIR, onde as médias de erro padrão relativo foram de 0,80% (com 4 variáveis latentes), 1,61% (com 3 variáveis latentes) e 0,63% (com 4 variáveis latentes), respectivamente para cada ácido graxo.

Para as farinhas de linhaça, o melhor modelo para teor de ácido linoleico foi obtido na região do infravermelho médio, sendo o erro padrão relativo de 0,76% (com 4 variáveis latentes), enquanto que para os ácidos oleico e linolênico, a região do infravermelho próximo forneceu os melhores modelos, com erro padrão relativo de 1,02% (com 4 variáveis latentes) e 1,21% (com 2 variáveis latentes), respectivamente.

A espectroscopia na região do infravermelho próximo e médio associada às técnicas de calibração multivariada mostrou-se eficiente na determinação do teor de ácidos graxos em farinhas e sementes de linhaça, podendo ser aplicada como método de controle de qualidade nas indústrias de alimentos.

7 SUGESTÕES PARA TRABALHOS FUTUROS

A respeito do estudo realizado sobre o desenvolvimento de modelos de calibração multivariada, baseados em análises de espectroscopia no infravermelho para determinação do teor de ácidos graxos, seguem-se algumas sugestões para trabalhos futuros:

- ✓ Comparar diferentes técnicas de extração de lipídeos totais a fim de obter o máximo rendimento.
- ✓ Realizar o processo de derivatização com diferentes agentes esterificantes com o intuito de obter o maior número possível de ácidos graxos.
- ✓ Desenvolver modelos com todos os ácidos graxos presentes nas farinhas e sementes de linhaça.
- ✓ Aplicar a metodologia espectroscópica multivariada para avaliar mudanças nos parâmetros modelados durante a vida de prateleira das farinhas e sementes de linhaça.
- ✓ Determinar os teores dos ácidos graxos através de Redes Neurais.

REFERÊNCIAS

ABDI, H.; Partial Least Squares (PLSR) Regression. In LEWIS-BECK, M.; BRYMAN, A.; FUTING, T. **Encyclopedia for Research Methods for the Social Sciences**. 2003.

APARÍCIO, S. M. **Metodologías analíticas basadas en espectroscopia de infrarrojo y calibración multivariante: aplicación a la industria petroquímica**. 222 f. Tesis (Doctoral) – Departament de Química Analítica i Química Orgànica, Universitat Rovira i Virgili, Tarragona, 2002.

AOCS. OFFICIAL METHODS AND RECOMMENDED PRACTICES OF THE AMERICAN OIL CHEMISTS SOCIETY. 5 th. ed. Champaign: AOCS, 2004, met. Af 3 – 54.

BASTIEN, P., ESPOSITO-VINZI, V., TENENHAUS, M. PLSR Generalised Regression. **Computational Statistics & Data Analysis**, v. 48, p. 17- 46, 2005.

BENATTI, P. *et al.* Polyunsaturated Fatty Acids: Biochemical, Nutritional and Epigenetic Properties. **Journal of the American College of Nutrition**, v. 23, n. 4, p. 281–302, 2004.

BLANCO, M. *et al.* Calibration in non-linear near infrared reflectance spectroscopy: a comparison of several methods. **Analytica Chimica Acta**. p. 207 – 214, 1999.

BOZAN, B.; TEMELLI, F. Supercritical CO₂ extraction of flaxseed. **Journal American Oil Chemists Society**. v. 79, n. 3, p. 231 – 235, 2002.

BROAD, N. W. *et al.* Application of transmission near-infrared spectroscopy to uniformity of content testing of intact steroid tablets. **Analyst**, p. 2207-2211, 2001.

BRUICE, P. Y. Química Orgânica. 4 ed. São Paulo: Pearson/Prentice Hall, 2006.

CAMPESTRINI, E. Utilização de Equipamento NIRS (Near Infrared Reflectance Spectroscopy) nos estudos de valores nutricionais (Composição química e digestibilidade) de Alimentos para não ruminantes. **Revista Eletrônica Nutritime**, v. 2, n.5, p. 240 – 251, 2005.

CANTARELLI, M. A. *et al.* Determination of oleic acid in sunflower seeds by infrared spectroscopy and multivariate calibration method. **Talanta**. v. 80, p. 489 – 492, 2009.

CASALE, M. *et al.* Characterization of table olive cultivar by NIR spectroscopy. **Food Chemistry**, v.122, p. 1261–1265, 2010.

CEN, H.; HE, Y. Theory and application of near infrared reflectance spectroscopy in determination of food quality trends. **Food Science and Technology**, v. 18, p. 72 – 83, 2007.

CHAMINADE, P. *et al.* Data treatment in near infrared spectroscopy. **Analysis**, v. 26, n. 4, p. 33 – 38, 1998.

CHEN, H. H.; XU, S. Y.; WANG, Z. Gelation properties of flaxseed gum. **Journal of Food Engineering**. v. 77, n. 2, p. 295 – 303, 2006.

CHIANG, L. H.; COLEGROVE, L. F. Industrial implementation of on-line multivariate quality control. **Chemometrics and Intelligent Laboratory Systems**., v. 88, p. 143-153, 2007.

CHUNG, M., LEI, B., LI-CHAN, E. Isolation and structural characterization of the major protein fraction from NorMan flaxseed (*Linum usitatissimum* L.). **Food Chemistry**. v. 90, p. 271–279, 2005.

COSKUNER, Y.; KARABABA, E. Some physical properties of flaxseed (*Linum usitatissimum* L.). **Journal of Food Engineering**, Mersin – Turquia, p. 1067 – 1073, 2007.

COZZOLINO, D. *et al.* Multivariate determination of free fatty acids and moisture in fish oils by partial least-squares regression and near-infrared spectroscopy. **LWT**. v. 38, p. 821 – 828, 2005.

CUI, W. S. Flaxseed: a functional food for the 21st century. **Canadian Chemical News**. v. 50, n. 5, p. 19, 1998.

DAGNEW, M.; CROWE, T. G.; SCHOENAU, J. J. Sensing of hog manure nutrients with reflectance spectroscopy. SAE/SCGR-NABEC. **Meeting Guelph**, Canada. 2001.

DAMODARAN, S.; PARKIN, K.L.; FENNEMA, O.R. **Química de Alimentos de Fennema**. Tradução BRANDELLI, Adriano *et al.* 4. ed. Porto Alegre: Artmed, 2010.

DAUN, J.K. *et al.* Structure, composition and variety development of flaxseed. In: THOMPSON, L.U., CUNNANE, S.C., **Flaxseed in Human Nutrition**, 2. ed. Champaign, Illinois, p. 1 – 40, 2003.

DAVIS, S. F. The effect of soy flour as a natural antioxidant on flaxseed in yeast bread. In: CONFORTI, F. D; MARCY, J. E.; NIBA, L. L. **Human Nutrition, Foods and Exercise**. Blacksburg, Virginia, p. 86, 2004.

DI – EGIDIO, V. *et al.*, Evaluation of shelf-life of fresh-cut pineapple using FT-NIR and FT-IR spectroscopy. **Postharvest Biology and Technology**. v. 54, p. 87– 92, 2009.

EL-NAGDY, G. A. *et al.* Response of Flax Plant (*Linum usitatissimum* L.) to Treatments with Mineral and Bio-Fertilizers from Nitrogen and Phosphorus. **Journal of American Science**, p. 207 – 217, 2010.

EPAMINONDAS, P. S. **Caracterização Físico-Química de Termo-Oxidativadas Sementes de Linhaça (*Linum usitatissimum* L.) e de seus Óleos**. Dissertação (Mestrado) Universidade Federal da Paraíba, 2009.

FAGAN, C. C., O'DONNELL, C. P. Application of mid-infrared spectroscopy to food processing system. In: IRUDAYARAJ, J., RECH, C. **Nondestructive Testing of Food Quality**. 1. ed. IFT Press, Iowa, p. 119 – 142, 2008.

FASSIO, A., D. COZZOLINO. Non-destructive prediction of chemical composition in sunflower seeds by near infrared spectroscopy. **Industrial Crops and Products**, v. 20, p. 321–329, 2004.

FERRARINI, H. Determinação dos Teores Nutricionais do Milho por Espectroscopia no Infravermelho e Calibração Multivariada. Dissertação (Mestrado) Universidade Federal do Paraná, 2004.

FERREIRA, M. M. C. *et al.* Quimiometria I: Calibração Multivariada, um tutorial. **Química Nova**, v. 5, n. 22, 1999.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). **Food and Agricultural Commodities Production**. Linseed statistics. 2010.

GALVÃO, E. L. **Extração supercrítica de óleo de linhaça: construção do extrator, estudo de parâmetros de processo, avaliação química e antioxidante do produto**. Tese – (Doutorado) Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, 2009.

GARCÍA, M. C.; TORRE, M.; MARINA, M. L.; Characterization of commercial soybean products by conventional and perfusion reversed-phase high-performance liquid chromatography and multivariate analysis. **Journal of Chromatography A.**, v. 881, n. 1-2 , p. 47-57. Jun., 2000.

GELADI, P. Chemometrics in spectroscopy. Part 1. Classical chemometrics. **Spectrochimica Acta Part B.** v. 58, p. 767 – 782, 2003.

GIVENS, D. I. *et al.* The principles, practices and some future applications of near infrared spectroscopy for predicting the nutritive value of foods for animals and humans. **Nutrition Research Reviews**, v. 10, p. 83 – 114, 1997.

GRAZIOLA, F.; SOLIS, V. S.; CURI, R. Estrutura Química e Classificação dos ácidos graxos. In: CURI, R.; POMPEIA, C. MIYASAKA, C. K.; PROCOPIO, J. **Entendendo a Gordura – os ácidos graxos**. São Paulo: Manole, 2002.

GUNSTONE, F. D. **The Chemistry of oils and fats: sources, composition, properties and uses**. Dundee: CRC Press, 2004.

GUNSTONE, F. D.; HARWOOD, J. L.; DIJKSTRA, A. J. **The lipid handbook**. 3. ed. New York: CRC Press, 2007.

GURDENIZ, G., OZEN, B., TOKATLI, F. Classification of Turkish olive oils with respect to cultivar, geographic origin and harvest year, using fatty acid profile and mid-IR spectroscopy. **European Food Research Technology**. v. 227, p. 1275-1281, 2008.

HARRIS, D. C. **Análise Química Quantitativa**. 7. ed. Rio de Janeiro: LTC, 2008.

HELPER, G. A. *et al.* Aplicação de Métodos de Análise Multivariada no Controle Qualitativo de Essências Alimentícias Empregando Espectroscopia no Infravermelho Médio. **Ciência e Tecnologia de Alimentos**, v. 26, n. 4, 2006.

HIGSON, S. P. J. Química Analítica. São Paulo: McGraw-Hill, 2009.

HOLLER, F. J.; SKOOG, D. A.; CROUCH, S. R. **Princípios de Análise Instrumental**. 6. ed. Porto Alegre: Bookman, 2009.

HU, F. B.; WILLETT, W. C. Optimal Diets for Prevention of Coronary Heart Disease. **Journal of the American Medical Association**, v. 288, n. 20, p. 2569 – 2578, 2002.

JHALA, A. J.; HALL, L. M. Flax (*Linum usitatissimum* L.): Current Uses and Future Applications. **Australian Journal of basic and Applied Sciences**, Edmonton – Canadá, p. 4304-4312, 2010.

KARÁSEK, P. *et al.* Direct continuous supercritical fluid extraction as a novel method of wine analysis. Comparison with conventional indirect extraction and implications for wine variety identification. **Journal of Chromatography**, v. 1002, n. 1- 2, p. 13 – 23, 2003.

KEEFE, S. F. Nomenclature and classification of lipids. In: AKOH, C. C.; MIN, D. B. **Food lipids: chemistry, nutrition and biotechnology**. 3. ed. New York: CRC Press, 2008.

KELLNER, R. *et al.* **Analytical Chemistry: A Modern Approach to Analytical Science**, WILEY-VCH, 2. ed., 2004.

KIM, K. S. *et al.* Use of Near-Infrared Spectroscopy for Estimating Fatty Acid Composition in Intact Seeds of Rapeseed. **Journal of Crop Science and Biotechnology**. v. 10, p. 15 – 20, 2007.

LAMPE, J. W. Isoflavonoid and Lignan Phytoestrogens as Dietary Biomarkers. **Journal of Nutrition**. v. 133, n. 3, p. 956-964, 2003.

LANKMAYR, E. *et al.* Chemometrical classification of pumpkin seed oils using UV–Vis, NIR and FTIR spectra. **Journal of Biochemical and Biophysical Methods**, v. 61, p. 95–106, 2004.

LAZZARI, M.; CHIANTORE, O. Drying and oxidative degradation of linseed oil. **Polymer Degradation and Stability**. v. 65, p. 303 – 313, 1999.

LUYPAERT, J. *et al.* An evaluation of direct orthogonal signal correction and other preprocessing methods for the classification of clinical study lots of a dermatological cream. **Journal of Pharmaceutical and Biomedical Analysis**, v. 30, p. 453-466, 2002.

MARQUES, A. C. **Propriedades funcionais da linhaça (*Linum usitatissimum* L.) em diferentes condições de preparo e de uso em alimentos.** Dissertação - (Mestrado) Centro de Ciências Rurais, Universidade Federal de Santa Maria, Rio Grande do Sul, 2008.

MENDHAM, J. *et al.* **VOGEL: Análise Química Quantitativa.** 6. ed. Rio de Janeiro: LCT, 2002.

MORRIS, D.H. Essential nutrients and other functional compounds in flaxseed. **Nutrition Today.** v. 36, n. 3, p.159-162, 2001.

MORRIS, D. H. *et al.* Dietary fats and the risk of incident Alzheimer disease. **Archives of Neurology.** v. 60, n. 8, p. 1072, 2003.

MORRIS, D. H. Description and Composition of Flax. In.____. **Flax – A Health and Nutrition Primer.** 4. ed, p. 9 – 21, 2007.

MUELLER, K. *et al.* Functional properties and chemical composition of fractionated brown and yellow linseed meal (*Linum usitatissimum* L.). **Journal of Food Engineering.** v. 98, p. 453–460, 2010.

NAES, T. *et al.* A User-friendly Guide to Multivariate Calibration and Classification. **NIR Publications,** 2002.

NAGATA, N. **Uso de métodos de calibração multivariada na análise de amostras de interesse industrial e espécies de importância ambiental por fluorescência e raios - X.** Tese (Doutorado) Universidade Estadual de Campinas, 2001.

NAGATA, N.; BUENO, M. I. M. S.; PERALTA-ZAMORA, P. Métodos Matemáticos para a correção de interferências espectrais e efeitos interelementos na análise quantitativa por fluorescência de raio X. **Química Nova,** v. 24, n. 4, p. 531 – 539, 2001.

NATALINO, R. **Caracterização de açúcar mascavo aplicando análise das componentes principais a dados espectrométricos.** Tese (Doutorado) Universidade Federal de Viçosa, 2006.

NELSON, D. L.; COX, M. M. **Lehninger – Princípios de Bioquímica**. Tradução: SIMÕES, Arnaldo Antonio; LODI, Wilson Roberto Navega. 3. ed. São Paulo: Sarvier, 2002.

OETTERER, M.; REGITANO-D'ARCE, M. A.B.; SPOTO, M. **Fundamentos de Ciência e Tecnologia de Alimentos**. São Paulo: Manole, 2006.

O'HAVER, T. **An introduction to signal processing in chemical analysis**. Department of Chemistry and Biochemistry, University of Maryland. 2001. Disponível em: <http://terpconnect.umd.edu/~toh/spectrum/IntroToSignalProcessing.pdf>> Acesso em 15/04/2011.

OOMAH, B.D.; MAZZA, G. Flaxseed proteins — A review. **Food Chemistry**. v. 48, p. 109-114, 1993.

OOMAH, B.D.; MAZZA, G. Productos de linaza para la prevención de enfermedades. In: MAZZA, G. **Alimentos funcionales: aspectos bioquímicos y de procesado**, Zaragoza: Acribia, 2000.

OTTO, M. **Chemometrics : statistics and computer application in analytical chemistry**. Weinheim : Wiley-VCH, 2007.

OZAKI, Y. *et al.* **Near-infrared spectroscopy in food science and technology**. Wiley Inter-science, 2007.

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

PATIL, A. G. Nondestructive estimation of fatty acid composition in soybean [*Glycine max* (L.) Merrill] seeds using Near-Infrared Transmittance Spectroscopy. **Food Chemistry**. v. 120, p. 1210 – 1217, 2010.

PAVIA, D. L. *et al.* **Introdução à Espectroscopia**. Tradução de: BARROS, Pedro. 4. ed. São Paulo: Cengage Learning, 2010.

PITA, M. C. G. *et al.* Efeito da suplementação de linhaça, óleo de canola, vitamina E na dieta sobre as concentrações de ácidos graxos poliinsaturados em ovos de galinha. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**. v. 58, n. 5, p. 925 – 931, 2006.

PITAT, B.; ZADERNOWSKI, R. Physicochemical characteristics of linseed oil and flour. **Polish Journal of Natural Sciences**. v. 25, n. 1, p. 106-113, 2010.

PRADHAN, R. C. *et al.* Supercritical CO₂ extraction of fatty oil from flaxseed and comparison with screw press expression and solvent extraction processes. **Journal of Food Engineering**. v. 98, p. 393 – 397, 2010.

PSOTA, T. L.; GEBAUER, S. K.; KRIS-ETHERTON, P. Dietary Omega-3 fatty acid intake and cardiovascular risk. **The American Journal of Cardiology**. v. 98, 2006.

RAJESHA, J. *et al.* Antioxidant potentials of flaxseed in vivo model. **Journal Agricultural Food Chemistry**, v. 54, p. 3794-3799, 2006.

REEVES III; J. B.; Mid-infrared diffuse reflectance spectroscopy: Is sample dilution with KBr necessary, and if so, when? **American Laboratory**, p. 24 – 28, 2003.

RUBILAR, M. *et al.* Flaxseed as a source of functional ingredients. **Journal Soil Science Plant Nutrition**. v.10. n. 3, p. 373 – 377, 2010.

SALMERÓN J, *et al.* **Journal of the American Medical Association**, v. 277 p.472-477, 1997.

SAUER, E. **Desenvolvimento de metodologias multivariadas para análise de queijos por espectroscopia DRIFT**. Tese (Doutorado) Universidade Federal do Paraná, p. 157, 2007.

SIESLER, H. W. *et al.* **Near-Infrared Spectroscopy**. Principles, Instruments, Applications, Wiley – VCH, 2002.

SILVERSTEIN, R. M. *et al.* **Identificação Espectrométrica de Compostos Orgânicos**. 7. ed. Rio de Janeiro: LTC, 2007.

SINELLI, *et al.* Varietal discrimination of extra virgin olive oils by near and mid infrared spectroscopy. **Food Research International**. v. 43, p. 2126 – 2131, 2010.

SHIN, E. C. *et al.* Chemometric approach to fatty acid profiles in Runner-type peanut cultivars by principal component analysis (PCA). **Food Chemistry**. v. 119, p. 1262 – 1270, 2010.

SHIROMA, C.; RODRIGUEZ – SAONA, L. Application of NIR and MID spectroscopy in quality control of potato chips. **Journal of Food Composition and Analysis**, v. 22, p. 596 – 605, 2009.

SKOOG, D. A *et al.* **Fundamentos de química analítica**. 8. ed. Madrid: Thomson-Paraninfo, 2005.

SKOOG, D. A, HOLLER, F. J., NIEMAN, T. A. **Princípios de Análise Instrumental**. 5. ed. Porto Alegre: Bookman, 2002.

SOTELO, F. F. **Aplicação da espectroscopia de infravermelho próximo na caracterização de petróleo**. Simulação de uma unidade de destilação atmosférica. Tese (Doutorado) Universidade de São Paulo, 2006.

SPILLER, R. C. Pharmacology of dietary fibre. **Pharmacology Therapeutics**, v. 62, p. 407– 427, 1994.

SPITELLER, G. Peroxyl radicals: Inductors of neurodegenerative and other inflammatory diseases. Their origin and how they transform cholesterol, phospholipids, plasmalogens, polyunsaturated fatty acids, sugars, and proteins into deleterious products. **Free Radical Biology & Medicine**. v. 41, p. 362–387, 2006.

SOLOMONS, T. W. G.; FRYHLE, C. B. **Química Orgânica**. 8. ed. Rio de Janeiro: LTC, 2006.

STRUIJS, K. *et al.* The flavonoid herbacetin diglucoside as a constituent of the lignan macromolecule from flaxseed hulls. **Phytochemistry**. v. 68, n. 8, p. 1227–1235, 2007.

SUNDARAM, J. *et al.* Determination of In-Shell Peanut Oil and Fatty Acid Composition Using Near-Infrared Reflectance Spectroscopy. **Journal of the American Oil Chemists' Society**, v.87, p. 1103–1114, 2010.

TARPILA, A.; WENNBERG, T.; TARPILA, S. Flaxseed as a functional food. **Current Topics in Nutraceutical Research**. v. 3, n. 3, p. 167-188, 2005.

THOMAS, E. V. A primer on multivariate calibration. **Analytical Chemistry**. v. 66, n. 15, p. 795-804,1994.

TOZETTO, A. **Controle de Qualidade de Edulcorantes em Adoçantes Comerciais via Espectrometria e Métodos de Calibração Multivariada**. Dissertação (Mestrado) Universidade Estadual de Ponta Grossa, 2005.

United States Department of Agriculture, National Agriculture Library (USDA). **Food and Nutrition**, 2012.

VAISEY-GENSER M, MORRIS DH. Introduction – history of the cultivation and uses of flaxseed. In: MUIR, A.D.; WESTCOTT, N. D. **Flax – The Genus Linum**, New York, p 1-21, 2003.

VANDEGNISTE, B. G. M. *et al.* **Handbook of Chemometrics and Qualimetrics – part B**, Amsterdam, 1998.

VELASCO, L.; BECKER, H, C. Estimating the fatty acid composition of the oil in intactseed rapeseed (*Brassica napus* L.) by nearinfrared reflectance spectroscopy. **Euphytica**, n.101, p. 221-230, 1998.

VISENTAINER, J. V.; FRANCO, M. R. B. **Ácidos graxos em óleos e gorduras: identificação e quantificação**. São Paulo: Varela, 2006.

VROMANS, J. **Molecular genetic studies in flax**. Wageningen: [S. l.: s.n.], 2006.

XIABO, Z. *et al.* Variables selection methods in near-infrared spectroscopy. **Analytica Chimica Acta**, v. 667, p. 14–32, 2010.

XU, Q. S.; LIANG, Y. Z. Monte Carlo cross validation. **Chemometrics and Intelligent Laboratory Systems**, v. 56, p.1-11, 2001.

WARD, O. P.; SINGH, A. Omega-3/6 fatty acids: Alternative sources of production. **Process Biochemistry**. v. 40, p. 3627–3652, 2005.

WHITE, P. J. Fatty acids in oilseeds (vegetable oils). In: CHOW, C. K. **Fatty acids in Foods and Their Health Implications**. 3. ed., CRC Press, p. 227 – 262, 2007.

WISE, B. M. *et al.* **PLSR toolbox for use with MATLAB**. version 3. engvevector Research, Inc., 2003.

YEHUDA, S. *et al.* The role of polyunsaturated fatty acids in restoring the aging neuronal membrane. **Neurobiology of Aging**. v. 23, p. 843 – 853, 2002.

YOUDIM, K. A.; MARTIN, A.; JOSEPH, J. A. Essential fatty acids and the brain: possible health implications. **International Journal of Developmental Neuroscience**. v. 18, p. 383 – 399, 2000.

ZHANG, G. *et al.* Effects of processing and storage condition on phenolic concentration and antioxidant activities of apple and apple juices. **Journal Food Science Technology**. v.45, p. 339–343, 2008.

ZOHARY, D.; HOPF, M. **Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe and the Nile Valley**. Oxford University Press, Oxford, 1993. E- book.