



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA
LABORATÓRIO DE QUIMIOMETRIA DO TRIÂNGULO

ADEMAR DOMINGOS VIAGEM MÁQUINA

ESPECTROMETRIA NO INFRAVERMELHO MÉDIO E MÉTODOS
QUIMIOMÉTRICOS PLS-DA E PLS: classificação e previsão do teor de biodiesel na
mistura de biodiesel/diesel de mafurra, moringa e algodão

UBERLÂNDIA

2017

ADEMAR DOMINGOS VIAGEM MÁQUINA

**ESPECTROMETRIA NO INFRAVERMELHO MÉDIO E MÉTODOS
QUIMIOMÉTRICOS PLS-DA E PLS: classificação e previsão do teor de biodiesel na
mistura de biodiesel/diesel de mafurra, moringa e algodão**

Dissertação apresentada ao Programa de Pós-Graduação em Química do Instituto de Química da Universidade Federal de Uberlândia, como requisito para obtenção de título de Mestre em Química.

Área de Concentração: Química Analítica

Orientador: Prof. Dr. Waldomiro Borges Neto

UBERLÂNDIA

2017

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

- M297e
2017 Máquina, Ademar Domingos Viagem, 1987-
Espectrometria no infravermelho médio e métodos quimiométricos
PLS-DA E PLS: classificação e previsão do teor de biodiesel na mistura
de biodiesel/diesel de mafurra, moringa e algodão / Ademar Domingos
Viagem Máquina. - 2017.
84 f. : il.
- Orientador: Waldomiro Borges Neto.
Dissertação (mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-Graduação em Química.
Inclui bibliografia.
1. Química - Teses. 2. Biodiesel - Teses. 3. Quimiometria - Teses. 4.
Biodiesel - Controle de qualidade - Teses. I. Borges Neto, Waldomiro. II.
Universidade Federal de Uberlândia. Programa de Pós-Graduação em
Química. III. Título.



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA



Ata da defesa de DISSERTAÇÃO DE MESTRADO junto ao Programa de Pós-Graduação em Química, do Instituto de Química da Universidade Federal de Uberlândia
DEFESA DE DISSERTAÇÃO DE MESTRADO ACADÊMICO, NÚMERO 246/PPQUI.

DATA: 20/01/2017

DISCENTE: Ademar Domingos Viagem Máquina

MATRÍCULA: 11512QMI001

TÍTULO DO TRABALHO: Espectrometria no infravermelho médio e métodos quimiométricos PLS-DA e PLS: classificação e previsão do teor de biodiesel na mistura de biodiesel/diesel de mafurra, moringa e algodão.

ÁREA DE CONCENTRAÇÃO: Química

LINHA DE PESQUISA: Espectroanalítica Aplicada

PROJETO DE PESQUISA DE VINCULAÇÃO: Controle de qualidade de biocombustíveis.

Às quatorze horas do dia vinte de janeiro do ano dois mil e dezessete, no Auditório Prof. Manuel Gonzalo Hernández Terrones, no Campus Santa Mônica, reuniu-se a Banca Examinadora composta pelos Professores Doutores Frederico Garcia Pinto, da Universidade Federal de Viçosa, Hélder Eterno da Silveira e Waldomiro Borges Neto, do Instituto de Química da UFU, professor(a) orientador(a) e presidente da mesa. Iniciando os trabalhos, o(a) presidente da mesa apresentou o(a) candidato(a) e a Banca Examinadora, agradeceu a presença do público e discorreu sobre as normas e critérios para a realização desta sessão, baseadas na Norma Interna nº 03/2013/PPQUI. Em seguida, o(a) presidente da mesa concedeu a palavra ao(a) candidato(a) para a exposição do seu trabalho e, em seqüência, aos examinadores, em ordem sucessiva, para arguir o(a) apresentador(a). A duração da apresentação e o tempo de arguição e resposta deram-se conforme as normas do Programa. Ultimada a arguição, desenvolvida dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu os conceitos finais e A provou o(a) candidato(a). Por sugestão da Banca Examinadora, o título do trabalho será

mantido conforme o original

Esta defesa de Dissertação de Mestrado Acadêmico é parte dos requisitos necessários à obtenção do título de Mestre. O competente diploma será expedido após cumprimento do estabelecido nas normas do Programa, legislação e regulamentação internas da UFU. As correções observadas pelos examinadores deverão ser realizadas no prazo máximo de 30 dias. Nada mais havendo a tratar, deu-se por encerrada a sessão às 16 horas e 30 minutos e lavrada a presente ata, que após lida e aprovada, foi assinada pela Banca Examinadora.

Prof. Dr. Frederico Garcia Pinto

UFV

Prof. Dr. Hélder Eterno da Silveira

UFU

Waldomiro Borges Neto
Orientador(a): Prof. Dr. Waldomiro Borges Neto – IQUFU

Dedico este trabalho aos meus pais Domingos Viagem Máquina (in memoriam) e Betia Domingos Araújo Máquina.

À minha esposa Agnes Mariza Paula Cordeiro Máquina.

Às minhas filhas Cynara Domingas Ademir Máquina e Yunaika Paula Ademir Máquina.

AGRADECIMENTOS

Agradeço:

À Deus Misericordioso, que em todos os dias da minha vida me deu força interior para superar as dificuldades e nunca desistir daquilo que acredito.

À minha família, que soube viver sem a minha presença física durante o período de formação.

Ao meu orientador, Prof. Dr. Waldomiro Borges Neto, pelo devido acompanhamento acadêmico e não só, mas como também pelo convívio, apoio, compreensão e, principalmente, pela confiança depositada em meu trabalho.

Aos colegas do Laboratório de Quimiometria do Triângulo: Sarmiento Junior Mazivila, Lucas Caixeta Gontijo, Baltazar Vasco Siteo, Hery Mitsutake, Letícia Maria de Souza, Eloíza Guimarães, Felipe Bachion de Santana, José Eduardo Buiatte e Edvando Souza Teles que, tanto me ajudaram durante os trabalhos acadêmicos.

Ao Programa do Convênio do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil/Ministério da Ciência e Tecnologia de Moçambique (Convênio CNPq/MCT-Mz), processo número 190279/2014-1, pelo suporte financeiro da bolsa.

À Universidade Federal de Uberlândia e aos seus funcionários que forneceram toda a estrutura física e tecnológica para a realização deste trabalho, principalmente à secretária do programa de pós-graduação em química, Mayta Mamede Negroto Peixoto, pelo acompanhamento em todos os processos técnico-administrativos.

À Universidade Pedagógica de Moçambique, dirigida pelo Reitor Prof. Dr. Rogério José Uthui, pelo afastamento para a realização do meu mestrado.

À TRANSPETRO S.A. (Uberlândia, Estado de Minas Gerais, Brasil) pelo fornecimento do diesel utilizado para a preparação das amostras.

Ao Prof. Dr. António Inácio Comando Suluda e ao MSc. Geraldo Alfredo Gueze, docentes e colegas da Universidade Pedagógica de Moçambique pelas cartas de recomendação fornecidas para a candidatura da bolsa de estudo.

A todos os meus colegas de serviço e alunos pela convivência profissional e apoio moral durante os meus estudos.

Finalmente, ressalto que, mencionar nomes de todos os que contribuíram para fazer deste trabalho uma realidade é uma tarefa difícil. Portanto, agradeço aos demais que não os mencionei aqui, mas que direta ou indiretamente contribuíram para o sucesso deste trabalho.

RESUMO

O biodiesel é um combustível renovável derivado de óleos vegetais ou gorduras animais que tem se destacado nas agendas dos Governos de Moçambique, Brasil e de vários países como opção energética e contraponto a dependência de combustíveis fósseis. Este combustível foi introduzido nas matrizes energéticas moçambicana e brasileira nos anos de 2009 e 2005, respectivamente, onde é comercializado em mistura com o diesel na proporção estabelecida nas leis dos dois países.

Diante do exposto, faz-se necessário o desenvolvimento de metodologias que forneçam respostas rápidas e eficientes para o controle da qualidade deste combustível. Assim sendo, o presente trabalho apresenta metodologias para classificar e prever o teor dos biodieseis metílicos de Mafurra, Moringa e Algodão em mistura com o diesel, utilizando Espectrometria no Infravermelho Médio aliada aos métodos quimiométricos de calibração multivariada por Quadrados Mínimos Parciais (PLS) e Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA).

Os modelos PLS desenvolvidos para a determinação do teor foram validados com base nas seguintes figuras de mérito: seletividade, sensibilidade, sensibilidade analítica, limite de detecção, limite de quantificação, e testes para erro sistemático (bias e t_{bias}). Também se fez a avaliação do ajuste dos modelos por meio da correlação dos valores reais e previstos dos conjuntos de calibração e previsão, onde se notou uma alta correlação existente entre ambos, com um coeficiente de correlação superior que 0,99 e com resultados satisfatórios para as figuras de mérito avaliadas.

Para o monitoramento qualitativo, foram desenvolvidos os modelos PLS-DA. A eficiência destes modelos foi avaliada com base nos critérios estatísticos de respostas verdadeiras, isto é, parâmetros de sensibilidade e especificidade. Estes parâmetros apresentaram valores de 1, o que representa 100% de classificação correta das amostras de calibração e previsão em todos os modelos.

Os bons resultados dos modelos comprovam que de fato, estas metodologias analíticas são viáveis e eficientes para o controle quantitativo e qualitativo destes combustíveis.

Palavras-Chave: Biodiesel. FT-MIR. Controle de Qualidade. Quimiometria.

ABSTRACT

Biodiesel is a renewable fuel derived from vegetable oils or animal fats that have been highlighted in the agendas of Governments of Mozambique, Brazil and several countries such as energy option and dependence counterpoint on fossil fuels. This fuel was introduced in the mozambican and brazilian energy matrix in 2009 and 2005, respectively, where it is marketed in combination with diesel in the proportion established in the laws of both countries.

Given the above, it is necessary to develop methodologies that provide rapid and efficient responses to control the quality of this fuel. However, this study presents methodologies to quantify and classify the content of methyl biodiesels of Mafurra, Moringa and cotton mixed with diesel, using Mid-Infrared Spectroscopy associated with chemometric methods of multivariate calibration by Partial Least Squares (PLS) and Partial Least Squares – Discriminant Analysis (PLS-DA).

The PLS models developed for the determination of the content has been validated based on the following figures of merit: selectivity, sensitivity, analytical sensitivity, detection limit, quantification limit, test to systematic error (bias e t_{bias}). Was also made the assessment of the adjustment of the models through the correlation of actual and forecast values of the sets of calibration and prediction, where were noted a high correlation between the two, with a correlation coefficient higher than 0, 99 and satisfactory results for the evaluated figures of merit.

For the qualitative monitoring, PLS-DA models were developed. The efficiency of these models was analyzed based on the criteria of true answers statistics, this is, sensitivity and specificity parameters. These parameters showed values of 1, which signify 100% correct classification of the calibration samples and forecast in all models.

The good results of the models show that in fact, these analytical methods are feasible and effective for the quantitative and qualitative control of these fuels.

Keywords: Biodiesel. FT-MIR. Quality Control. Chemometrics.

LISTA DE FIGURAS

Figura 1 –	Matérias-primas utilizadas na produção de biodiesel no Brasil.....	24
Figura 2 –	Árvores, frutos e sementes de Mafurra.....	24
Figura 3 –	Árvore, frutos e sementes de Moringa.....	25
Figura 4 –	Plantas e sementes de Algodão.....	26
Figura 5 –	Mecanismo da Reação de Hidrólise de triacilgliceróis na presença de umidade.....	27
Figura 6 –	Mecanismo da formação de alcóxido.....	27
Figura 7 –	Mecanismos de saponificação de triglicerídeos.....	28
Figura 8 –	Mecanismo de reação de esterificação via catálise ácida.....	29
Figura 9 –	Mecanismo de reação de transesterificação via catálise alcalina.....	29
Figura 10 –	Índice de não conformidade com relação a natureza do diesel.....	30
Figura 11 –	Processo de reflexão no ATR horizontal.....	34
Figura 12 –	Cristal de ZnSe do acessório ATR.....	35
Figura 13 –	Organização dos dados para a construção da matriz X	37
Figura 14 –	Construção do modelo de regressão por Quadrados Mínimos Parciais.....	38
Figura 15 –	Validação do modelo de regressão por Quadrados Mínimos Parciais.....	38
Figura 16 –	Gráfico de <i>leverage</i> versus <i>Q</i> Residuals.....	40
Figura 17 –	Representação da matriz de dados X e o vetor y usados no PLS-DA.....	45
Figura 18 –	Representação gráfica do resultado do modelo PLS-DA.....	47
Figura 19 –	Etapas do processo de esterificação.....	49
Figura 20 –	Etapas do processo de transesterificação.....	50
Figura 21 –	Espectros MIR de: a) óleos de Mafurra, Moringa e Algodão; b) biodieseis metílico de Mafurra, Moringa e Algodão.....	54
Figura 22 –	Espectros MIR de 120 amostras de misturas B7 de BMMA, BMMo e BMA...	55
Figura 23 –	Espectros MIR das amostras dos conjuntos de calibração e previsão dos biodieseis em mistura com diesel: (a) BMMA; (b) BMMo.....	57
Figura 24 –	Espectros MIR das amostras dos conjuntos de calibração e previsão do BMA em mistura com diesel.....	58
Figura 25 –	Gráfico de <i>leverage</i> versus <i>Q</i> residuals com 95% de confiança para o modelo PLS de BMMA.....	59
Figura 26 –	Gráfico de <i>leverage</i> versus <i>Q</i> residuals com 95% de confiança para o modelo PLS de BMMo.....	59

Figura 27 –	Gráfico de <i>leverage</i> versus <i>Q residuals</i> com 95% de confiança para o modelo PLS de BMA.....	60
Figura 28 –	(a) Gráfico de ajuste do modelo PLS de BMMA e (b) gráfico de resíduo do modelo PLS de BMMA.....	62
Figura 29 –	(a) Gráfico de ajuste do modelo PLS de BMMo e (b) gráfico de resíduo do modelo PLS de BMMo.....	63
Figura 30 –	(a) Gráfico de ajuste do modelo PLS de BMA e (b) gráfico de resíduo do modelo PLS de BMA.....	64
Figura 31 –	Estimativas do modelo PLS-DA para a classificação de amostras de B7 e BX do BMMA.....	66
Figura 32 –	Estimativas do modelo PLS-DA para a classificação de amostras de B7 e BX do BMMo.....	66
Figura 33 –	Estimativas do modelo PLS-DA para a classificação de amostras de B7 e BX do BMA.....	67
Figura 34 –	Gráfico dos pesos de LV1, LV2, LV3, LV4, LV5 e LV6 <i>versus</i> variáveis do modelo PLS-DA de BMMA.....	68
Figura 35 –	Gráfico dos pesos de LV1, LV2, LV3, LV4, LV5 e LV6 <i>versus</i> variáveis do modelo PLS-DA de BMMo.....	68
Figura 36 –	Gráfico dos pesos de LV1, LV2, LV3, LV4 e LV5 <i>versus</i> variáveis do modelo PLS-DA de BMA.....	69
Figura 37 –	Gráfico de sensibilidade <i>versus</i> especificidade do modelo PLS-DA de BMMA para amostras de: (a) B7 e (b) BX.....	70
Figura 38 –	Gráfico de sensibilidade <i>versus</i> especificidade do modelo PLS-DA de BMMo para amostras de: (a) B7 e (b) BX.....	71
Figura 39 –	Gráfico de sensibilidade <i>versus</i> especificidade do modelo PLS-DA de BMA para amostras de: a) B7 e b) BX.....	72
Figura 40 –	Estimativas do modelo PLS2-DA para a classificação das três classes de B7 de BMMA, BMMo e BMA.....	74
Figura 41 –	Gráfico dos pesos de LV1, LV2 e LV3 <i>versus</i> variáveis do modelo PLS2-DA.....	75
Figura 42 –	VIP scores <i>versus</i> variáveis.....	76

LISTA DE TABELAS

Tabela 1 –	Equações usadas para o cálculo das figuras de mérito.....	43
Tabela 2 –	Parâmetros e valores de exatidão para os modelos PLS.....	56
Tabela 3 –	Resultados das figuras de mérito para o modelo PLS.....	61
Tabela 4 –	Parâmetros de classificação obtidos pelo modelo PLS-DA.....	65
Tabela 5 –	Parâmetros de classificação obtidos pelo modelo PLS2-DA.....	73

LISTA DE ABREVIATURAS E SIGLAS

ABNT NBR	Norma Brasileira da Associação Brasileira de Normas Técnicas
AGL	Ácidos Graxos Livres
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ASTM	Sociedade Americana de Testes e Materiais (do inglês, <i>American Society of Testing and Materials</i>)
ATR	Refletância Total Atenuada (do inglês, <i>Attenuated Total Reflectance</i>)
B100	Biodiesel Puro
B7	Mistura Biodiesel/Diesel com 7% de Biodiesel
BX	Mistura Biodiesel/Diesel com X% de Biodiesel
BMA	Biodiesel Metílico de Algodão
BMMa	Biodiesel Metílico de Mafurra
BMMo	Biodiesel Metílico de Moringa
CNPE	Conselho Nacional da Política Energética
EN	Norma Europeia (do inglês, <i>European Norm</i>)
FT–MIR	Infravermelho Médio com Transformada de Fourier (do inglês, <i>Fourier Transform Mid Infrared</i>)
GC–MS	Cromatografia Gasosa – Espectrometria de Massas (do inglês, <i>Gas Chromatography – Mass Spectrometry</i>)
LC–MS	Cromatografia Líquida – Espectrometria de Massas (do inglês, <i>Liquid Chromatography – Mass Spectrometry</i>)
MIR	Infravermelho Médio (do inglês, <i>Mid Infrared</i>)
NAS	Sinal Analítico Líquido (do inglês, <i>Net Analyte Signal</i>)
NIR	Infravermelho Próximo (do inglês, <i>Near Infrared</i>)
PCA	Análise de Componentes Principais (do inglês, <i>Principal Components Analysis</i>)
PLS	Quadrados Mínimos Parciais (do inglês, <i>Partial Least Squares</i>)
PLS–DA	Análise Discriminante por Quadrados Mínimos Parciais (do inglês, <i>Partial Least Squares – Discriminant Analysis</i>)
PNPB	Programa Nacional de Produção e Uso do Biodiesel
RMN	Ressonância Magnética Nuclear (do inglês, <i>Nuclear Magnetic Resonance</i>)

RMSEC	Erro Quadrático Médio de Calibração (do inglês, <i>Root Mean Square Error of Calibration</i>)
RMSECV	Erro Quadrático Médio de Validação Cruzada (do inglês, <i>Root Mean Square Error of Cross-Validation</i>)
RMSEP	Erro Quadrático Médio de Previsão (do inglês, <i>Root Mean Square Error of Prediction</i>)
SADC	Comunidade para o Desenvolvimento da África Austral (do inglês, Southern Africa Development Community)
SDV	Desvio Padrão da Validação (do inglês, Standard Deviation of Validation)
TAG	Triacilglicerol
VIP	Variável Importante na Projeção
VL	Variável Latente

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 Objetivos.....	19
1.1.1 <i>Objetivo Geral.....</i>	19
1.1.2 <i>Objetivos Específicos.....</i>	19
2 BIODIESEL.....	20
2.1 Histórico.....	20
2.2 Panorama moçambicano.....	21
2.3 Panorama brasileiro.....	22
2.4 Matérias-primas.....	23
2.5 Produção de biodiesel.....	26
2.6 Índice de não conformidades (teor e normas)	30
3 ESPECTROMETRIA NO INFRAVERMELHO.....	33
3.1 Espectrometria no Infravermelho Médio.....	33
3.1.1 <i>Refletância Total Atenuada.....</i>	34
4 MÉTODOS QUIMIOMÉTRICOS.....	36
4.1 Organização de dados.....	36
4.2 Regressão por Quadrados Mínimos Parciais (PLS).....	37
4.2.1 <i>Deteção de Amostras Anômalas (Outlier).....</i>	39
4.2.2 <i>Validação do modelo PLS.....</i>	40
4.3 Métodos de classificação.....	43
4.3.1 <i>Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA).....</i>	44
4.3.1.1 <i>Parâmetros de Avaliação do Modelo PLS-DA.....</i>	46
4.3.1.2 <i>Representação gráfica do resultado do modelo PLS-DA.....</i>	47
5 PROCEDIMENTO EXPERIMENTAL.....	48
5.1 Produção do Biodiesel.....	48
5.2 Preparação das Amostras.....	50
5.3 Obtenção dos Espectros MIR.....	51
5.4. Construção dos modelos.....	51
6 RESULTADOS E DISCUSSÃO.....	53
6.1. Caracterização dos Espectros MIR.....	53
6.2. Modelos PLS.....	55

6.3. Modelos PLS-DA.....	65
6.4. Modelo PLS2-DA.....	73
7 CONCLUSÕES.....	77
REFERÊNCIAS.....	78

1. INTRODUÇÃO

À medida que a população mundial cresce, aumenta a demanda por combustíveis fósseis, no entanto, esses combustíveis constituem recursos não renováveis e causam preocupações ambientais. Assim, por serem recursos esgotáveis, surgem necessidades de se potenciar o uso de fontes de energia renováveis (VERRASTRO; LADISLAW, 2007; ACHTEN et al., 2008).

Portanto, na busca de fontes alternativas de energia renovável, os biocombustíveis têm se destacado nas agendas dos Governos de Moçambique, Brasil e de vários outros países como alternativa a dependência por combustíveis fósseis e a busca para reduzir impactos ambientais na produção e combustão desses combustíveis (MOÇAMBIQUE, 2009; GERMANA e SCHONEVELD, 2012), isso é devido a algumas vantagens, tais como: ser biodegradável, livre de enxofre e de compostos aromáticos, conter alto número de cetano, teor médio de oxigênio, maior ponto de fulgor e menor índice de emissão de gases no processo produtivo. Outro fator de grande importância consiste na geração de oportunidades de emprego e renda para a população rural, principalmente os incentivados pelo governo como a agricultura familiar (FERRARI et al., 2005).

Moçambique, país localizado no sudeste do Continente Africano, banhado pelo oceano Índico a leste e que faz fronteira com a Tanzânia ao norte; Malawi e Zâmbia a noroeste; Zimbabwe a oeste; Suazilândia e África do Sul a sudoeste (CUMBE, 2007), possui uma abundância relativa de terras com condições agroclimáticas favoráveis para o cultivo de várias matérias-primas usadas para a produção de biocombustíveis, tornando-se desta forma, um potencial produtor de biomassa (BATIDZIRAI et al., 2006; ARNDT et al., 2008). As semelhanças agroclimáticas, geopolítica e geoestratégica existente entre Moçambique e Brasil fizeram com que os dois países firmassem um acordo de cooperação na área de biocombustíveis, assinando um memorando de entendimento em 2007, com os seguintes objetivos (BAMBO, 2014):

- Promover trocas de missões técnico-empresariais, em datas de conveniência mútua;
- Apoiar o desenvolvimento de programa de produção e uso de biocombustíveis em Moçambique;
- Elaborar, conjuntamente, projetos de cooperação técnica para capacitar recursos humanos do setor de biocombustíveis em Moçambique;

- Identificar parcerias com outros países e organismos internacionais interessados em apoiar projetos específicos para a implantação do programa de desenvolvimento dos biocombustíveis em Moçambique.

Um ano após a assinatura do memorando, Moçambique registrou solicitações de cerca de 12 milhões de hectares de terra para o cultivo das matérias primas usadas para a produção de biocombustíveis (ARNDT et al., 2008).

Em 2009, foi aprovada a política de biocombustíveis em Moçambique, destacando o compromisso do governo em explorar oportunidades de desenvolvimento econômico para as comunidades rurais oferecido pela crescente demanda mundial de biocombustíveis (MOÇAMBIQUE, 2009). No entanto, a produção de biocombustíveis em Moçambique ainda não é significativa, pois usa como matéria-prima principal o óleo de Pinhão Manso, cuja produção é insuficiente para suprir a demanda existente (AVINASH et al., 2014).

Ao contrário de Moçambique, a produção de biodiesel no Brasil cresceu, tendo atingido 3.419.838 m³ em 2014 contra 2.917.488 m³ do ano anterior, aumentando desta forma em 17,2% a quantidade de biodiesel disponibilizada no mercado interno (BRASIL, 2015), onde é obrigatoriamente comercializado em mistura com o diesel na proporção de 7% de biodiesel para 93% de diesel puro, formando um combustível denominado de B7 (BRASIL, 2014). A principal matéria-prima usada para a produção de biodiesel no Brasil é o óleo de soja (69,2%), seguido do sebo bovino (17,0%) (BRASIL, 2015). O uso do óleo de soja para a produção de biodiesel porém estabelece uma concorrência ao seu uso como alimento o que pode acarretar aumentos nos custos de produção.

Diante do exposto, faz-se necessário o desenvolvimento de pesquisas visando o melhoramento da matriz energética, principalmente no que se refere a estudos de fontes alternativas para produção de biocombustíveis, que não possam competir com culturas alimentares em áreas com insegurança alimentar (MAKKAR; BECKER, 2009). Portanto, para evitar tal competição, culturas energéticas devem ser não comestíveis e adaptáveis em terras marginais (SENGO et al., 2010; BANKOVIĆ-ILIĆ et al., 2012).

A Mafurra (*Trichilia emetica*), Moringa (*Moringa oleifera*) e Algodão (*Gossypium hirsutum* L.) são culturas que de uma forma geral não são usadas como alimentos. Assim sendo, estas plantas constituem alternativas viáveis que se pode potencializar para produção do biodiesel, pois contêm teores adequados de óleo de 55-65%, 35-40% e 20%, nas suas sementes (GRUNDY; CAMPBELL, 1993; CARDELLO et al., 1995; OLIVEIRA et al., 2012).

Ressalta-se, que na medida em que a produção do biodiesel cresce, aumenta também a possibilidade da sua comercialização com teor fora do exigido pela norma. Este cenário é comprovado com os dados estatísticos de monitoramento da qualidade dos combustíveis líquidos automotivos do mês de Julho de 2016, onde mostram claramente que o teor de biodiesel representou 55,8% das não conformidades observadas para o diesel comercializado (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP, 2016a). Neste contexto, torna-se importante o desenvolvimento de metodologias que forneçam respostas rápidas e eficientes para monitorar quantitativamente e qualitativamente o teor do biodiesel metílico de Mafurra (BMMa), Moringa (BMMo) e Algodão (BMA) em mistura com o diesel. Assim sendo, muitos estudos relatam a aplicação de métodos instrumentais para o Controle de Qualidade de misturas de biodiesel/diesel, usando Espectrometria no Infravermelho Médio (MIR) (PIMENTEL et al., 2006; SITOIE et al., 2016).

Importa salientar que para agilizar as tomadas de decisões em tempo hábil tanto na linha de produção quanto no controle de qualidade do produto, é necessário e desejável que se desenvolvam métodos analíticos rápidos, de baixo custo e com menor impacto ambiental. A Espectrometria no Infravermelho Médio é uma das técnicas vibracionais mais utilizadas em análises químicas visando atender tais requisitos, pois principalmente com a amostragem sendo realizada com acessórios de Reflectância, torna possível a análise direta, não destrutiva, com pequeno volume de amostra, rápida e, no caso de alguns equipamentos a análise *in situ* (BORIN; POPPI, 2004; PIMENTEL et al., 2006; KAROUI et al., 2010).

Na literatura pode-se encontrar relatos sobre a aplicação frequente da Espectrometria no Infravermelho Médio associada aos métodos quimiométricos, como metodologia alternativa as análises quantitativas e qualitativas convencionais (BUIATTE et al., 2016; SANTANA et al., 2016). No caso das análises quantitativas, o método de calibração multivariado mais utilizado é a regressão por Quadrados Mínimos Parciais (PLS), como por exemplo, Souza et al. (2014) desenvolveram metodologias para a quantificação de óleo lubrificante automotivo residual como adulterante em diesel S-10 brasileiro usando Espectrometria MIR e PLS, tendo obtido valores de erros quadráticos médios de calibração e previsão relativamente baixos e de acordo com a norma ABNT NBR 15568. Em análises qualitativas, são usados comumente os métodos quimiométricos da Análise de Componentes Principais (PCA) e a Análise de Agrupamentos Hierárquicos (HCA), que por se tratarem de métodos não supervisionados, não podem usar a informação do modelo construído para previsão de novas amostras. Baseada no PLS, a Análise Discriminante por Quadrados

Mínimos Parciais (PLS-DA), trata-se de um método supervisionado capaz de prever a qual classe pertence novas amostras, como no trabalho de Mazivila et. al. (2015) que realizaram o controle de qualidade de biodiesel de soja e pinhão manso classificando as amostras de biodiesel conforme sua origem e possíveis processos de adulterações.

A presente dissertação está dividida em cinco capítulos, a saber:

O primeiro capítulo, Biodiesel, apresenta o histórico do biodiesel nos panoramas moçambicano e brasileiro; as matérias-primas utilizadas para produção de biodiesel em Moçambique e no Brasil; a produção do biodiesel e índice de não conformidade com relação à natureza do diesel.

O segundo capítulo, Espectrometria no Infravermelho, descreve a região espectral do infravermelho médio, incluindo as vantagens do uso do espectrômetro com Transformada de Fourier e da Refletância Total Atenuada para o controle de qualidade de biocombustíveis.

O terceiro capítulo, Métodos Quimiométricos, apresenta o princípio geral de análise multivariada, incluído as áreas de aplicação da Quimiometria; a organização dos dados e os métodos de calibração e de classificação (reconhecimento de padrões).

O quarto capítulo, Procedimento Experimental, descreve o processo de produção de biodiesel, preparo de amostras dos conjuntos de calibração e previsão, aquisição dos respectivos dados espectrais na Espectrometria no Infravermelho Médio (MIR) e uma breve descrição da construção dos modelos PLS e PLS-DA.

O quinto capítulo, Resultados e Discussão, apresentam e analisam os resultados obtidos em cada um dos modelos PLS e PLS-DA de BMMA, BMMo e BMA.

As referências usadas para a produção deste trabalho são apresentadas no final da dissertação.

1.1. Objetivos

1.1.1. *Objetivo Geral*

- Desenvolver metodologias analíticas para classificar e prever o teor de biodieseis metílicos de Mafurra, Moringa e Algodão em mistura com diesel usando a técnica de Espectrometria no Infravermelho Médio aliada aos métodos quimiométricos PLS-DA e PLS.

1.1.2. *Objetivos Específicos*

- Construir modelos de PLS-DA e PLS para classificar e prever o teor de biodieseis metílicos de Mafurra, Moringa e Algodão em mistura com diesel B7, respectivamente;
- Classificar tipos de biodieseis presentes na mistura biodiesel/diesel (B7) de acordo com o tipo de matéria-prima usada (óleo), usando modelos PLS-DA;
- Validar os modelos PLS construídos de acordo com as normas ASTM E1655 e ABNT NBR 15568;
- Validar os modelos PLS-DA construídos, através dos valores de sensibilidade e especificidade.

2. BIODIESEL

2.1. Histórico

A utilização de matérias-primas não derivadas de petróleo como combustível não é recente (NITSKE; WILSON, 1965). A companhia francesa Otto apresentou um motor que funcionou com óleo de amendoim, durante a Exposição Universal de 1900, em Paris. Este motor tinha sido construído para operar a óleo mineral e, no entanto, funcionou com óleo de amendoim sem que este tivesse sofrido qualquer tipo de alteração (DIESEL, 1912). A ideia de se utilizar óleo de amendoim partiu do Governo Francês, uma vez que esta matéria-prima abundava nas suas colônias africanas e, assim, poderiam ser fornecidas como energia da sua própria indústria, sem que fosse necessário importar combustíveis fósseis (SIMÕES, 2013). Mais tarde, Rudolf Diesel, inventor do motor a diesel, revelou que seus motores também funcionaram bem experimentalmente com outros óleos orgânicos, como o óleo de mamona e gorduras animais. Após a revelação do Rudolf Diesel, vários países europeus com colônias na África, tais como: França, Itália, Reino Unido e Bélgica, tiveram diferentes interesses em combustíveis de óleo vegetal, mas a questão do baixo preço internacional do barril de petróleo na época firmou a preferência desta fonte de energia e seus derivados para a alimentação dos motores a combustão (KNOTHE et al., 2006).

No entanto, durante a Segunda Guerra Mundial (1939-1945) as trocas comerciais foram interrompidas e o consumo de derivados de petróleo passou a ser descontrolado e, conseqüentemente, a acessibilidade do petróleo voltou a escassear (RODRIGUES, 2007). Este fato, fez com que vários países utilizassem os óleos vegetais como combustíveis de emergência. Por exemplo, nos Estados Unidos começou-se a desenvolver projetos sobre o uso de misturas binárias (bicomcombustíveis), com principal foco para óleos de algodão, milho e misturas destes com o óleo diesel convencional; a China produziu a gasolina e querosene por processos de craqueamento a partir dos óleos de tungue e de outras matérias-primas oleaginosas; a Índia expandiu suas pesquisas a dez novos tipos de óleos vegetais para o desenvolvimento de combustíveis domésticos; o Japão utilizou óleo de soja refinado como óleo combustível para os motores de navio de guerra; o Brasil proibiu a exportação de óleo de algodão para que este substituísse o óleo diesel importado (CHOWHURY et al., 1942; CHANG; WAN, 1947; KNOTHE et al., 2006).

Embora os óleos vegetais tenham sido considerados como uma alternativa ao óleo diesel, a sua utilização apresenta alguma limitação relacionada com a sua elevada viscosidade

cinética. No entanto, a alta viscosidade dos óleos vegetais conduz a depósitos de carbono nos cilindros e nos injetores dificultando, desta forma, a queima desse combustível (KNOTHE; STEIDLEY, 2005). Para ultrapassar essa limitação, várias pesquisas foram desenvolvidas tal que se culminou no uso da transesterificação, que é o processo de conversão dos óleos vegetais e gorduras animais em um produto com viscosidade cinética similar a do óleo diesel, o alquil éster (RODRIGUES, 2007). Devido a esta e outras similaridades, os autores chineses denominaram o alquil éster de biodiesel (WANG, 1988). Com esse avanço, países como Bélgica, Estados Unidos, Áustria, Alemanha, Brasil, França dentre outros, começaram a produzir o biodiesel em grande escala para o consumo interno.

2.2. Panorama moçambicano

Em resposta aos condicionalismos impostos pelas vicissitudes do mercado internacional do petróleo, o Governo de Moçambique aprovou a Política e Estratégias de Biocombustíveis no dia 24 de Março de 2009, introduzindo desta forma os biocombustíveis na sua matriz energética, tendo em conta que os mesmos proporcionariam vários benefícios, a saber (MOÇAMBIQUE, 2009):

- A substituição gradual dos combustíveis fósseis, sem nenhuma ou com uma adaptação tecnológica mínima do equipamento, com impactos positivos na balança de pagamentos e na redução da dependência externa do país;
- A geração de emprego e auto-emprego na agricultura e no processamento agroindustrial, particularmente nas zonas rurais, gerando ocupação para a abundante, disciplinada e treinável mão de obra moçambicana;
- A possibilidade de produção de biocombustíveis seguindo um modelo de desenvolvimento agrícola apropriado à prevenção da predominância excessiva de monoculturas e que assegure, ao mesmo tempo, a produção de outras culturas;
- Melhor uso das infraestruturas existentes, incluído os portos das águas profundas, ferrovias, estradas, oleodutos e gasodutos;
- Melhor aproveitamento da água e recursos climáticos favoráveis para produzir biocombustíveis em grande escala para o consumo interno e exportação;
- Melhor aproveitamento das oportunidades oferecidas pelo mercado regional da Comunidade para o Desenvolvimento da África Austral (SADC).

A Política e Estratégias de Biocombustíveis aprovadas têm como foco a promoção do etanol e do biodiesel produzidos a partir de matérias primas agrícolas adequadas aos ambientes agroclimáticos variados do país, para serem usados principalmente nas áreas de transporte, bem como para outros fins energéticos.

Em 2011 foi publicado o Decreto nº 58/2011 que aprovou o Regulamento de Biocombustíveis e suas Misturas com combustíveis fósseis. Este regulamento estabelece que a partir de Janeiro de 2012 a mistura obrigatória de biodiesel com diesel a ser comercializado em território nacional será feita com as seguintes porcentagens (MOÇAMBIQUE, 2011):

- 3% de biodiesel, de Janeiro de 2012 até 31 de Dezembro de 2015;
- 7,5% de biodiesel, de Janeiro de 2016 até 31 de Dezembro de 2020;
- 10% de biodiesel, a partir de Janeiro de 2021.

2.3. Panorama brasileiro

Antes da introdução do biodiesel na matriz energética brasileira, a Presidência da República instituiu por meio de Decreto de 02 de Julho de 2003, um Grupo de Trabalho Interministerial encarregado de apresentar estudos sobre a viabilidade de utilização de biodiesel como fonte alternativa de energia (BRASIL, 2003a). Como resultado foi elaborado um relatório que deu embasamento ao Presidente da República para estabelecer o Programa Nacional de Produção e Uso do Biodiesel (PNPB) como ação estratégica e prioritária para o Brasil, com objetivo de implementar de forma sustentável, tanto técnica, como economicamente, a produção e uso do biodiesel, com enfoque na inclusão social e no desenvolvimento regional, via geração de emprego e renda (BRASIL, 2003b).

No entanto, a introdução deste biocombustível na matriz energética brasileira foi feita mediante a publicação da Lei nº 11.097, no Diário Oficial da União, em 2005 (BRASIL, 2005). A partir de 2008, a mistura de biodiesel puro (B100) ao óleo diesel passou a ser obrigatória, com as seguintes porcentagens (CNPE, 2008; CNPE, 2009; BRASIL, 2014):

- 2% de biodiesel, de Janeiro até Junho de 2008;
- 3% de biodiesel, de Julho de 2008 até Junho de 2009;
- 4% de biodiesel, de Julho até Dezembro de 2009;
- 5% de biodiesel, de Janeiro de 2010 até Junho de 2014;
- 6% de biodiesel, de Julho até Outubro de 2014.

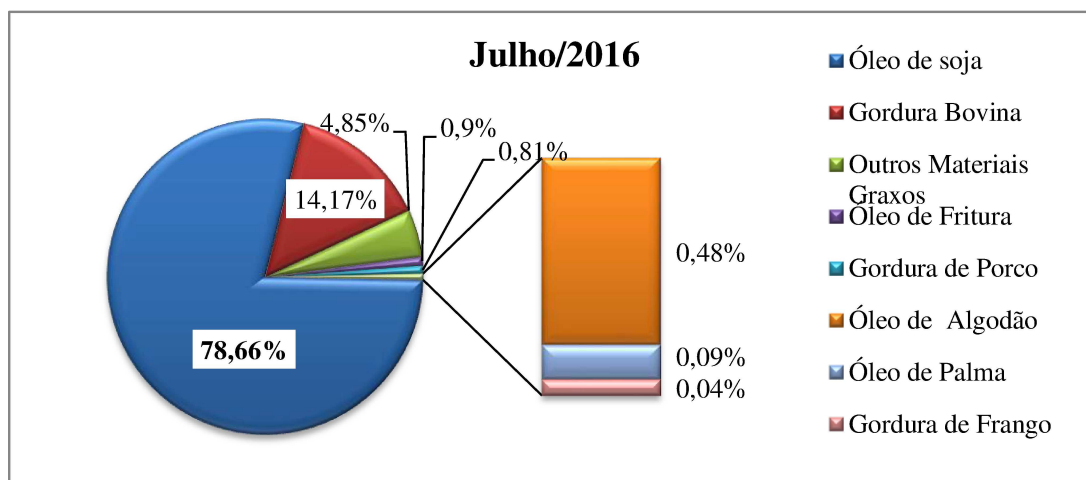
Essa percentagem foi crescendo tal que, a partir de Novembro de 2014 até hoje, a mistura passou a ser feita na seguinte proporção: 7% de biodiesel e 93% de diesel puro, formando um combustível denominado de B7 (BRASIL, 2014). Entretanto, a contínua elevação do percentual de adição de biodiesel ao diesel demonstra o sucesso do Programa Nacional de Produção e Uso do Biodiesel e da experiência acumulada pelo Brasil na produção e no uso em larga escala de biocombustíveis. Com isso, o Brasil se destaca entre os maiores produtores e consumidores de biodiesel do mundo, com uma produção anual, em 2014, de 3.419.838 m³ (BRASIL, 2015).

2.4. Matérias-primas

O biodiesel pode ser produzido a partir de muitas matérias-primas, a saber: óleos vegetais, gorduras animais, óleos de frituras residuais e, até mesmo matérias graxas de alta acidez (KNOTHE et al., 2006). No entanto, a maior parte do biodiesel atualmente produzido no mundo deriva dos óleos vegetais (GONTIJO, 2016). Assim sendo, fatores como a geografia, o clima e a economia determinam quais óleos vegetais apresentam maior potencial para produzir biodiesel num determinado local, como, por exemplo, nos Estados Unidos o óleo mais utilizado é o de soja; na Europa é óleo de colza; nos países tropicais a preferência recai sobre o óleo de palma (KNOTHE et al., 2006).

Moçambique dispõe de várias matérias-primas para produção de biodiesel: pinhão manso, coco, girassol, amendoim, soja, dentre outros, levando em conta aspectos como a sustentabilidade como matéria-prima, a avaliação dos seus rendimentos, custos de produção, impactos socioeconômicos e ambientais, atualmente são preferidas o pinhão manso e o coco. Devido aos problemas relacionados com o aparecimento da doença do amarelecimento letal do coqueiro e o seu alto preço de mercado, fazem com que atualmente a matéria-prima mais usada seja o pinhão manso. No entanto, segundo a Resolução nº 22/2009 de 21 de Maio, os produtores locais dessas matérias-primas, devem entregar sua produção exclusivamente aos titulares licenciados de produção, armazenagem e distribuição de biocombustíveis, para consequente introdução na cadeia produtiva do biodiesel.

No Brasil, as matérias-primas comuns, empregadas na produção de biodiesel são: óleos de soja, algodão, palma e de fritura, gordura bovina, gordura de porco e de frango. Dessas matérias-primas, a mais usada é o óleo de soja com 78,66%, seguido do sebo bovino (14,17%), como se pode ver na Figura 1 (ANP, 2016b).

Figura 1 – Matérias-primas utilizadas na produção de biodiesel no Brasil

Fonte: ANP (2016b).

A procura de novas matérias-primas, alternativas aos óleos vegetais alimentares, tem sido objeto de investigação nos últimos anos. Culturas como a Mafurra (*Trichilia emetica*), Moringa (*Moringa oleifera*) e Algodão (*Gossypium hirsutum* L.) que são muito produzidas nas regiões sul, centro e norte de Moçambique, constituem alvo de estudo neste trabalho por serem oleaginosas que de uma forma geral não são usadas como alimentos.

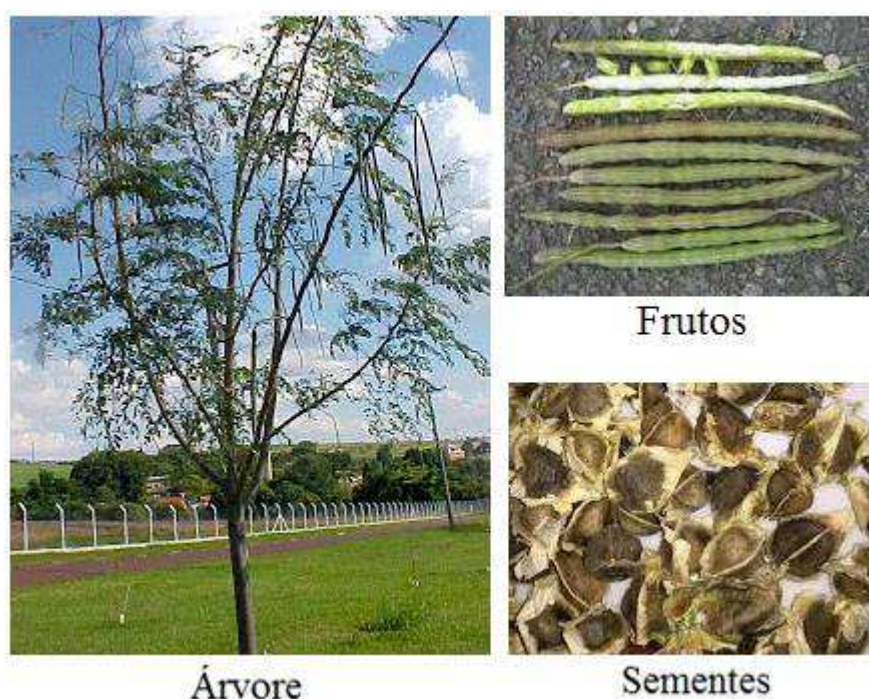
A árvore da Mafurra (Figura 2) é uma planta da família das meliáceas de folha perene, tem uma preferência por áreas com uma elevada precipitação pluviométrica e é abundante ao longo de áreas costeiras (KOMANE et al., 2011). As suas sementes contém um teor de óleo de 55-65% (GRUNDY; CAMPBELL, 1993).

Figura 2 – Árvores, frutos e sementes de Mafurra.

Fonte: O autor.

A árvore de Moringa (Figura 3) pertence a família Moringaceae, adapta-se bem às condições climáticas tanto do clima semiárido como subúmido, não sendo necessários altos volumes pluviométricos por ano, pode desenvolver com precipitações anuais a partir de 250 mm, é tolerante à seca e não é exigente em termos de fertilidade dos solos, desenvolvendo-se também em solos pobres (FOIDL et al., 2001). As suas sementes contém um teor de óleo de 35-40% (OLIVEIRA et al., 2012).

Figura 3 – Árvore, frutos e sementes de Moringa.



Fonte: O autor.

O algodoeiro (Figura 4) pertence a família Malvaceae, é uma planta de clima tropical cujo plantio não é recomendado em altitudes superiores a 1500 m, a média mensal de temperatura deve permanecer entre 18 e 30 °C, nunca ultrapassando o limite inferior de 14 °C e superior a 40 °C, necessita de precipitações pluviométricas em torno de 700 a 1300 mm anuais. Os solos para plantio do algodoeiro devem ser profundos, de textura média, ricos em matéria orgânica, permeáveis, bem drenados e de boa fertilidade (RICHETTI et al., 2003). As suas sementes contém um teor de óleo em média de 20% (CARDELLO et al., 1995).

Figura 4 – Plantas e sementes de Algodão.



Plantas



Sementes

Fonte: O autor.

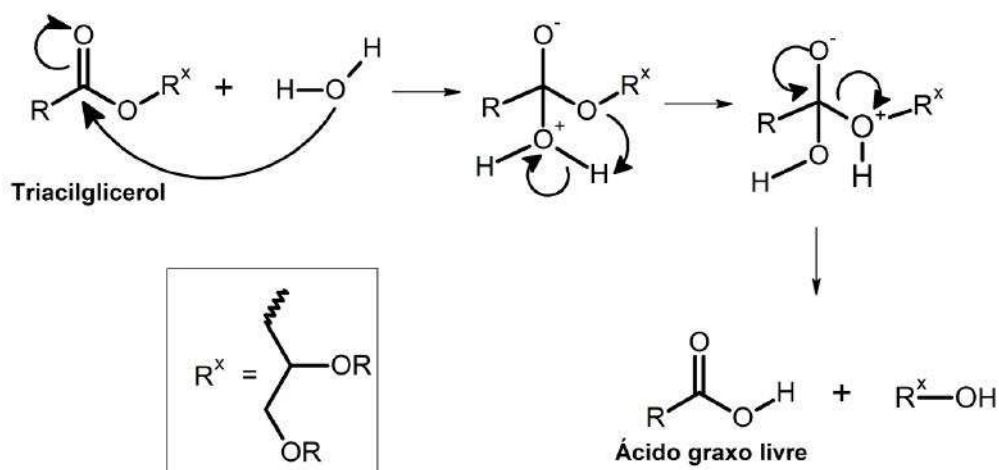
2.5. Produção de biodiesel

Para que o biodiesel seja produzido, os óleos vegetais e gorduras animais são submetidos a uma reação química denominada de transesterificação. Nesta reação, óleos vegetais e gorduras de origem animal reagem na presença de um catalisador (ácido ou básico) com um álcool de cadeia curta, geralmente metanol ou etanol (KNOTHE et al., 2006; BRASIL, 2012). No entanto, os catalisadores alcalinos (hidróxido de sódio e de potássio; ou os alcóxidos correspondentes) proporcionam processos muito mais rápidos que catalisadores ácidos (LAM et al., 2010).

Além do tipo de catalisador (alcalino ou ácido), outros parâmetros de reação como a razão molar entre o álcool e o óleo vegetal, a temperatura, o tempo de reação, o grau de refino do óleo vegetal, o efeito da presença de umidade e ácidos graxos livres, devem ser levados em conta no processo de transesterificação alcalina (FREEDMAN et al., 1984). Assim, para se obter rendimentos máximos no processo de transesterificação, o álcool deve ser livre de umidade e o conteúdo de ácidos graxos livres (AGL) do óleo vegetal deve ser menor que 0,5% (FREEDMAN et al., 1984). A ausência de umidade torna-se importante para evitar a ocorrência da hidrólise dos ésteres alquílicos sintetizados a ácidos graxos livres. Da mesma

forma, a reação de triacilgliceróis (TAG) com água pode formar ácidos graxos livres. A reação de formação dos ácidos graxos livres é mostrada na Figura 5 (KNOTHE et al., 2006).

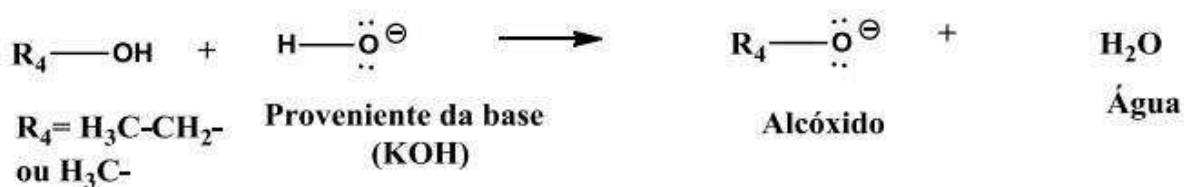
Figura 5 – Mecanismo da reação de hidrólise de triacilgliceróis na presença de umidade.



Fonte: Clayden et al., (2011)

Em temperaturas $\geq 60^\circ\text{C}$, empregando óleos vegetais refinados em razões molares álcool:óleo de pelo menos 6:1, a reação pode se completar em 1 h, fornecendo ésteres metílicos ou etílicos. Soluções de alcóxidos (Figura 6) com o álcool correspondente apresentam vantagens sobre os hidróxidos porque não permitem ocorrer a reação de formação de água, assegurando que o processo de transesterificação permaneça livre de água tanto quanto possível (LAM et al., 2010).

Figura 6 – Mecanismo da formação de alcóxido.

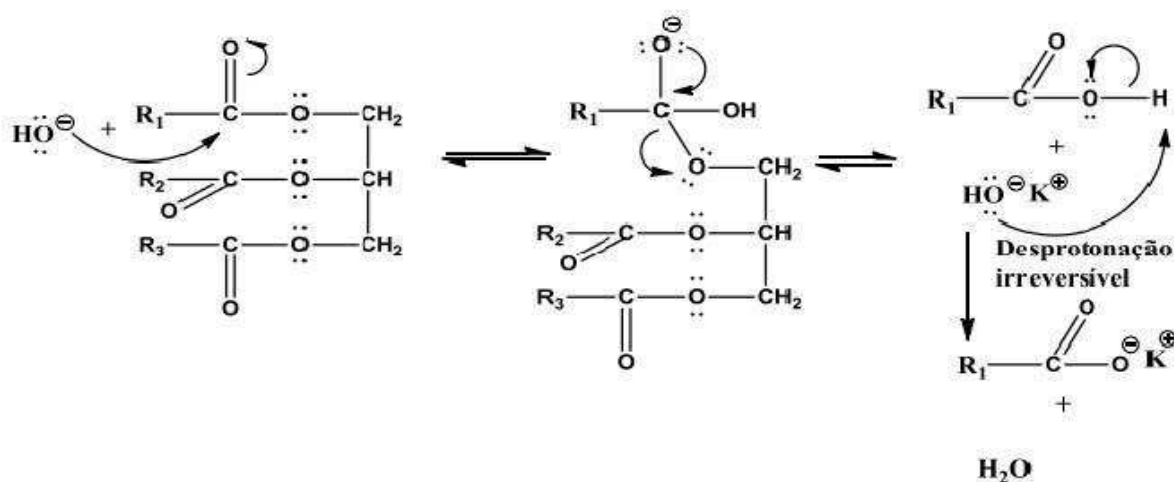


Fonte: Clayden et al., (2011)

Os óleos brutos também podem ser transesterificados pela catálise básica direta, mas os rendimentos de reação são geralmente reduzidos, devido à presença de impurezas e de

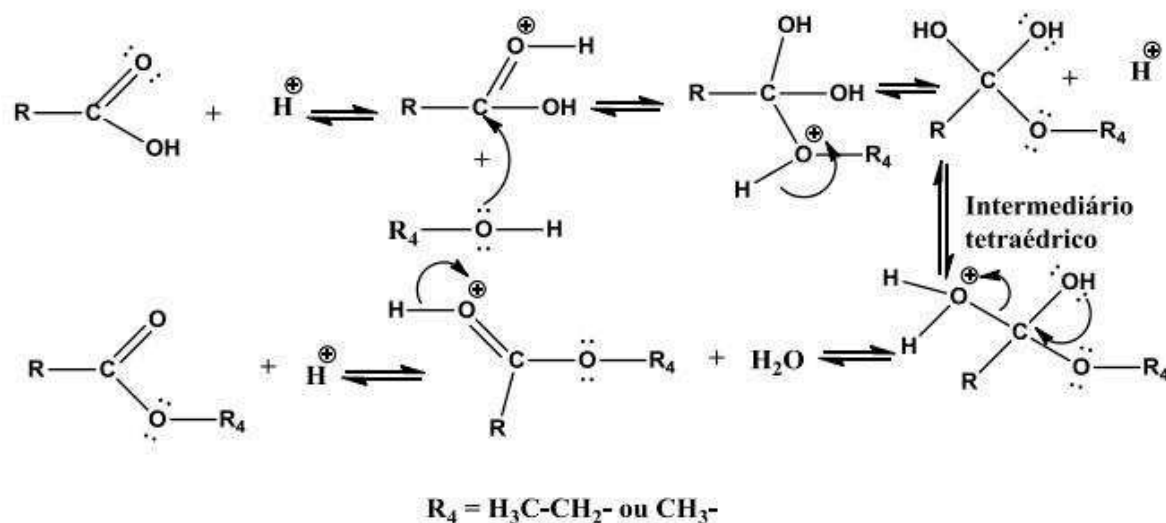
quantidades significativas de ácidos graxos livres (FREEDMAN et al., 1984). Ora, quando um catalisador alcalino é adicionado às matérias-primas com quantidades significativas de ácidos graxos livres, ocorre formação de sabões e água, devido à reação dos ácidos graxos livres com o catalisador (KNOTHE, 2007), como demonstra o mecanismo da reação apresentado na Figura 7.

Figura 7 – Mecanismos de saponificação de triglicerídeos.



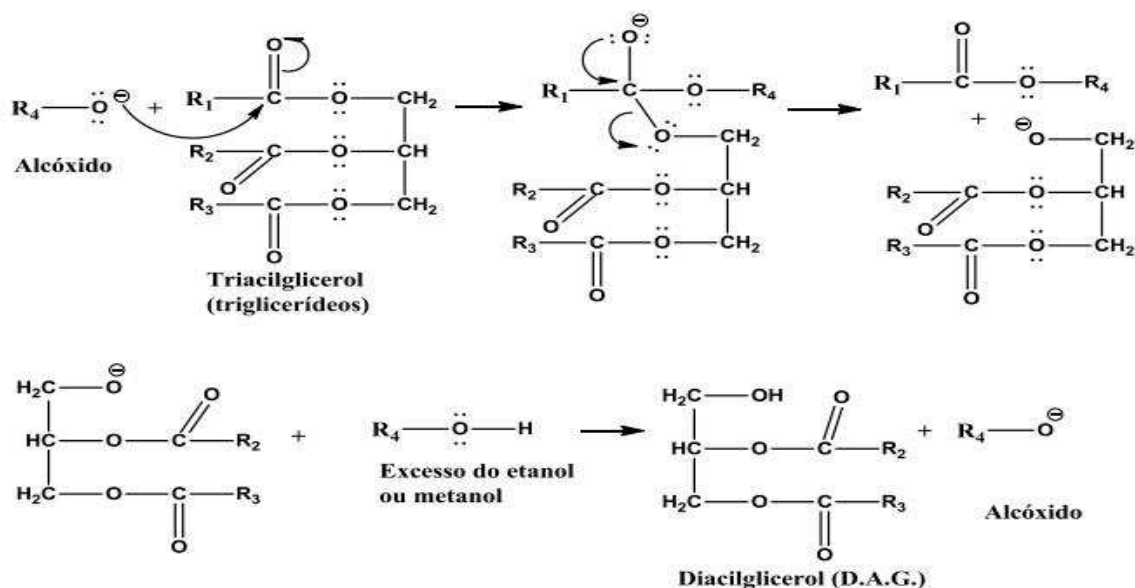
Fonte: Clayden et al., (2011)

Até o nível de 5% de ácidos graxos livres, a reação ainda pode ser realizada com catalisadores alcalinos, mas para tal uma quantidade adicional de catalisador deve ser utilizada para compensar a perda para reações de saponificação (KNOTHE et al., 2006). Os sabões formados durante a reação são removidos com o glicerol ou são eliminados durante a etapa de lavagem aquosa. Quando a concentração de ácidos graxos livres é superior a 5%, os sabões inibem a separação de fases entre o glicerol e os ésteres metílicos e contribuem para a formação de emulsões durante a lavagem aquosa (LEUNG et al., 2010). Nestes casos, é necessária a utilização de um catalisador ácido como o ácido sulfúrico para esterificar os ácidos graxos livres a ésteres metílicos, como demonstra o mecanismo da reação apresentado na Figura 8.

Figura 8 – Mecanismo de reação de esterificação via catálise ácida.

Fonte: Clayden et al., (2011)

Portanto, este processo pode ser utilizado como um pré-tratamento da matéria-prima, na qual os ácidos graxos livres são convertidos em ésteres metílicos, proporcionando assim uma redução nos níveis de ácidos graxos livres. Assim sendo, matéria-prima pré-tratada, com baixo teor de ácidos graxos livres, pode ser transesterificada com um catalisador alcalino para converter os triacilgliceróis em ésteres metílicos, como demonstra o mecanismo da reação apresentado na Figura 9 (DONG et al., 2013).

Figura 9 – Mecanismo de reação de transesterificação via catálise alcalina.

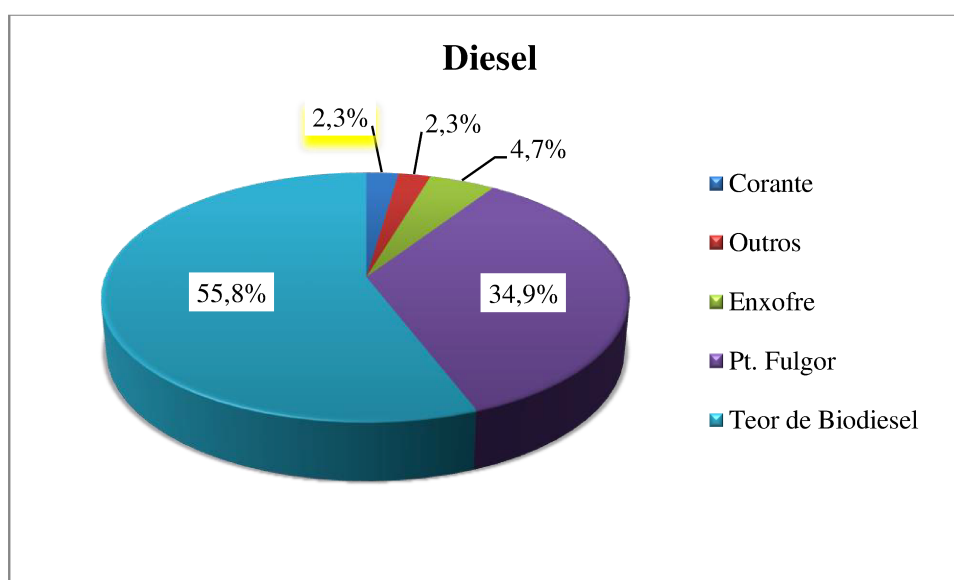
Fonte: Clayden et al., (2011)

2.6. Índice de não conformidades (teor e normas)

O biodiesel é miscível com o diesel em qualquer proporção. Em muitos países, esta propriedade levou ao uso de misturas binárias biodiesel/diesel, ao invés do biodiesel puro. No caso do Brasil, esta mistura é comercializado na seguinte proporção: 7% de biodiesel a 93% de diesel puro, formando um combustível denominado de B7 (BRASIL, 2014).

Atualmente, a produção do biodiesel no Brasil é subsidiada pelo governo, razão pela qual a sua comercialização é feita por meio de editais de leilões como forma de garantir que o biodiesel arrematado seja destinado à mistura com o óleo diesel de petróleo nas condições previstas na legislação vigente (ANP, 2016c). Mesmo assim, a comercialização do biodiesel com teores fora do que é exigido pela norma ainda é feita em grande percentagem. Este cenário é comprovado com os dados estatísticos de monitoramento da qualidade dos combustíveis líquidos automotivos do mês de Agosto de 2016, em que mostram que a característica teor de biodiesel representou 55,8% das não conformidades observadas para o diesel comercializado. As outras são devido aos corantes, aspectos, ponto de fulgor, teor de enxofre e outros, como se pode ver na Figura 10 (ANP, 2016a).

Figura 10 – Índice de não conformidade com relação a natureza do diesel.



Fonte: ANP (2016a).

Diante do exposto, faz-se necessário o desenvolvimento de metodologias que forneçam resposta rápida e eficiente para monitorar quantitativamente e qualitativamente o teor de biodiesel na mistura binária com diesel em concentrações principalmente em torno dos 7% (v/v).

O principal critério para a monitoração do teor de biodiesel é o atendimento a um método padrão apropriado. No entanto, existem três métodos padrão para quantificar biodiesel em misturas com diesel, a saber:

- Norma Européia EN 14078 (2004) (do inglês, European Standard);
- Associação Brasileira de Normas Técnicas (ABNT) NBR 15568 (2008);
- Sociedade Americana de Testes e Materiais (ASTM) D7371 (2007) (do inglês, American Society for Testing and Materials).

A Norma Européia EN 14078 (2004) que utiliza a Espectrometria na região do Infravermelho Médio (MIR) baseia-se na obtenção de uma curva de calibração (univariada) e emprega o sinal de absorção associado à vibração de estiramento da carbonila (C=O) presente nos ésteres ($\sim 1745\text{ cm}^{-1}$).

A norma ABNT NBR 15568 sugere a determinação do teor de biodiesel metílico e/ou etílico em óleo diesel usando modelo de calibração multivariada, baseado no método quimiométrico de regressão por Quadrados Mínimos Parciais (PLS), que emprega medidas de absorvância no infravermelho médio nas regiões espectrais de 1100 cm^{-1} à 1200 cm^{-1} e de 1735 cm^{-1} à 1750 cm^{-1} , atribuídas ao estiramento das ligações C-O e C=O de ésteres, respectivamente. Para tal, são apresentados dois modelos de regressão, em que um é para a faixa de concentração de biodiesel de 0 % (v/v) à 8 % (v/v) e outro para a faixa de 8 % (v/v) à 30 % (v/v).

A ASTM D7371 também utilizam a Espectrometria no Infravermelho Médio, através de medições de transmitância ou Reflectância Total Atenuada (ATR) aliada ao método quimiométrico de regressão por Quadrados Mínimos Parciais (PLS). No entanto, para a obtenção da exatidão necessária é recomendada a utilização de dois ou três modelos analíticos. Assim, o método ASTM sugere o desenvolvimento de três modelos dependendo da faixa de concentração, ou seja, 0-10% (v/v), 10-30% (v/v) e 30-100% (v/v).

Para a validação do modelo de calibração é recomendado seguir as orientações da norma ASTM E1655-05 que descreve um guia prático de procedimentos para coleta e tratamento de dados na quantificação de analitos quando se utiliza a Espectrometria no Infravermelho.

Portanto, vários autores têm publicado trabalhos relacionados à quantificação utilizando estas normas, como são os casos: Gontijo et al. (2014) aplicaram a Espectrometria MIR aliada ao PLS para quantificar biodieseis metílicos e etílicos provenientes de óleo de soja, na faixa de concentração de 1,00 a 30,00% (v/v) em amostras de misturas biodiesel/diesel, seguindo as orientações da ASTM E1655-05; Buiatte et al., (2016) desenvolveram metodologias multivariadas de acordo com a norma ABNT NBR 15568 para a identificação e quantificação do teor de biodiesel metílico de Algodão em mistura com diesel usando Espectrometria no infravermelho médio e métodos quimiométrico PLS e cartas de controle multivariadas, em que os modelos foram validados a partir de figuras de mérito multivariadas conforme diretrizes da ASTM E1655-05; Guimarães et al. (2014) utilizaram PLS para quantificar etanol em amostras de biodieseis etílicos B100 provenientes de óleos de soja e de fritura usado, em que os modelos obtidos também foram validados a partir de figuras de mérito multivariadas de acordo com as diretrizes da ASTM E1655-05.

Os modelos apresentados neste trabalho foram desenvolvidos e validados tendo como base os princípios das normas ABNT NBR 15568, ASTM D7371 e ASTM E1655-05, com destaque para:

- A construção de um único modelo de calibração multivariada para a faixa de concentração de 0,25 % (v/v) a 30 % (v/v);
- O não uso de solventes;
- Uso de pequenas quantidades de amostra;
- A não seleção de regiões espectrais para a construção dos modelos.

3. ESPECTROMETRIA NO INFRAVERMELHO

3.1. Espectrometria no Infravermelho Médio

A Espectrometria MIR refere-se a transições vibracionais e rotacionais nas ligações das moléculas na região de 4.000 a 200 cm^{-1} e, tem sua utilização em grande parte voltada para a caracterização de compostos orgânicos (BARBOSA, 2008). Dentro desta região, existe uma faixa espectral específica de 1200 a 700 cm^{-1} , conhecida como região de impressão digital ou “*fingerprint*”, em que pequenas diferenças na estrutura e na constituição de uma molécula resultam em mudanças significativas nas intensidades ou perfis das bandas de absorção e por esse fato são usadas como parâmetro de identificação ou distinção entre compostos (BARBOSA, 2008).

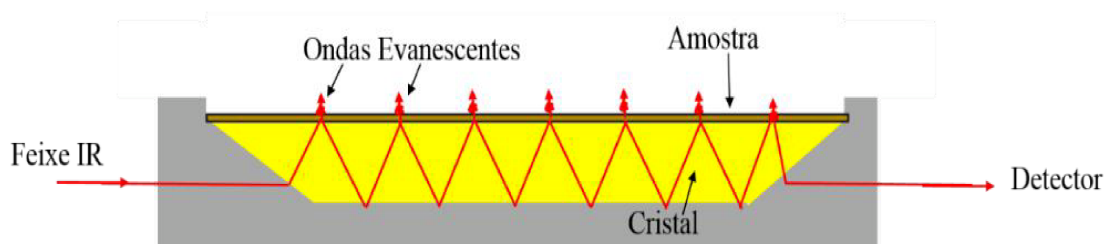
A primeira vantagem da Espectrometria MIR é que trata-se de um método quase universal; muitas moléculas, sejam sólidos, líquidos, gases, semi-sólidos, pós, polímeros, substâncias orgânicas, inorgânicas, materiais biológicos, substâncias puras e misturas apresentam intensas bandas de absorção nessa região. A segunda vantagem é que os espectros são ricos em informação, uma vez que as bandas fornecem as estruturas das moléculas em uma amostra, enquanto que suas intensidades fornecem as concentrações, e as larguras das bandas são sensíveis a matriz química da amostra, incluindo pH e ligações de hidrogênio. A terceira vantagem consiste na rapidez e facilidade de se obter um espectro de infravermelho. Embora o tempo de medição varie conforme a natureza e técnica de amostragem, em geral, os espectros são medidos em cinco minutos ou menos. A quarta vantagem está no custo relativamente baixo em relação as outras técnicas instrumentais, tais como Ressonância Magnética Nuclear (NMR), Cromatografia Gasosa – Espectrometria de Massa (GC-MS) e Cromatografia Líquida – Espectrometria de Massa (LC-MS) (SMITH, 2011).

Além disso, no MIR pode-se empregar o acessório de Reflexão Total Atenuada (ATR) em que os espectros de absorção são obtidos rapidamente em uma variedade de tipos de amostras com um mínimo de preparação. Assim, linhas, fios, tecidos, fibras, pós, pastas, suspensões, polímeros, borrachas e outros materiais podem ser analisados pressionando as amostras sobre o cristal ATR. Para amostras líquidas, o líquido pode ser adicionado diretamente no acessório ATR, como por exemplo, amostras de biodiesel em diesel. Soluções aquosas podem também ser usadas, desde que o cristal não seja solúvel em água (HOLLER et al., 2009). Utilizando-se este tipo de acessório a amostra pode ser preservada, ou seja, utilizando-se o ATR a técnica tem a vantagem de ser não destrutiva.

3.1.1. Refletância Total Atenuada

A técnica de Refletância Total Atenuada (ATR, do inglês, *Attenuated Total Reflectance*) baseia-se no fato da ocorrência da reflexão da radiação infravermelha quando ela passa de um meio mais denso (cristal de ATR) para um menos denso (a amostra). Durante o processo de reflexão, o feixe se comporta como se penetrasse uma pequena distância no meio menos denso antes de ser refletido. A profundidade de penetração, que varia de uma fração até vários comprimentos de onda, depende do comprimento de onda da radiação incidente, dos índices de refração dos dois materiais e do ângulo do feixe incidente em relação à interface. Essa radiação penetrante é denominada de onda evanescente (HARRICK, 1964). Nos comprimentos de onda onde o meio menos denso absorve essa radiação, o feixe é atenuado. Esse processo é conhecido como Refletância Total Atenuada (ATR), representado na Figura 11 (HARRICK, 1964; HOLLER et al., 2009).

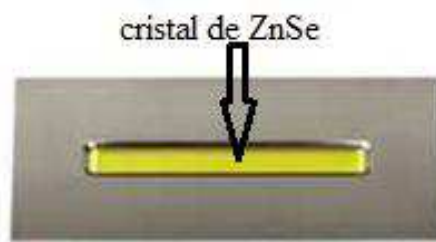
Figura 11 – Processo de reflexão no ATR horizontal.



Fonte: O autor.

A Figura 12, ilustra o acessório de ATR produzido de seleneto de zinco (ZnSe) que foi usado neste trabalho. Ele é insolúvel em água, mais duro que o KRS-5, e é transparente para a maioria da faixa do infravermelho médio. Porém, pós duros e sólidos podem arranhá-lo, ele quebrará se muita pressão for aplicada e é atacado por ácidos e bases fortes. Este material trabalha bem com líquidos, polímeros e sólidos macios (COATES, 2006; PAVIA et al., 2010).

Figura 12 – Cristal de ZnSe do acessório ATR.



Fonte: O autor.

A Reflectância Total Atenuada (ATR) apresenta várias vantagens em relação aos outros métodos citados anteriormente (SMITH, 2011). Para muitas análises, não é necessário o preparo de amostras; esta é colocada sobre o cristal, obtém-se o espectro, remove-se a amostra e limpa-se o cristal, fazendo esta técnica rápida e fácil de ser utilizada.

No entanto, muitos estudos encontrados na literatura de análises quantitativas e qualitativas a partir de dados de Espectrometria MIR envolvem o uso de métodos quimiométricos. Isso ocorre devido a complexidade dos dados obtidos a partir destes espectros, os quais visualmente, isto é, sem qualquer tratamento, é difícil extrair informações sobre a qualidade das amostras.

4. MÉTODOS QUIMIOMÉTRICOS

Hoje em dia, os instrumentos analíticos geram grande quantidade de informação química (variáveis ou características) para um grande número de amostras. No entanto, a necessidade de extrair ao máximo a informação química útil que de outra forma estaria soterrada na avalanche de dados produzidos pela moderna instrumentação foi a causa principal do surgimento e rápido desenvolvimento da Quimiometria (FERREIRA, 2015).

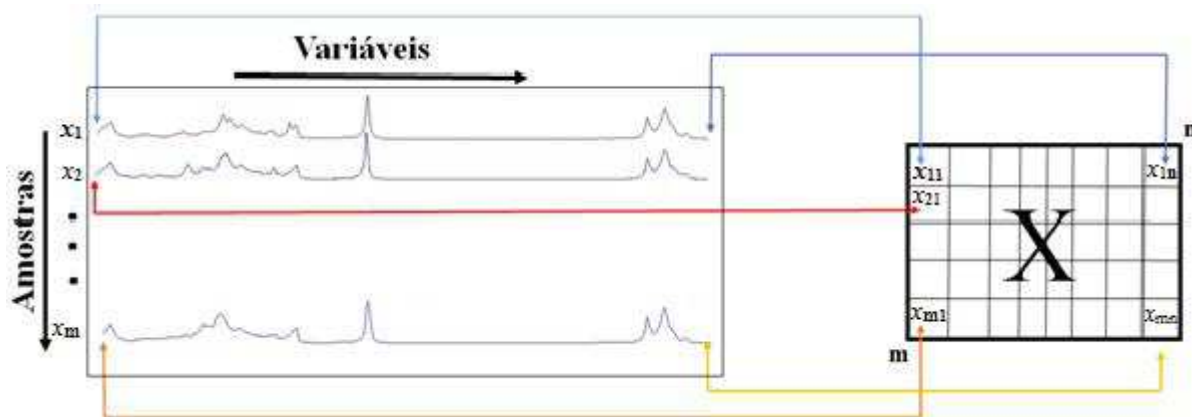
A Quimiometria é definida como a aplicação de métodos matemáticos, estatísticos e de lógica formal para o tratamento de dados químicos (MASSART et al., 1998). Ela é dividida de acordo com a área em que é aplicada, sendo as principais: processamento de sinais analíticos, planejamento e otimização de experimentos, reconhecimento de padrões e classificação de dados, calibração multivariada, monitoramento e modelagem de processos multivariados, métodos de inteligência artificial, entre outros (VANDEGINSTE et al., 1988; NETO, 2005).

Neste trabalho, foi dado foco em duas áreas, a saber: calibração multivariada, que tem por objetivo a construção de um modelo matemático que correlacione a resposta de um dado instrumento analítico com alguma(s) propriedade(s) de interesse, de um conjunto de amostras conhecidas para posteriormente tornar possível a quantificação dessa(s) propriedade(s) em novas amostras desconhecidas; a outra área foi de reconhecimento de padrões e classificação de dados, em que a propriedade de interesse é categórica (discreta), isto é, cada uma das amostras é descrita por um conjunto de medidas experimentais, denominado de “padrão” e são classificadas de acordo com uma propriedade de interesse (bom/ruim; falso/verdadeiro; ativo/não ativo, etc.). Assim, a determinação da propriedade de interesse ao atribuir uma amostra à sua respectiva classe é denominada de reconhecimento (FERREIRA, 2015).

4.1. Organização de dados

Uma vez coletados, os dados devem ser organizados em um arranjo ordenado de linhas e colunas, constituindo uma matriz **X**, em que cada linha corresponde a uma amostra “*m*” e cada coluna corresponde a uma variável “*n*”, na qual $m = 1, 2, 3, \dots, m$ e $n = 1, 2, 3, \dots, n$ (FERREIRA, 2015).

A Figura 13 ilustra como uma matriz de dados **X** de dimensão ($m \times n$) pode ser construída a partir de um vetor de respostas instrumentais.

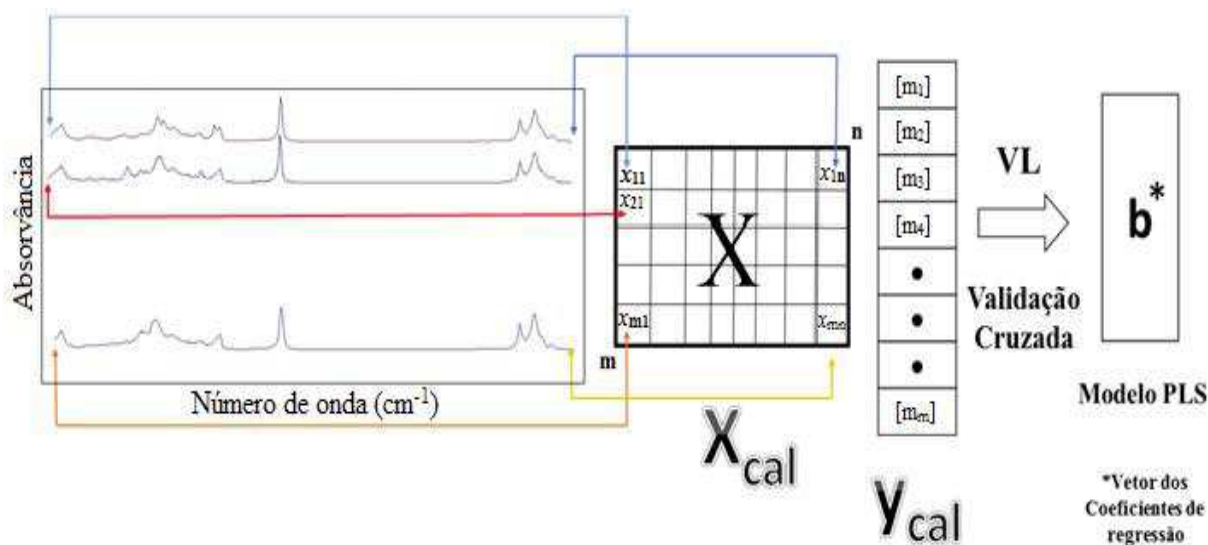
Figura 13 – Organização dos dados para a construção da matriz **X**.

Fonte: O autor.

4.2 Regressão por Quadrados Mínimos Parciais (PLS).

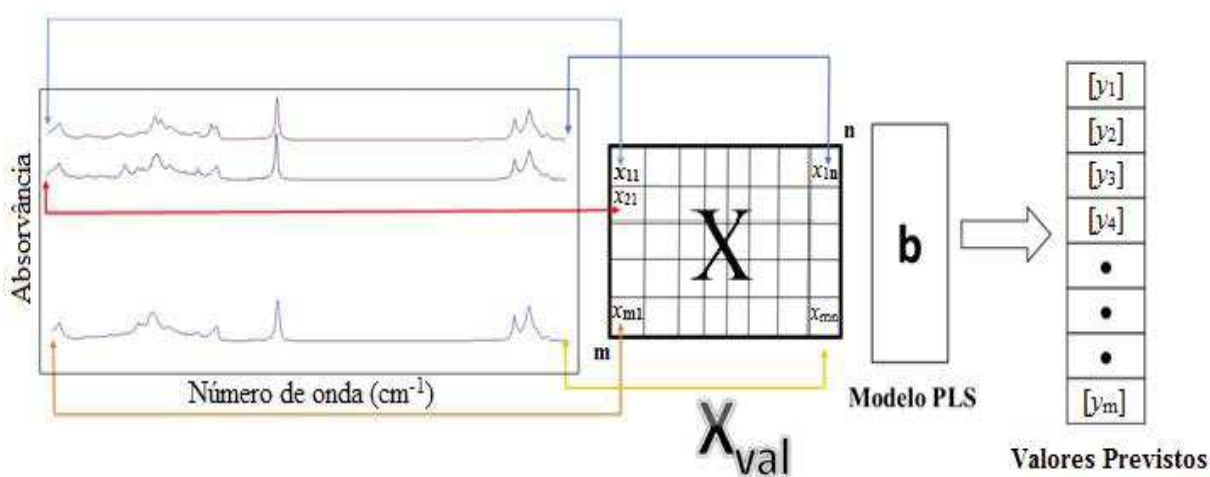
A regressão por Quadrados Mínimos Parciais é um método amplamente utilizado nos dados de calibração multivariada que emprega a análise de fatores (FERREIRA, 2015). Este método consiste na retirada de informações do conjunto de dados da matriz **X** e as correlaciona, através de operações matemáticas, com as informações retiradas do conjunto de dados de referência (vetor **y**) (VANDEGINSTE et al., 1988). Através da combinação da matriz **X** com o vetor **y** são obtidas as variáveis latentes (VL), as quais têm por finalidade alcançar a máxima covariância entre, no nosso caso, os espectros e as concentrações da espécie de interesse.

Na construção do modelo de calibração (Figura 14) é usado um número de VL que proporcione o menor erro de previsão, ou seja, que a diferença entre os valores de referência e os valores previstos pelo modelo seja o menor possível. Assim, têm-se os vetores dos coeficientes de regressão.

Figura 14 – Construção do modelo de regressão por Quadrados Mínimos Parciais.

Fonte: O autor.

Na etapa de validação, prepara-se outro conjunto de amostras na mesma faixa de concentração das amostras de calibração e faz-se uma correlação entre essa nova matriz X e o vetor de regressão, prevendo-se então a concentração dessas amostras conforme é mostrado na Figura 15 (WOLD et al., 2001).

Figura 15 – Validação do modelo de regressão por Quadrados Mínimos Parciais.

Fonte: O autor.

A determinação do número correto de VL é realizada através do método de validação cruzada, o qual se baseia na habilidade de previsão de um modelo construído por parte de um conjunto de dados seguido pela previsão do restante do conjunto de dados.

A validação cruzada pode ser realizada em blocos, ou seja, um número determinado de amostras é deixado de fora no processo de construção do modelo e a seguir essas amostras são previstas pelo modelo construído. Isso é semelhante ao método *leave-one-out*, em que uma amostra é deixada de fora no processo de construção do modelo e a seguir essa amostra é prevista pelo modelo construído. Na sequência, esta amostra retorna ao conjunto de calibração, e outra é retirada e prevista pelo modelo construído pelas restantes, e esse processo é repetido até que todas as amostras tenham sido retiradas do modelo uma vez (VANDEGINSTE et al., 1988). Em ambos os casos, o processo é repetido até que todas as amostras tenham previstas e o Erro Quadrático Médio de Validação Cruzada (RMSECV) é calculado de acordo com a Equação 1:

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

Onde y_i , \hat{y}_i e n são, os valores reais, valores previstos pelo modelo PLS e número de amostras de calibração, respectivamente.

4.2.1. Detecção de Amostras Anômalas (Outlier)

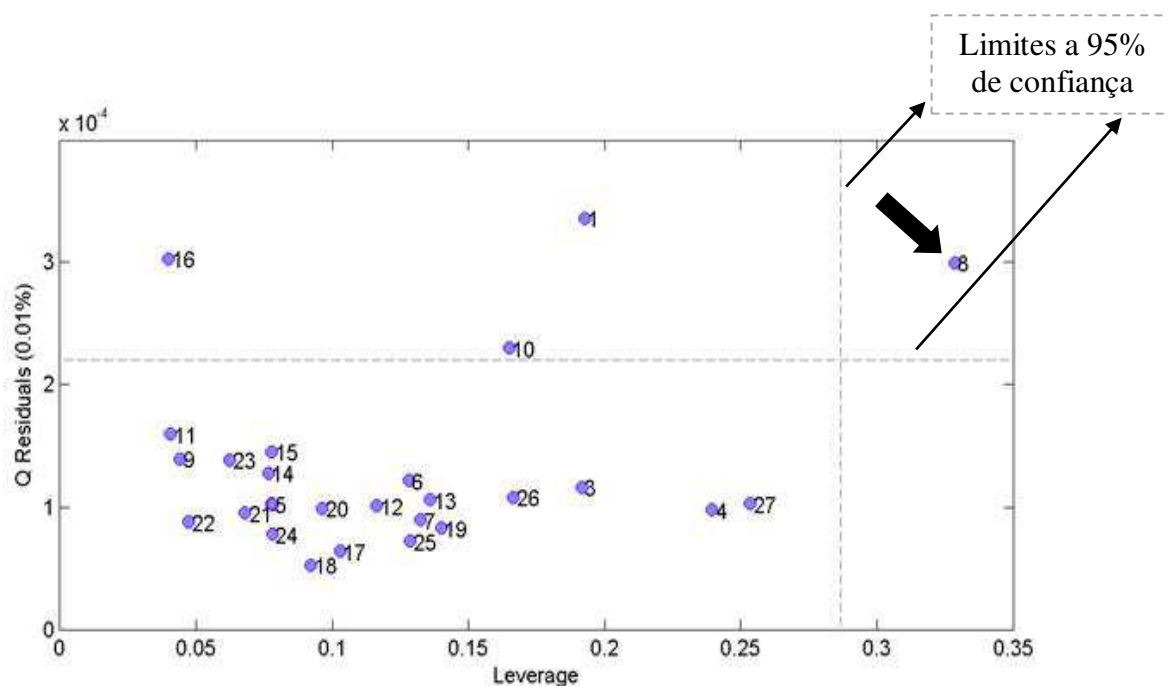
Ao verificar a qualidade do conjunto de amostras utilizadas na etapa de calibração, deve-se assegurar que as amostras formem um conjunto homogêneo. Portanto, se o perfil de uma amostra é diferente das restantes, ela será considerada de amostra anômala ou *outlier* e poderá ter uma influência inadequada no modelo.

As amostras anômalas ocorrem por diferentes razões. Dentre elas, destacam-se as seguintes: os erros de digitação; os erros de laboratório, como por exemplo, na determinação da propriedade de interesse por um método de referência; e os erros instrumentais, quando o espectro da amostra em questão é de baixa qualidade. Uma vez que este tipo de amostra pode afetar a qualidade global do modelo PLS, faz-se necessário a sua retirada do conjunto de calibração (VALDERRAMA et al., 2009).

A identificação dessas amostras anômalas é feita analisando o gráfico dos valores de *leverage* (o grau que uma amostra está distante da média do conjunto de dados) versus valores

de resíduos (a parte não modelada nos dados espectrais ou nos valores de concentração analítica), conforme mostra a Figura 16 (VANDEGINSTE et al., 1988). Nesta figura, pode se observar que apenas a amostra de nº 8 encontra-se fora dos limites a 95% de confiança (linhas tracejadas), isto é, apresenta maior número de *leverage* e maior número de resíduos. Portanto, ela deve ser retirada do conjunto de calibração, pois sua permanência implica no aumento de informação errônea, podendo diminuir a eficiência de previsão caso tal informação seja correlacionada com as amostras do conjunto de previsão. Caso seja necessária a informação sobre esta amostra, o analista deve refazer a sua medição até esta fique dentro dos limites a 95% de confiança.

Figura 16 – Gráfico de *leverage* versus Q Residuals.



Fonte: O autor.

4.2.2. Validação dos modelos PLS

Para a validação dos modelos PLS desenvolvidos neste trabalho, seguiu-se as orientações da norma ASTM E1655-05 (2012), que descrevem especificamente o desenvolvimento e validação de modelos quantitativos multivariados quando se utiliza a Espectrometria no Infravermelho. Assim, foi analisada a relação sinal-ruído e as seguintes figuras de mérito: seletividade, sensibilidade, sensibilidade analítica, limite de detecção,

limite de quantificação, teste para erro sistemático (bias, SDV e t_{bias}), de acordo com as equações 2 a 9, respectivamente, apresentadas na Tabela 1 (VALDERRAMA et al., 2009; ROCHA et al., 2012; SILVA et al., 2012; FERREIRA et al., 2013).

A relação sinal-ruído foi analisada, tendo em conta que, todo espectro medido experimentalmente possa ser representado pela soma de dois vetores ortogonais entre si, onde um deles é vetor NAS, que é exclusivo do constituinte de interesse (sinal analítico), e outro é o vetor interferente, que contém as informações de todos os outros constituintes presentes (LORBER, 1986). No entanto, o sinal analítico líquido (NAS) (\mathbf{r}_k^*) foi calculado para analitos decompondo a matriz \mathbf{X} das variáveis preditoras do conjunto de calibração em duas outras: \mathbf{X}_k (contribuição exclusiva do constituinte de interesse) e \mathbf{X}_{-k} (contribuição dos interferentes), de acordo com a Equação (10) (LORBER et al., 1997; FERREIRA, 2015). Onde, a matriz \mathbf{P}_{NAS} (de dimensões $J \times J$) faz a projeção dos espectros no espaço ortogonal, no qual se encontra o vetor NAS; k identifica o analito de interesse; \mathbf{x}_i representa o espectro de uma dada amostra; \mathbf{I} é uma unidade de matriz adequadamente dimensionada; \mathbf{X}_{-k} é uma matriz dos sinais espectrais gerados por todos os outros analitos, com exceção de k ; $(\mathbf{X}_{-k})^+$ é a pseudo-inversa de \mathbf{X}_{-k} , normalmente calculado pela decomposição em valores singulares usando A fatores.

$$\mathbf{r}_k^* = \mathbf{P}_{\text{NAS},k} \mathbf{x}_i = [\mathbf{I} - \mathbf{X}_{-k}(\mathbf{X}_{-k})^+] \mathbf{x}_i \quad (10)$$

A concentração de k na amostra desconhecida foi obtida a partir do espectro MIR de \mathbf{r} , usando a Equação (11) (ROCHA et al., 2012).

$$\mathbf{Y}_{\text{un},k} = \frac{\mathbf{S}_k^T \mathbf{P}_{\text{NAS},k} \mathbf{r}}{\mathbf{S}_k^T \mathbf{P}_{\text{NAS},k} \mathbf{s}_k} = \frac{\mathbf{S}_k^T \mathbf{P}_{\text{NAS},k} \mathbf{P}_{\text{NAS},k} \mathbf{r}}{\mathbf{S}_k^T \mathbf{P}_{\text{NAS},k} \mathbf{P}_{\text{NAS},k} \mathbf{s}_k} = \frac{(\mathbf{S}_k^*)^T \mathbf{r}_k^*}{\|\mathbf{S}_k^*\|^2} \quad (11)$$

Seletividade

É a medida do grau de sobreposição entre o sinal da espécie de interesse e os interferentes presentes na amostra indicando, também, a parte do sinal que é perdida por essa sobreposição (VALDERRAMA et al., 2009; SOUZA et al., 2015).

Sensibilidade

Corresponde à fração do sinal responsável pelo acréscimo de uma unidade de concentração à propriedade de interesse (ROCHA et al., 2012; SILVA et al., 2012).

Sensibilidade analítica

Apresenta a sensibilidade do método em termos da unidade de concentração que é utilizada, sendo definida como a razão entre a sensibilidade e o desvio padrão do sinal de referência (VALDERRAMA et al., 2009; ROCHA et al., 2012).

Limite de detecção (LD)

Equivale à menor concentração da substância de interesse que pode ser detectada, mas não necessariamente quantificada (LORBER et al., 1997).

Limite de quantificação (LQ)

Representa a menor concentração da substância de interesse que pode ser medida com uma incerteza máxima de 10% (VALDERRAMA et al., 2007; SILVA et al., 2012).

Teste para erro sistemático

Para este teste, primeiro é calculado o *bias* (a diferença entre a média da população e o valor verdadeiro) para o conjunto de validação (VALDERRAMA et al., 2009). A seguir, é calculado o desvio padrão dos erros de validação (SDV) (FERREIRA et al., 2013). Por fim, é determinado o valor de t_{bias} (ROCHA et al., 2012).

Caso o valor de t_{bias} resultar em maior valor do que o valor de $t_{crítico}$ para n graus de liberdade e 95% de confiança, isso é uma evidência de que erros sistemáticos presentes no modelo multivariado são significativos. No entanto, se o valor de t_{bias} apresentar valor menor do que $t_{crítico}$, então, o erro sistemático no modelo pode ser considerado desprezível (ASTM, 2012).

Tabela 1 – Equações usadas para o cálculo das figuras de mérito.*

Designação	Figura de mérito	Equação
Eq. (2)	Seletividade	$\hat{S}E_{L_i} = \frac{n\hat{s}_i}{\ x_i\ }$
Eq. (3)	Sensibilidade	$\hat{S}E_N = \frac{1}{\ b_k\ }$
Eq. (4)	Sensibilidade analítica	$\gamma = \frac{\hat{S}E_N}{\ \delta_x\ }$
Eq. (5)	Limite de detecção	$LD = 3.3\delta_x \frac{1}{\hat{S}E_N}$
Eq. (6)	Limite de quantificação	$LQ = 10\delta_x \frac{1}{\hat{S}E_N}$
Eq. (7)	bias	$bias = \frac{\sum_{i=1}^{n_{val}}(y_i - \hat{y}_i)}{n_{val}}$
Eq. (8)	SDV	$SDV = \sqrt{\frac{\sum[(y_i - \hat{y}_i) - bias]^2}{n_{val} - 1}}$
Eq. (9)	t_{bias}	$t_{bias} = \frac{ bias \sqrt{n_{val}}}{SDV}$

*onde, y_i e \hat{y}_i correspondem aos valores de referência e aos previstos pelo modelo, respectivamente; n_{val} é o número de amostras de validação; $n\hat{s}_i$ é o valor escalar do sinal analítico líquido para a amostra i ; $\|x_i\|$ é a norma de cada espectro; b_k é o vetor dos coeficientes de regressão para a espécie de interesse k estimada pelo modelo de calibração multivariada; δ_x é o desvio padrão do sinal de referência.

Fonte: O autor.

4.3. Métodos de classificação

Os métodos de classificação, também conhecidos como métodos de reconhecimento de padrões são divididos em dois tipos, a saber: métodos não supervisionados e métodos supervisionados (FERREIRA et al., 1999; FERREIRA, 2015).

Os métodos não supervisionados de reconhecimento de padrões são usados basicamente para a análise exploratória dos dados. Os mais usados são: Análise de Componentes Principais (PCA) e Análise de Agrupamentos Hierárquicos (HCA) (FERREIRA, 2015).

Os métodos supervisionados de reconhecimento de padrões são usados quando a propriedade de interesse é categórica (discreta), isto é, cada uma das amostras é descrita por um conjunto de medidas experimentais, denominada de “padrão” e são classificadas de acordo com uma propriedade de interesse (bom/ruim; falso/verdadeiro; ativo/não ativo, etc.). Os mais utilizados na química são: A Modelagem Independente e Flexível por Analogia de Classe (SIMCA), K-ésimo Vizinhos mais Próximos (KNN) e Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) (Ferreira, 2015). Estes métodos são aplicados a uma ampla variedade de dados químicos com diversas finalidades, tais como: identificação de perfis e impressões digitais, autenticação, detecção de falsificação, avaliação da qualidade de vários produtos alimentares, interpretação de dados, etc (MASSART et al., 1998).

Para a classificação dos dados deste trabalho, foi usado o método de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) porque com esse método pode se usar a informação do modelo construído para previsão de novas amostras.

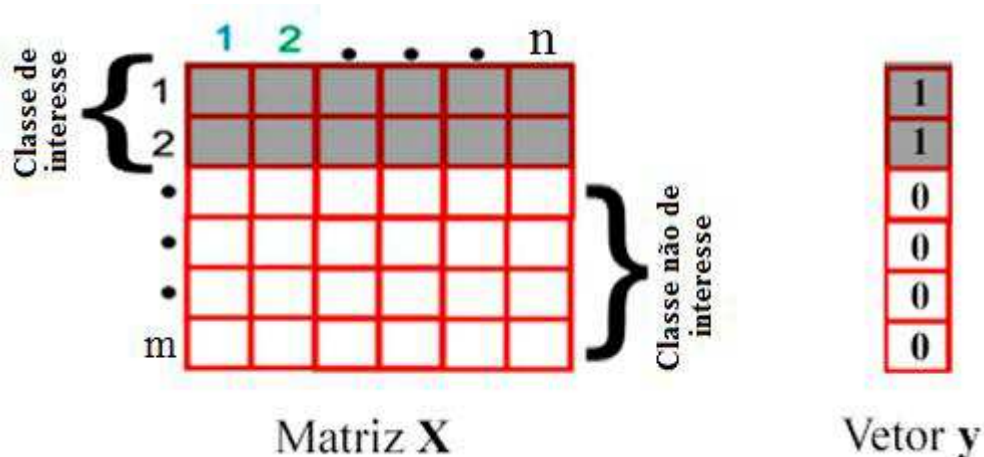
4.3.1. Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA)

O método PLS-DA é uma variação do algoritmo de regressão por (PLS), onde as variáveis da matriz \mathbf{X} (dados espectrais) são relacionadas com a propriedade de interesse categórica (discreta) contidas em um vetor \mathbf{y} , como mostra a Figura 17 (WOLD et al., 2001). O vetor \mathbf{y} é constituído por valores inteiros (0 e 1), onde 0 é usado para classe não de interesse e 1 para a classe de interesse (SERRANO-LOURIDO et al., 2012).

Idealmente, os valores previstos pelo modelo PLS-DA seriam 0 ou 1, mas na prática os valores previstos são próximos de 0 e 1. Tendo em conta que os valores previstos não são 0 ou 1, faz-se necessário o cálculo de um valor limite para separar as classes, denominado de “threshold” representado normalmente em gráfico por uma linha tracejada (WISE et al., 2006).

O valor de limiar “*threshold*” é previsto entre 0 e 1, com base no teorema de Bayes, segundo o qual, o limiar assume que os valores de \mathbf{y} previstos seguem uma distribuição semelhante ao que vai ser observado para as amostras futuras. Quando as duas distribuições estimadas se cruzam, é selecionado um limite, isto é, um valor de \mathbf{y} em que o número de falsos positivos e falsos negativos devem ser minimizados para previsões futuras (ALMEIDA et al., 2013). Quando a amostra apresentar um valor previsto acima do *threshold*, ela é considerada como pertencente à classe 1 e as amostras abaixo deste valor limite são pertencentes à classe 0.

Figura 17 – Representação da matriz de dados **X** e o vetor **y** usados no PLS-DA.



Fonte: O autor.

No entanto, quando as variáveis da matriz **X** são relacionadas com as classes contidas em um vetor **y**, o seu modelo é denominado de PLS-DA, enquanto que quando são relacionadas com as classes contidas em uma matriz **Y** é denominado de PLS2-DA.

O modelo PLS-DA é usado quando se tem apenas duas classes para se distinguir e, quando se tem mais de duas classes para se distinguir simultaneamente é usado o modelo PLS2-DA.

Para a construção do modelo PLS-DA, procura-se encontrar a quantidade de variáveis latentes (VL) que descrevam a covariância nas amostras das matrizes e que tenham a correlação máxima com a classe de valores conhecidos, dando menos peso para a classe irrelevante ou a variância do ruído (BARKER; RAYENS, 2003). O número de VL é determinado a partir da validação das amostras do conjunto de treinamento pelo método de validação cruzada “*leave one out*” (deixar uma de fora por vez), como acontece para os modelos PLS. Portanto, para cada variável latente a eficiência da validação é expressa pelo Erro Quadrático Médio de Validação Cruzada (RMSECV), calculada a partir da Equação 1.

Após a construção do modelo PLS-DA, pode-se fazer a análise da Variável Importante na Projeção (VIP) com intuito de identificar quais variáveis dos espectros MIR foram mais relevantes para separação das classes das amostras (ALMEIDA et al., 2013). Este parâmetro é determinado de acordo com a Equação 12 (MEHMOOD et al., 2012).

$$v_j = \sqrt{p \sum_{a=1}^A [(q^2 a t' a t a) (w a j / \| w a \|)^2] / \sum_{a=1}^A (q^2 a t' a t a)} \quad (12)$$

Onde: v_j é a medida de VIP; q são os pesos de \mathbf{Y} ; $(waj/\|wa\|)^2$ representa a importância da variável j ; $(q^2at'ata)$ é a variância explicada por cada VL; t são os escores.

4.3.1.1. Parâmetros de Avaliação do Modelo PLS-DA

A avaliação do desempenho do modelo de classificação é feita através dos critérios de estatísticas de respostas verdadeiras: sensibilidade e especificidade (ALMEIDA et al., 2013).

A sensibilidade é a capacidade que o modelo tem de classificar corretamente as amostras verdadeiras positivas e, a especificidade é a capacidade que o modelo tem de classificar corretamente as amostras verdadeiras negativas. Os parâmetros sensibilidade e especificidade são calculados de acordo com as equações (13) e (14), respectivamente (XU et al., 2012).

$$\text{Sensibilidade} = \frac{VP}{VP+FN} \quad (13)$$

$$\text{Especificidade} = \frac{VN}{VN+FP} \quad (14)$$

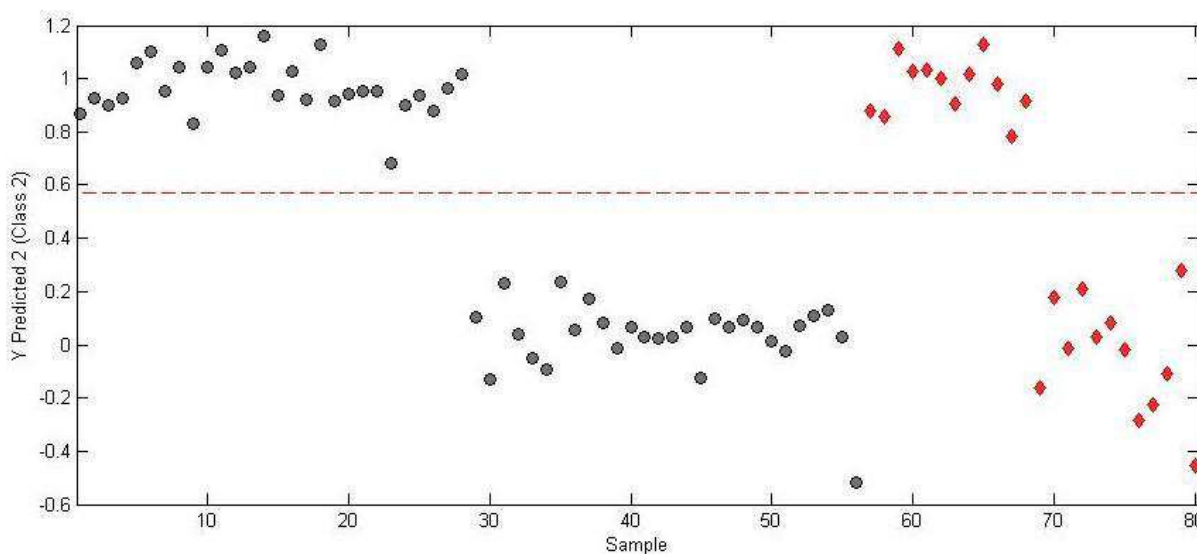
Onde:

- VP (Verdadeiros Positivos) é a probabilidade de uma amostra positiva (pertencente à classe) ser classificada como uma amostra positiva;
- VN (Verdadeiros Negativos) é a probabilidade de uma amostra negativa (não pertencente à classe) ser classificada como uma amostra negativa;
- FP (Falsos Positivos) é a probabilidade de uma amostra negativa (não pertencente à classe) ser classificada como uma amostra positiva (pertencente à classe);
- FN (Falsos Negativos) é a probabilidade de uma amostra positiva (pertencente à classe) ser classificada como uma amostra negativa (não pertencente a classe).

4.3.1.2. Representação gráfica do resultado do modelo PLS-DA

O resultado do modelo PLS-DA geralmente é expresso através de um gráfico de estimativas de classes, onde as amostras posicionadas acima da linha tracejada (*threshold*) são classificadas como pertencentes à classe 1 (classe de interesse) e, as amostras posicionadas abaixo do valor de *threshold* são classificadas como pertencentes à classe 0, ou seja, a classe que não é de interesse. A Figura 18 mostra um exemplo do gráfico de estimativas para um modelo PLS-DA, em que nela pode se observar que as amostras do conjunto de treinamento (●) e de teste (◆) que pertencem à classe de interesse estão acima do *threshold*, separando-se das amostras do conjunto de treinamento (●) e de teste (◆) da classe não de interesse que estão abaixo do *threshold*.

Figura 18 – Representação gráfica do resultado do modelo PLS-DA.



Fonte: O autor.

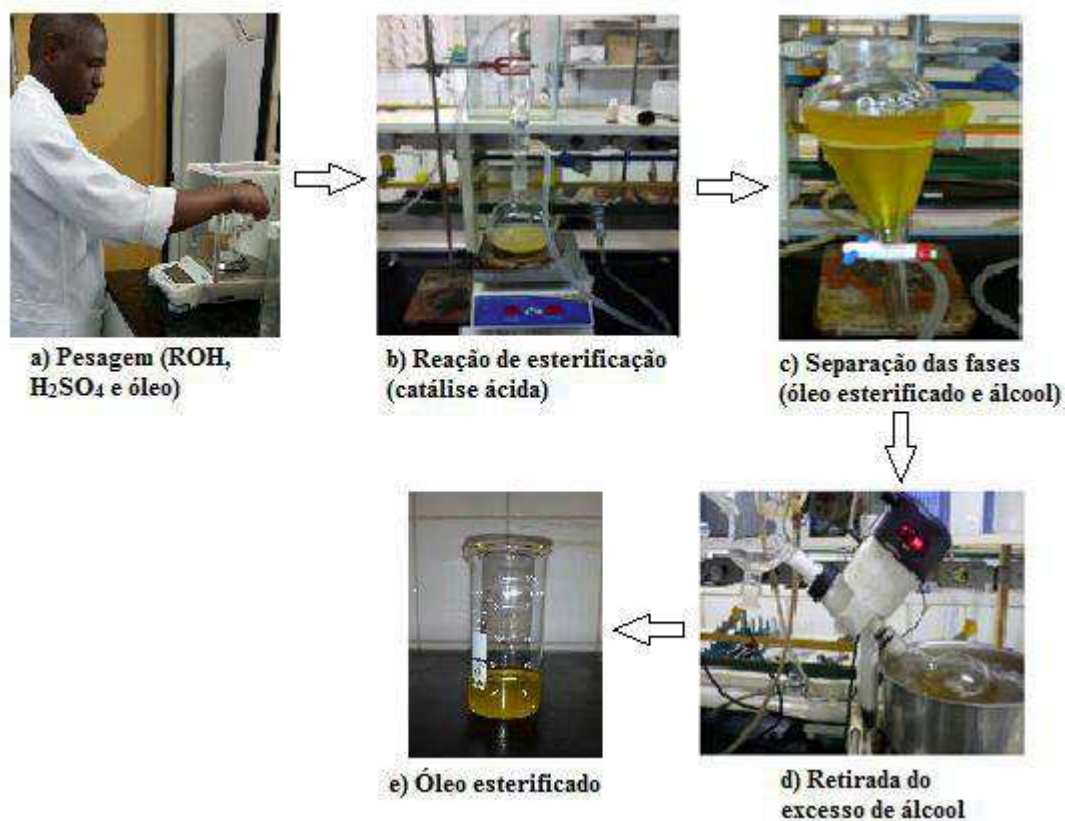
5. PROCEDIMENTO EXPERIMENTAL

5.1. Produção do Biodiesel

Os óleos brutos de Mafurra e Moringa foram obtidos em Moçambique, a partir de método de prensagem a frio, isto é, sem o uso de solventes. No entanto, não houve restrições em relação ao transporte desses óleos para Brasil por serem óleos comerciais que foram transportados em pequena quantidade (500 ml de cada). O óleo de Algodão foi fornecido pela indústria Triângulo Alimentos S/A. Estes óleos foram usados para a produção de biodieseis metílicos puros (B100), no Laboratório de Biocombustíveis do Instituto de Química, da Universidade Federal de Uberlândia.

Os óleos de mafurra e de moringa obtidos em Moçambique, não foram purificados. No entanto, os seus respectivos biodieseis foram produzidos realizando a esterificação antes do processo de transesterificação com o objetivo de reduzir o alto índice de acidez e aumentar a eficiência do processo de conversão dos ácidos graxos livres em ésteres (DONG et al., 2013; RAMOS et al., 2015). Para o óleo de algodão não foi necessário a etapa da esterificação (catálise ácida) por ser óleo refinado e com baixo índice de acidez. Assim, fez-se somente o processo de transesterificação (catálise básica).

Para o processo de esterificação, ilustrado na Figura 19, adiciona-se 3 g de ácido sulfúrico concentrado à 100 g de álcool metílico e, em seguida adiciona-se 100 g do óleo (Mafurra ou Moringa). A mistura deve ser aquecida a uma temperatura de 60 °C sob agitação magnética a 400 rpm e sob refluxo durante 3 horas. Em seguida, coloca-se a mistura num funil separador por 1 hora de repouso para decantação das fases: inferior (óleo esterificado) e superior (misturas de álcool e ácido sulfúrico). Por fim, o óleo separado deve ser transferido para um balão de fundo redondo para a destilação a vácuo usando o evaporador rotatório a 90 rpm e 80 °C durante 1 hora, com o objetivo de remover o álcool em excesso (Portela, 2011; Siteo, 2016).

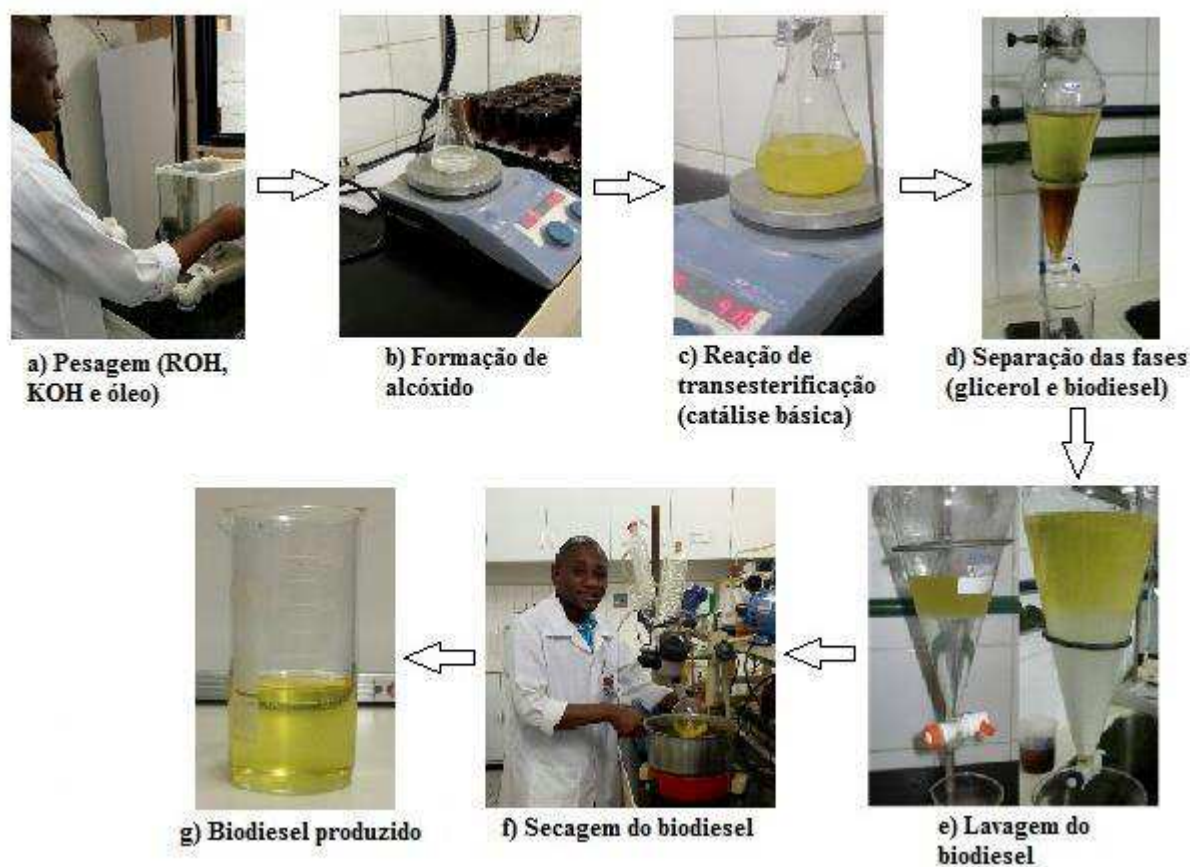
Figura 19 – Etapas do processo de esterificação.

Fonte: O autor.

Para o processo de transesterificação, ilustrado na Figura 20, deve ser usada a proporção de 100,0 g de cada óleo e 30,0 g de metanol na presença de 1,0 g de KOH como catalisador. Para tal, adiciona-se em primeiro lugar, o álcool metílico no hidróxido de potássio para formar o alcóxido. Após a formação do alcóxido, é adicionado o óleo para ocasionar o processo de transesterificação ocorra (Portela, 2011; Siteo, 2016).

A reação de transesterificação deve ser mantida sob agitação magnética a 400 rpm e à temperatura ambiente durante 1 h. Após este tempo, coloca-se a solução resultante num funil de separação por 24 h de repouso para se separar as fases: inferior (glicerol) e superior (biodiesel). Após a separação, o biodiesel obtido deve ser purificado por lavagem com água destilada a 90 °C e, em seguida, seco utilizando um evaporador rotativo durante 1 h, à 88 rpm e 90 °C (Portela, 2011; Siteo, 2016).

Figura 20 – Etapas do processo de transesterificação.



Fonte: O autor.

5.2. Preparação das Amostras

O diesel puro usado no preparo das amostras foi fornecido pela empresa brasileira Transpetro S/A. As amostras foram preparadas à temperatura ambiente ($25 \pm 2^\circ\text{C}$) utilizando-se uma balança analítica (Sartorius, modelo BP211D).

Foram preparadas 85, 81 e 82 amostras de BMMA, BMMo e BMA, respectivamente, por adição do biodiesel B100 ao diesel puro, em um intervalo de concentração de 0,25% a 30,00% (v/v). As concentrações das amostras de 0,25 à 1% (v/v) e depois de 1% (v/v) em diante (final da curva de calibração), foram preparadas com incrementos de 0,25% e 0,35%, respectivamente. Para a construção do modelo PLS de BMMA, 59 amostras foram utilizadas para o conjunto de calibração e 26 amostras para o conjunto de previsão; para o modelo PLS de BMMo, 54 amostras foram utilizadas para o conjunto de calibração e 27 amostras para o

conjunto de previsão e, para o modelo PLS de BMA, 55 amostras foram utilizadas para o conjunto de calibração e 27 amostras para o conjunto de previsão.

Para a construção do modelo PLS-DA, preparou-se dois lotes de amostras para cada tipo de biodiesel (BMMa, BMMo e BMA), isto é, um lote de 40 amostras de B7 e outro de 40 amostras de BX (B3 à B6 e B8 à B19). Em que, 28 amostras de cada lote (B7 e BX) foram utilizadas para o conjunto de treinamento e 12 amostras de cada (B7 e BX) foram utilizadas para o conjunto de teste.

Para a construção do modelo PLS2-DA, produziu-se 40 biodieseis (B100) para cada tipo de biodiesel (BMMa, BMMo e BMA). Desses biodieseis produzidos, preparou-se 40 amostras de mistura de biodiesel/diesel (B7), das quais 28 amostras de cada uma das três classes foram utilizadas para o conjunto de treinamento e 12 amostras de cada uma das três classes foram utilizadas para o conjunto de teste.

5.3. Obtenção dos Espectros MIR

Os dados espectrais MIR das amostras de todos os modelos foram obtidos em quintuplicata e com 16 varreduras, na região de 4.000 a 600 cm^{-1} , resolução de 4 cm^{-1} , usando espectrômetro de marca Perkin Elmer, modelo Spectrum Two (Perkin Elmer, Waltham, MA, EUA) com Transformada de Fourier, equipado com um acessório amostrador (Pike Technologies, Fitchburg, WI, EUA) de Refletância Total Atenuada Horizontal (HATR). Após o registro do espectro de cada amostra, o ATR horizontal foi limpo com álcool isopropílico.

Os espectros foram corrigidos pelo método de *baseline* aplicado nas regiões de 2570 a 1850 cm^{-1} e 4000 a 3200 cm^{-1} (regiões sem bandas de absorção). Após esta correção da linha de base, a região abaixo de 680 cm^{-1} e acima de 3100 cm^{-1} foram descartadas, por apresentarem grande quantidade de ruído e nenhuma banda de absorção, respectivamente.

5.4. Construção dos modelos

Para executar os procedimentos multivariados na construção dos modelos PLS e PLS-DA, foi usado o software MATLAB versão 7.5 (Mathworks Inc.) e PLS_Toolbox, versão 8.0.2 (Eigenvetor Research).

O pré-processamento foi realizado centrando os dados pela média, com objectivo de minimizar as variações sistemáticas indesejáveis nos dados. O método de validação cruzada

leave-one-out foi usado para a validação interna dos modelos. No entanto, para este trabalho foram construídos:

- Três modelos PLS e três modelos PLS-DA para quantificar e classificar o teor dos três tipos de biodiesel (BMMa, BMMo e BMA) em misturas com diesel puro (B7), respectivamente;
- Um modelo PLS2-DA para classificar simultaneamente os três tipos de biodiesel (BMMa, BMMo e BMA) em mistura com o diesel puro (B7).

6. RESULTADOS E DISCUSSÃO

6.1. Caracterização dos Espectros MIR

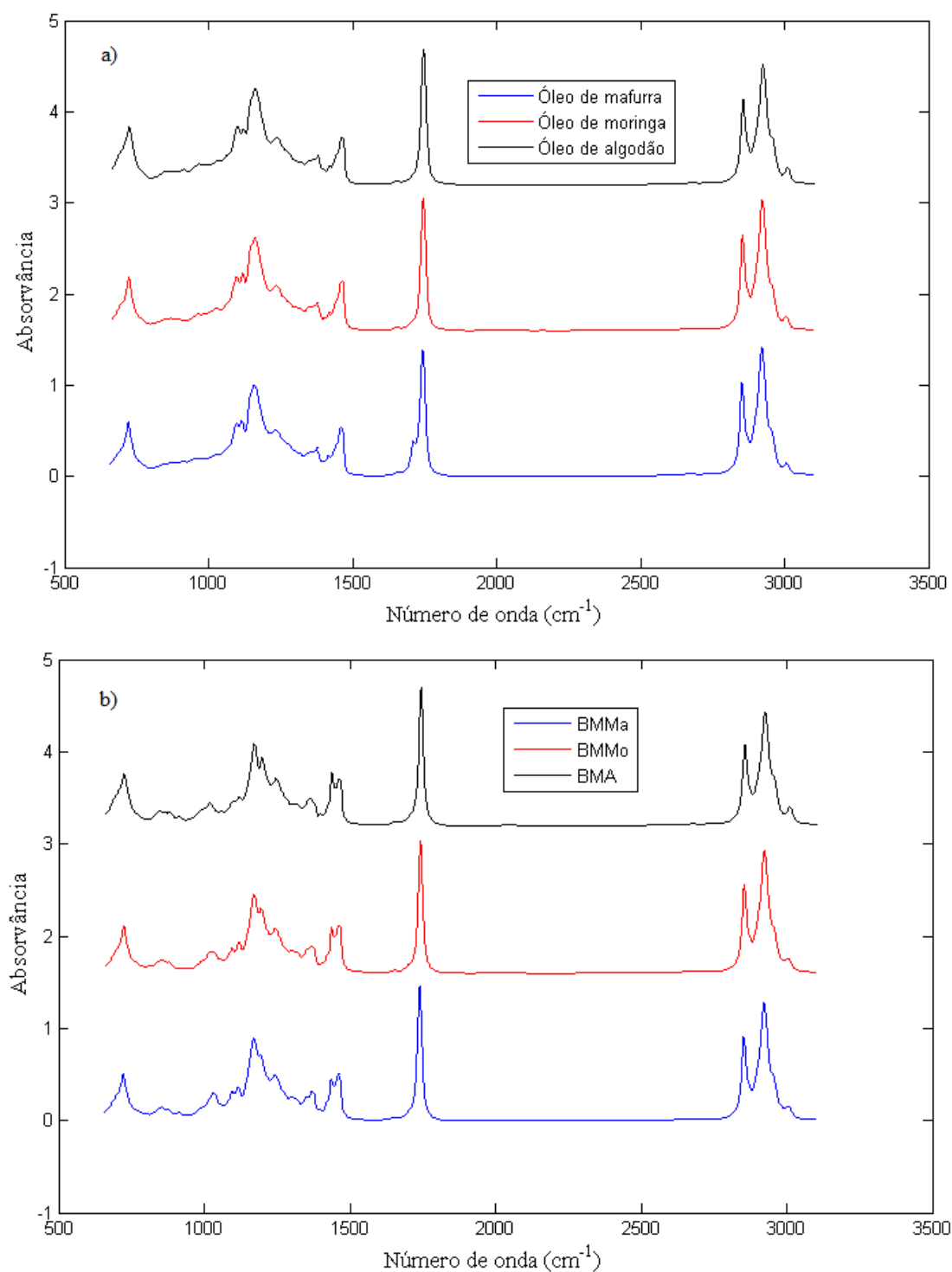
A Figura 21 a ilustra os espectros MIR dos óleos de mafurra, moringa e algodão. Nestes espectros, notam-se semelhanças nas seguintes bandas de absorções significativas (GUILLÉN; CABO, 1997):

- As bandas de absorção de maior intensidade na região espectral entre 3008 à 2853 cm^{-1} , devido às vibrações de estiramentos simétricos e assimétricos da ligação C-H de grupos metila ($-\text{CH}_3$), metileno ($-\text{CH}_2$) e de carbonos insaturados;
- A banda de absorção de maior intensidade na região espectral de 1744 cm^{-1} , atribuído a $\nu\text{C}=\text{O}$ dos acilgliceróis;
- As bandas de intensidades intermédias nas regiões de 1464 cm^{-1} atribuído a deformação angular assimétrica ($\delta_{\text{as}} \text{CH}_3$) e 1378 cm^{-1} atribuído a deformação angular simétrica ($\delta_{\text{s}} \text{CH}_3$);
- As bandas de absorção observadas nas regiões espectrais de 1236, 1160, 1119, e 1098 cm^{-1} , são atribuídas às vibrações de estiramento da ligação do grupo CO em ésteres: vibrações de acoplamento assimétricos C-C(=O)-O e O-C-C;
- A banda de baixa intensidade, relevante na região de 720 cm^{-1} é resultante da sobreposição da vibração de deformação angular tipo balanço de C-H do grupo metileno $\rho[\text{CH}_2]_n$, $n \geq 4$ e da vibração de deformação angular fora do plano sacudida de olefinas cis-disubstituído.

Porém, o espectro do óleo de mafurra apresenta mais uma banda de $\nu\text{C}=\text{O}$ em 1712 cm^{-1} proveniente dos ácidos graxos livres, razão pela qual se efetuou a catálise ácida do óleo antes da transesterificação, no processo de produção do biodiesel. Apesar desta ligeira diferença espectral existente entre o óleo de mafurra e os demais óleos, a análise visual dos espectros dos respectivos biodieseis, ilustrados na Figura 21 b, não apresentam diferenças consideráveis. Contudo, os espectros apresentados na Figura 21 b diferem dos espectros dos óleos, principalmente nas regiões de 1460 à 1425 cm^{-1} e 1200 à 1119 cm^{-1} , isto é, nos espectros do biodiesel: a banda em 1436 cm^{-1} é atribuído às vibrações de ligações C-H de olefinas cis – dissustituído (GAYDOU et al., 2011); as bandas que aparecem abaixo de 1244 cm^{-1} são bandas típicas da vibração axial assimétrica de alongamento de C–C da $\text{CC}(=\text{O})-\text{O}$

dos ésteres alifáticos saturado e as bandas em 1169 cm^{-1} e 1196 cm^{-1} referem-se deformação axial assimétrica do CO do grupo $\text{C}=\text{O}(\text{OR})$ dos ésteres alifáticos (SILVERSTEIN; WEBSTER, 1998).

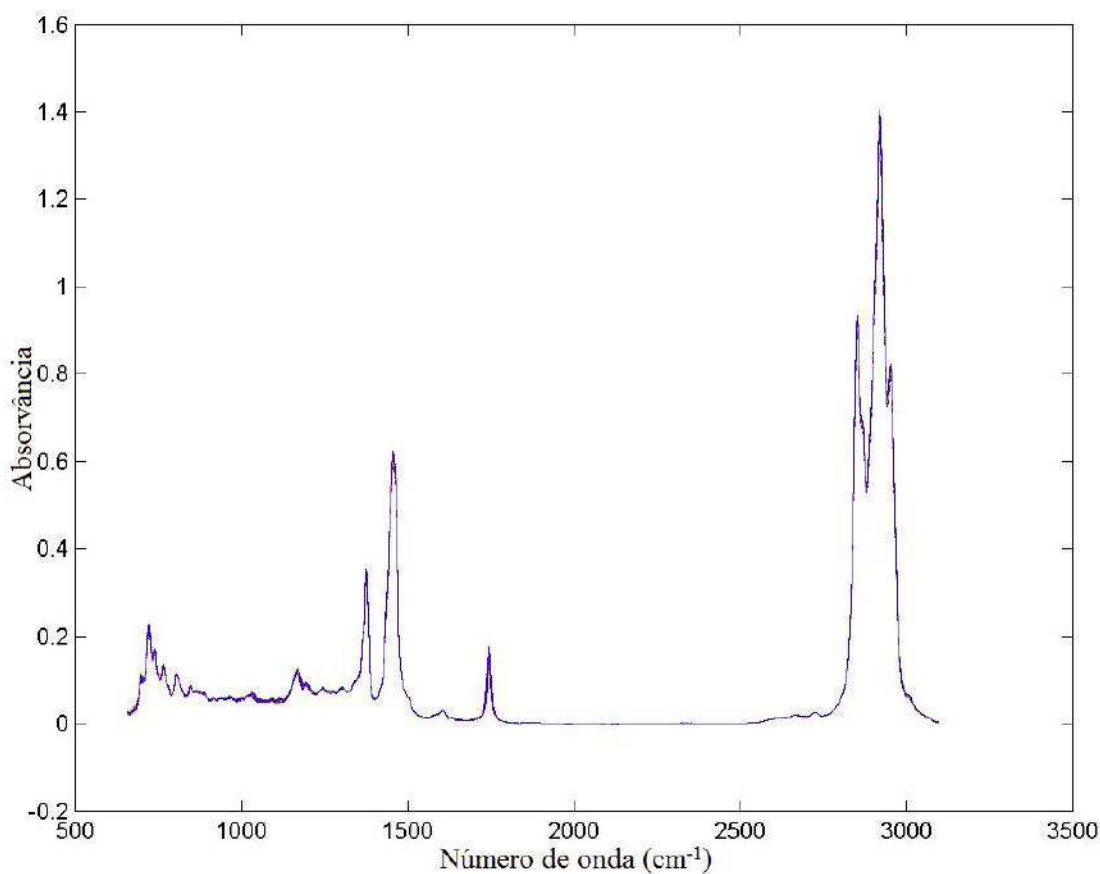
Figura 21 – Espectros MIR de: a) óleos de mafurra, moringa e algodão; b) biodieseis metílico de mafurra, moringa e algodão



Fonte: O autor.

Fazendo uma análise visual dos espectros MIR das 120 amostras de misturas B7 de BMMa, BMMo e BMA apresentados na Figura 22, pode-se notar que existe uma grande semelhança entre os perfis espectrais dos três tipos de biodiesel em misturas com diesel puro (B7). Fato este, que torna difícil classificar simultaneamente cada tipo de mistura de biodiesel/diesel de acordo com a sua matéria-prima usada na produção. Assim sendo, existindo muita informação nestes espectros que não é enxergada visualmente, o uso de ferramentas quimiométricas se torna extremamente útil para a substituição da análise visual dos espectros na classificação das três classes.

Figura 22 – Espectros MIR de 120 amostras de misturas B7 de BMMa, BMMo e BMA.



Fonte: O autor.

6.2. Modelos PLS

A quantidade das amostras de calibração e previsão utilizadas para a construção dos três modelos PLS (BMMa, BMMo e BMA) são apresentadas na Tabela 2. Como pode se observar nesta tabela, tanto o conjunto calibração quanto de previsão dos três modelos, o

número de amostra segue as exigências da norma ASTM E1655-55 (2012), segundo o qual estabelece que para o conjunto de calibração o número mínimo de amostras deve ser $6(k+1)$ para os dados centrados na média e, para o conjunto de previsão, $4k$. O valor de k refere-se ao número de VLs escolhido para construção do modelo PLS.

Tabela 2 – Parâmetros e valores de exatidão para os modelos PLS.

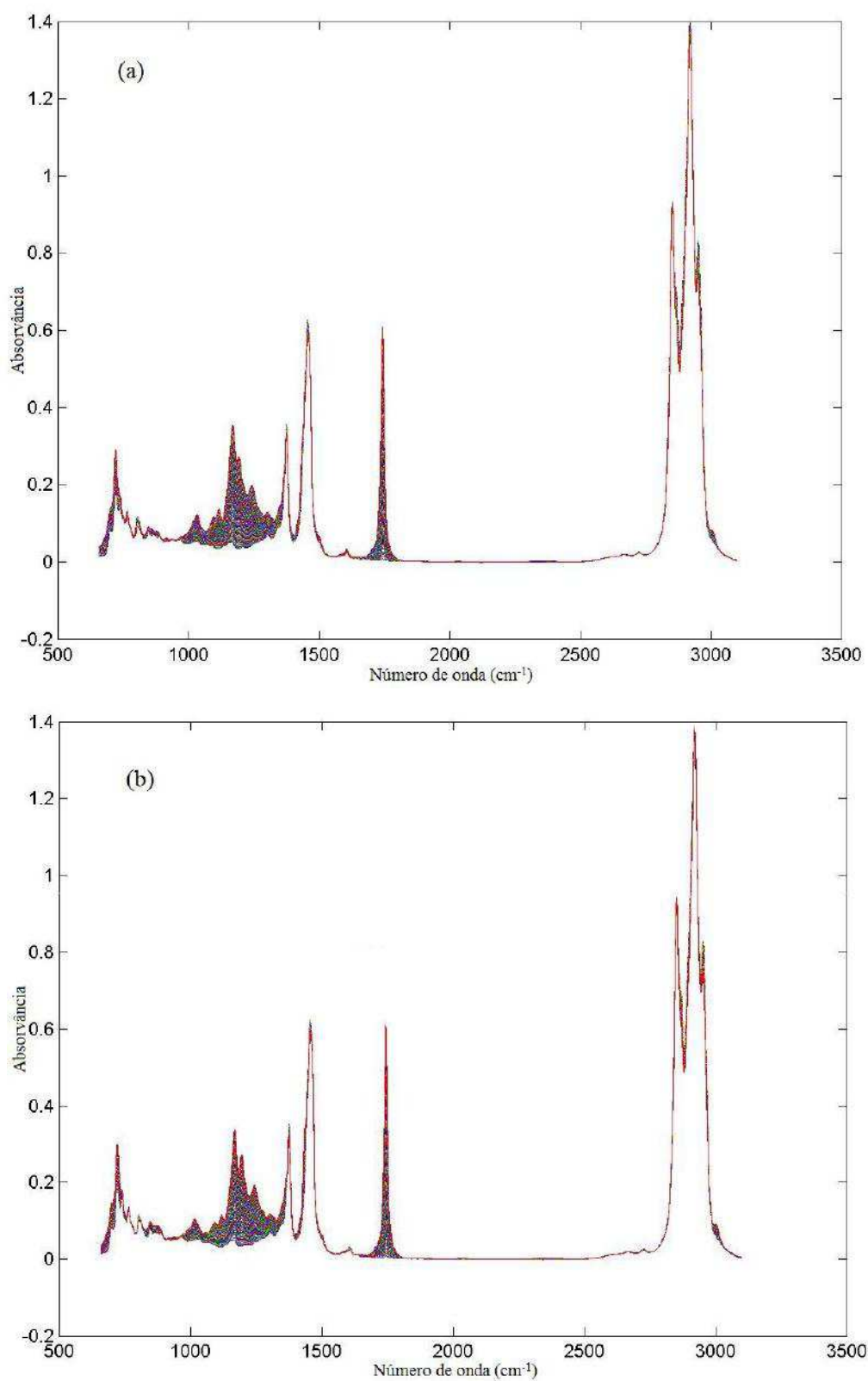
MODELO	BMMa	BMMo	BMA
Nº amostras de calibração	59	54	55
*Mínimo de amostras de calibração	30	42	36
Nº amostras de previsão	26	27	27
*Mínimo de amostras de previsão	16	24	20
VLs	4	6	5
Variância em X (%)	99,98	99,99	99,99
Variância em y (%)	100,00	100,00	100,00
RMSECV	0,0644	0,0648	0,0397
RMSEC	0,0568	0,0459	0,0331
RMSEP	0,0617	0,0485	0,0370

*Conforme norma ASTM E1655-05.

Fonte: O autor.

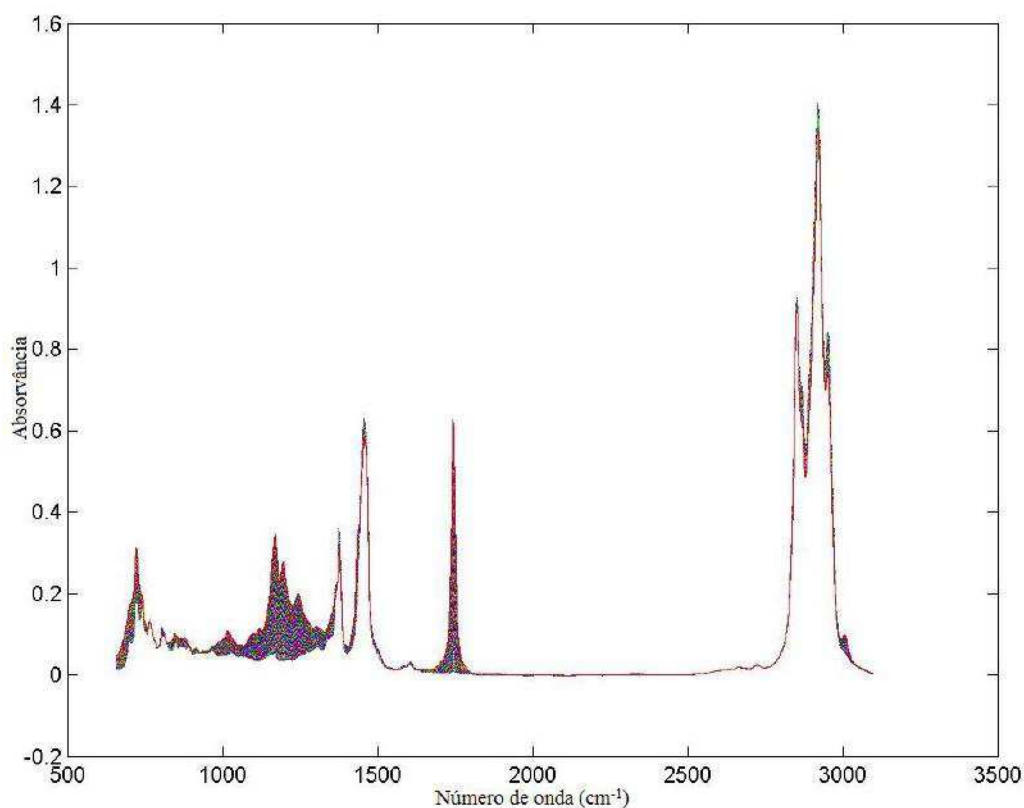
As Figuras 23 e 24 apresentam os espectros obtidos das amostras apresentadas na Tabela 2, em que é notório que para todos os conjuntos espectrais, as regiões de maior variação ocorrem na região de impressão digital de $1000\text{-}1300\text{ cm}^{-1}$ e na variação do grupo carbonila de $1735\text{-}1750\text{ cm}^{-1}$.

Figura 23 – Espectros MIR das amostras dos conjuntos de calibração e previsão dos biodieseis em mistura com diesel: (a) BMMa; (b) BMMo.



Fonte: O autor.

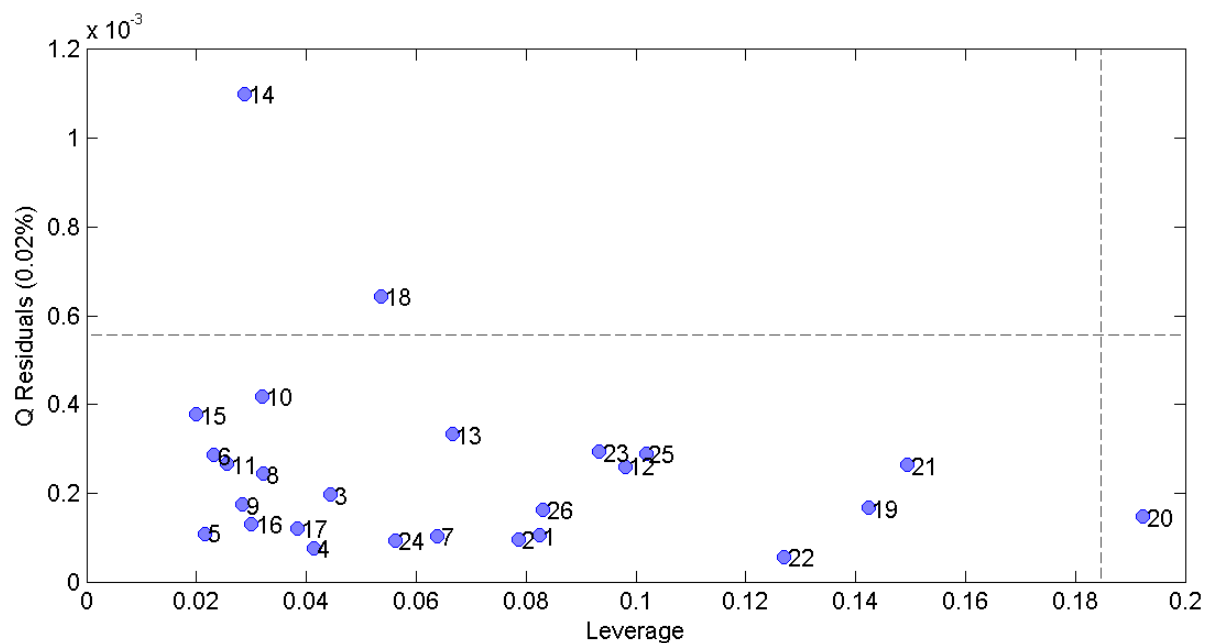
Figura 24 – Espectros MIR das amostras dos conjuntos de calibração e previsão do BMA em mistura com diesel.



Fonte: O autor.

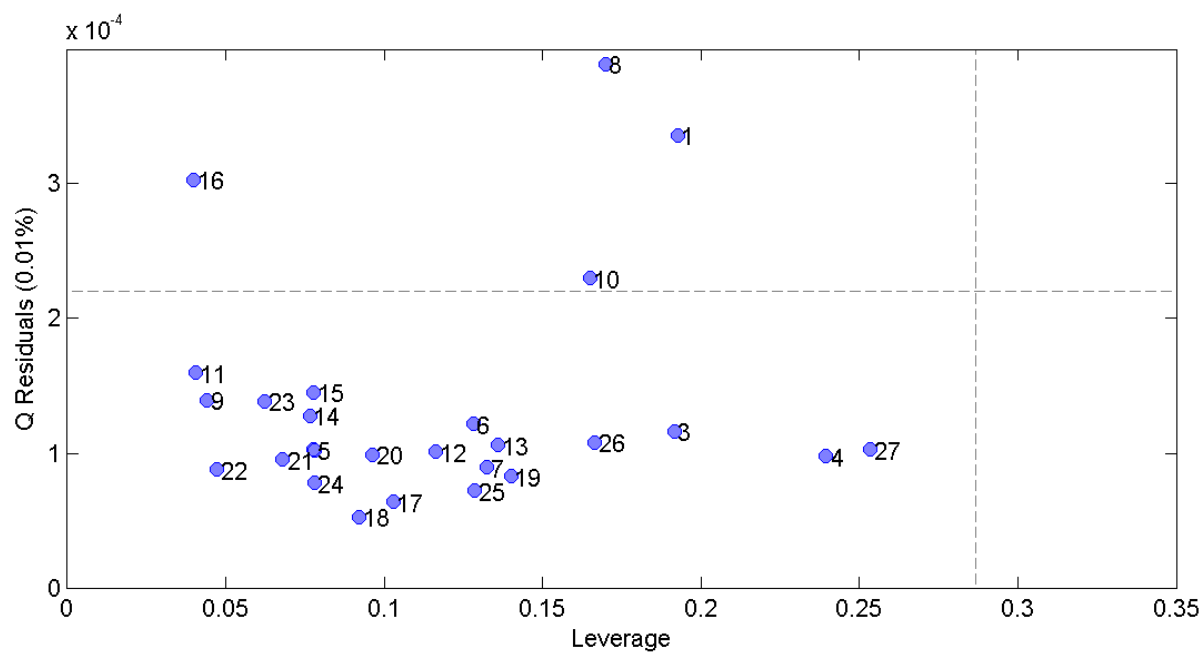
Os gráficos dos valores de *leverage* (o grau que uma amostra está distante da média do conjunto de dados) *versus* valores de resíduos (a parte não modelada nos dados espectrais ou nos valores de concentração analítica), apresentados nas Figuras 25, 26 e 27, foram analisados para identificar a presença de amostras anômalas nos três modelos. A linha tracejada horizontal representa o limite de resíduos, a 95% confiança, enquanto que a linha tracejada vertical representa a influência crítica do valor de *leverage*. Nestas figuras, observam-se que existem algumas amostras que apresentaram altos valores de resíduos, isto é, estão fora do limite de resíduos. Estas amostras não são consideradas de anômalas porque apresentam valores de *leverage* dentro do seu limite. O mesmo acontece com as amostras 20 do modelo BMMa e 1 do modelo BMA que apresentam altos valores de *leverage* e não são consideradas de anômalas porque apresentam valores de resíduos baixos, não prejudicando a qualidade do modelo.

Figura 25 – Gráfico de *leverage* versus *Q residuals* com 95% de confiança para o modelo PLS de BMMa.



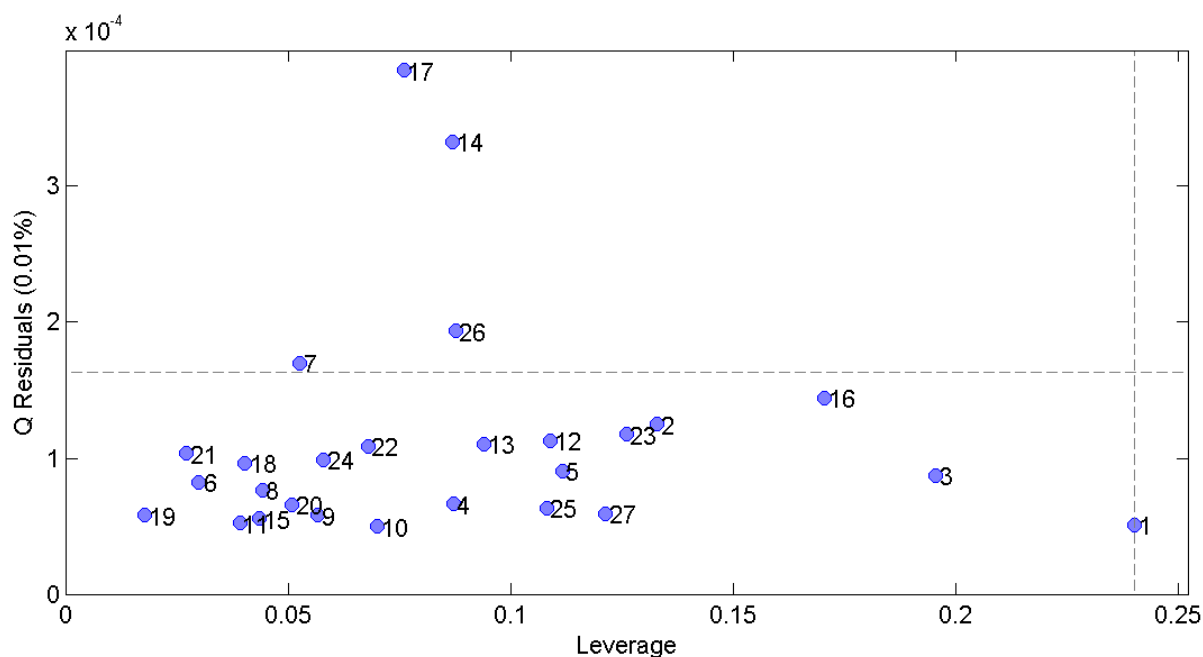
Fonte: O autor.

Figura 26 – Gráfico de *leverage* versus *Q residuals* com 95% de confiança para o modelo PLS de BMMo.



Fonte: O autor.

Figura 27 – Gráfico de *leverage* versus *Q residuals* com 95% de confiança para o modelo PLS de BMA.



Fonte: O autor.

Os modelos PLS de BMMA, BMMo e BMA foram construídos utilizando 4, 6 e 5 VLs, respectivamente. Estas variáveis latentes, representam 99,98, 99,99 e 99,99% da variância explicada na matriz **X**, respectivamente. As mesmas variáveis também representam 100,00, 100,00 e 100,00 % da variância explicada no vetor **y**, respectivamente.

Os valores de RMSECV, RMSEC e RMSEP dos três modelos, apresentados na Tabela 2, são relativamente baixos e considerados aceitáveis de acordo com a norma ABNT NBR 15568, ao qual dita que o erro médio quadrático de previsão não deve ser superior que 1 %. Os valores de RMSEC são inferiores que os valores de RMSEP, indicando uma boa concordância entre os dois parâmetros, ou seja, que não há sobreajuste (LIU et al., 2008).

Buiatte, et al., (2016) quantificaram o teor de biodiesel metílico de algodão na mistura biodiesel/diesel usando Espectrometria no Infravermelho Médio associado ao método quimiométrico PLS, em que os valores de RMSECV, RMSEC e RMSEP foram 0,04, 0,05 e 0,05, respectivamente. Portanto, esses valores são superiores que os valores obtidos neste trabalho.

Todos os modelos PLS desenvolvidos não apresentaram erros sistemáticos porque os valores de $t_{calculado}$ é menor do que $t_{crítico}$, para graus de liberdade (GL) igual ao número de amostras de previsão, como se pode ver na Tabela 3. Assim, os modelos estão de acordo com

as recomendações da norma ASTM E1655-05, a qual sugere a realização de teste-*t* para avaliar se o “*bias*” incluso no modelo é significativo.

Tabela 3 – Resultados das figuras de mérito para o modelo PLS.

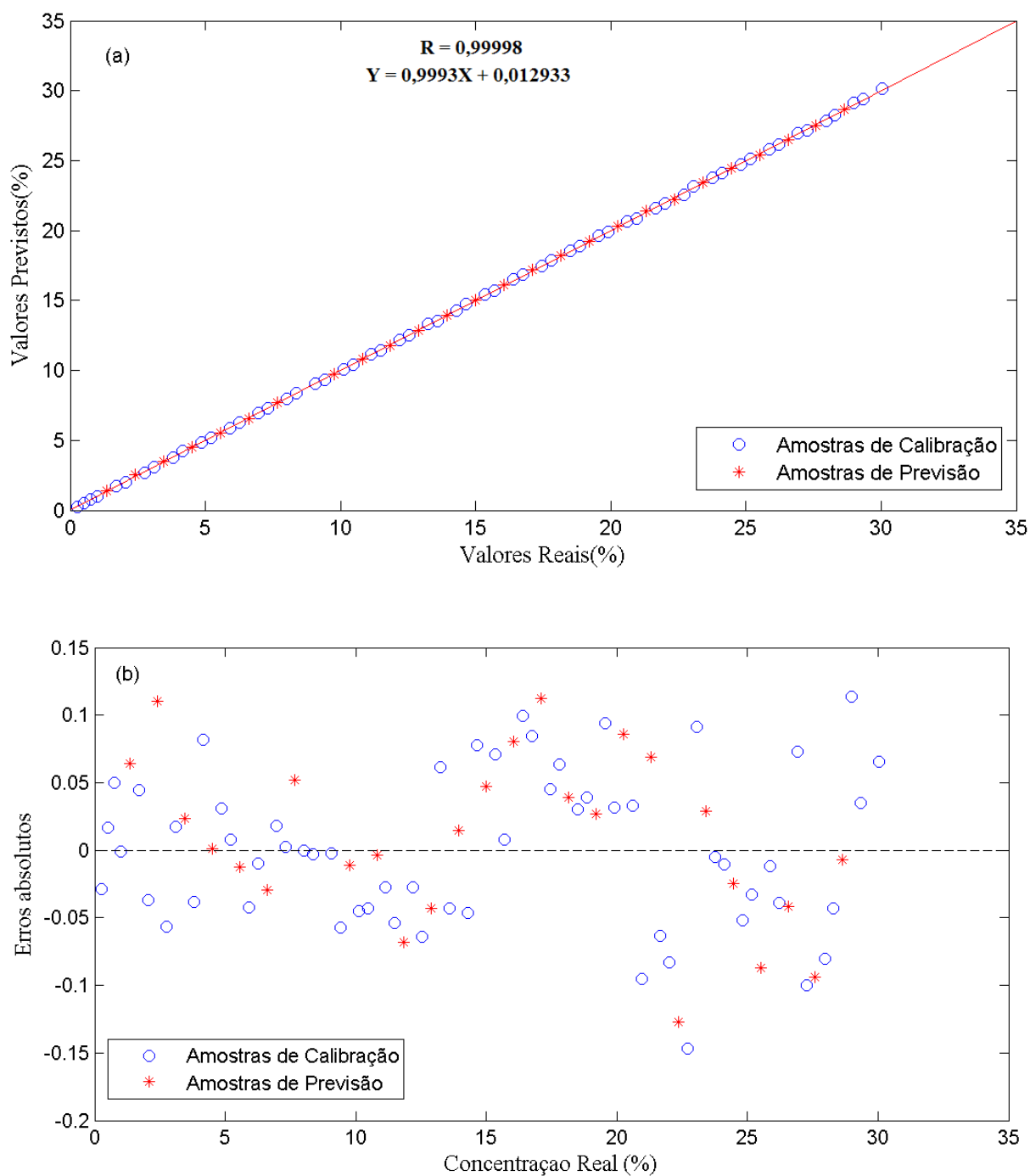
Figuras de mérito	Valores dos modelos PLS		
	BMMa	BMMo	BMA
Sensibilidade analítica / % (v/v)	49,458	25,348	33,407
Sensibilidade / % (v/v) ⁻¹	0,0877	0,4708	0,6580
Selectividade	0,2376	0,1551	0,2194
Limite de detecção / % (v/v)	0,0667	0,1302	0,0988
Limite de quantificação / % (v/v)	0,2022	0,3945	0,2993
SDV	0,0624	0,0490	0,0360
<i>t</i> _{bias}	0,6519	0,6599	1,5901
<i>t</i> _{crítico}	2,0555	2,0518	2,0518

Fonte: O autor.

As Figuras 28 a, 29 a e 30 a mostram a avaliação do ajuste do modelo por meio da correlação dos valores reais e previstos dos conjuntos de calibração e previsão dos modelos PLS de BMMa, BMMo e BMA, respectivamente, em que é visível a alta correlação existente entre os valores reais e previstos dos conjuntos de calibração e previsão, com um coeficiente de correlação superior a 0,99, o que significa que qualquer erro sistemático presente é desprezível, comparando ao erro aleatório.

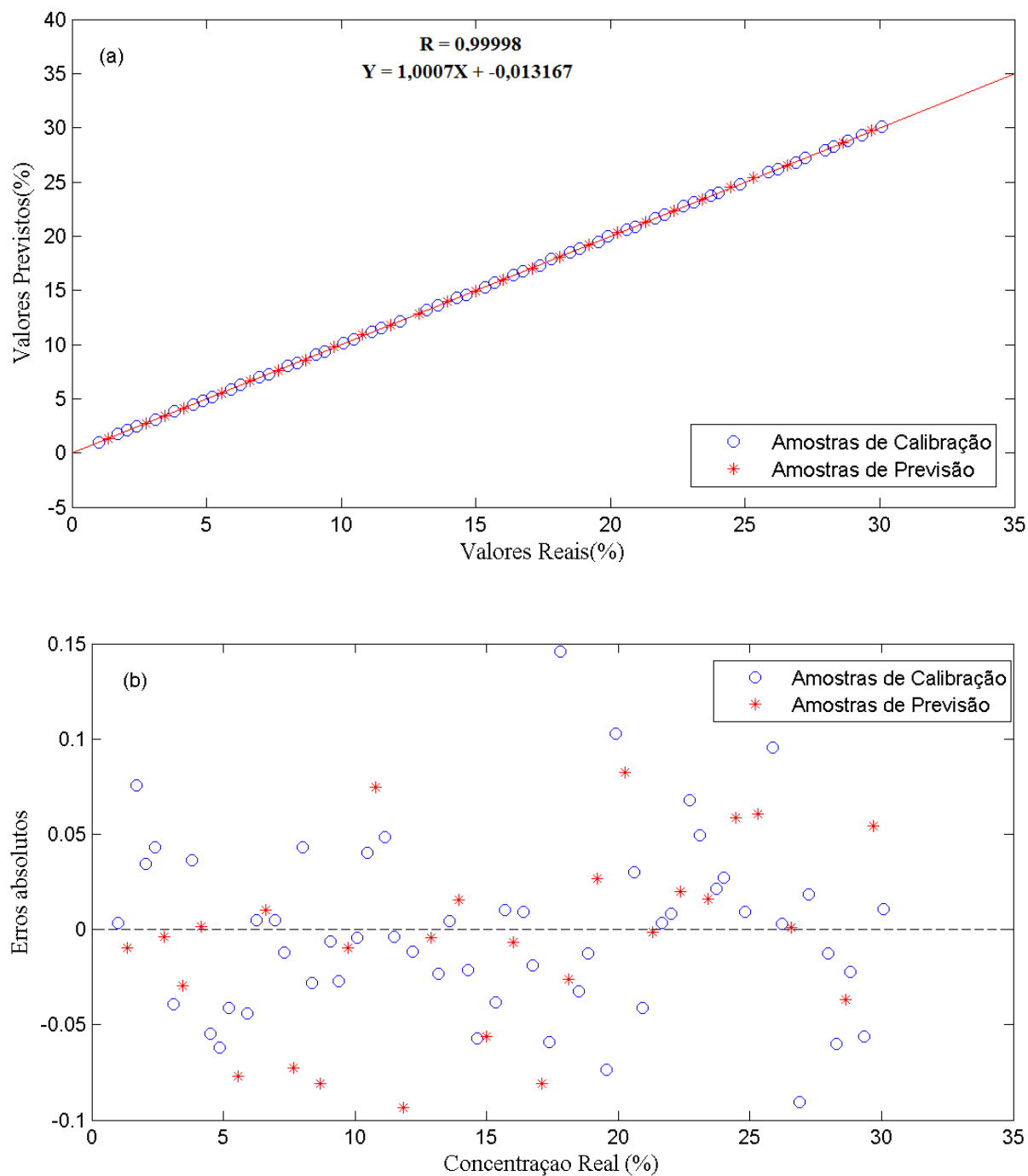
Importa referir que, um alto coeficiente de correlação é necessário, mas não é suficiente para garantir a boa qualidade do modelo, portanto, os resíduos também devem ser analisados (FERREIRA, 2015). Assim sendo, as Figuras 28 b, 29 b e 30 b ilustram os gráficos de valores de concentrações reais *versus* erros absolutos dos modelos PLS de BMMa, BMMo e BMA, respectivamente, com distribuição uniforme numa faixa horizontal, indicando assim, que os resíduos têm um comportamento aleatório e que sua variância é constante no domínio da concentração, ou seja, demonstram que a curva está bem ajustada.

Figura 28 – (a) Gráfico de ajuste do modelo PLS de BMma e (b) gráfico de resíduo do modelo PLS de BMma.



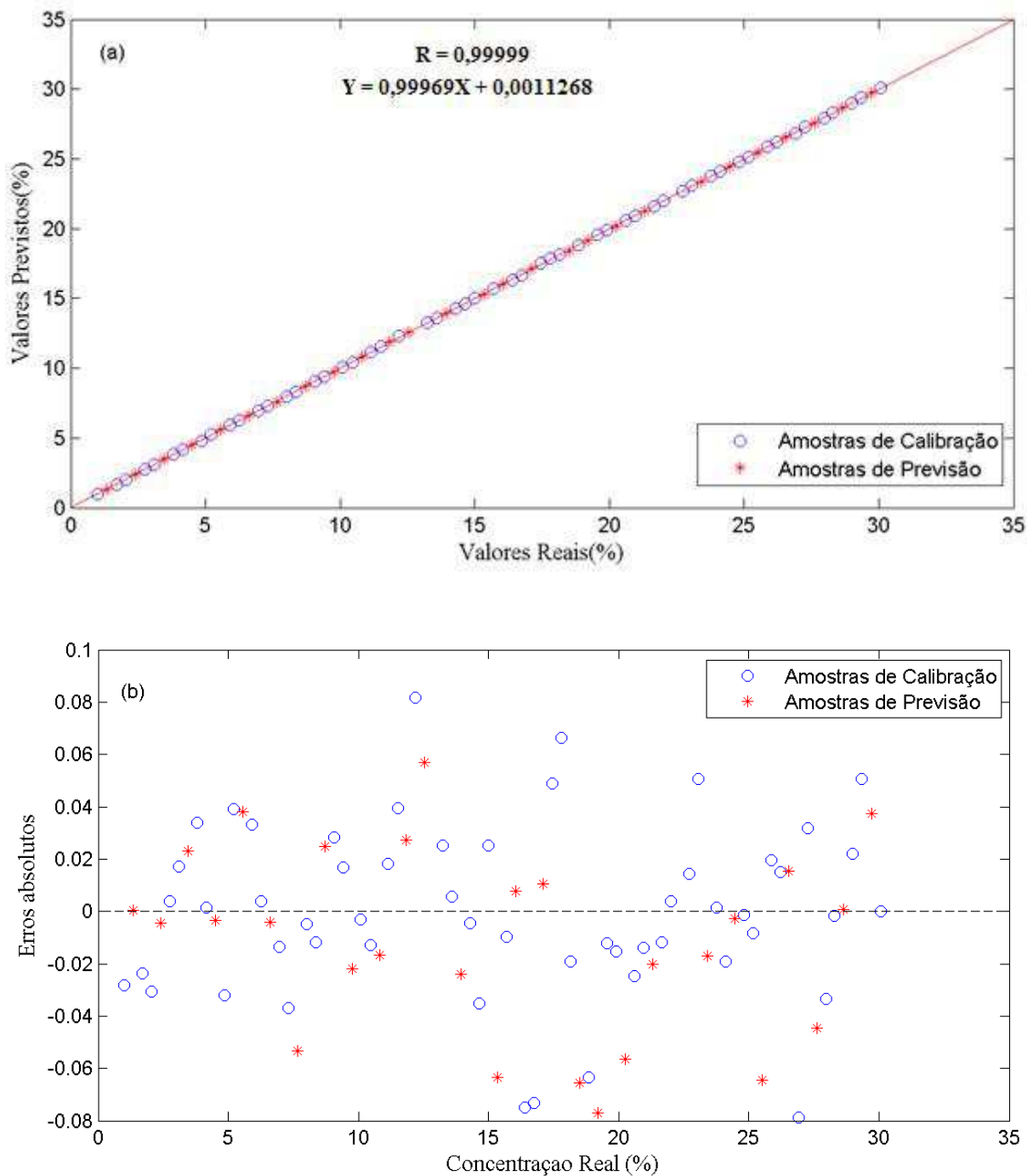
Fonte: O autor.

Figura 29 – (a) Gráfico de ajuste do modelo PLS de BMMo e (b) gráfico de resíduo do modelo PLS de BMMo.



Fonte: O autor.

Figura 30 – (a) Gráfico de ajuste do modelo PLS de BMA e (b) gráfico de resíduo do modelo PLS de BMA.



Fonte: O autor.

6.3. Modelos PLS-DA

Para a classificação do teor de BMMa, BMMo e BMA na mistura B7/BX com o diesel, foram construídos três modelos PLS-DA utilizando 6, 6 e 5 VLs, respectivamente. Estas variáveis latentes, representam 99,94; 99,89 e 99,95% da variância explicada na matriz **X**, respectivamente. As mesmas variáveis também representam 90,86; 87,74 e 93,95% da variância explicada no vetor **y**, respectivamente.

A Tabela 4 apresenta os resultados dos parâmetros de classificação obtidos pelos modelos, onde nota-se que os valores de RMSEC, RMSECV e RMSEP são relativamente baixos e considerados aceitáveis de acordo com a literatura (WISE et al., 2006), mostrando que existe uma boa concordância entre os parâmetros RMSEC e RMSEP, pois o RMSEC apresenta valor inferior ao RMSEP.

Tabela 4 – Parâmetros de classificação obtidos pelo modelo PLS-DA.

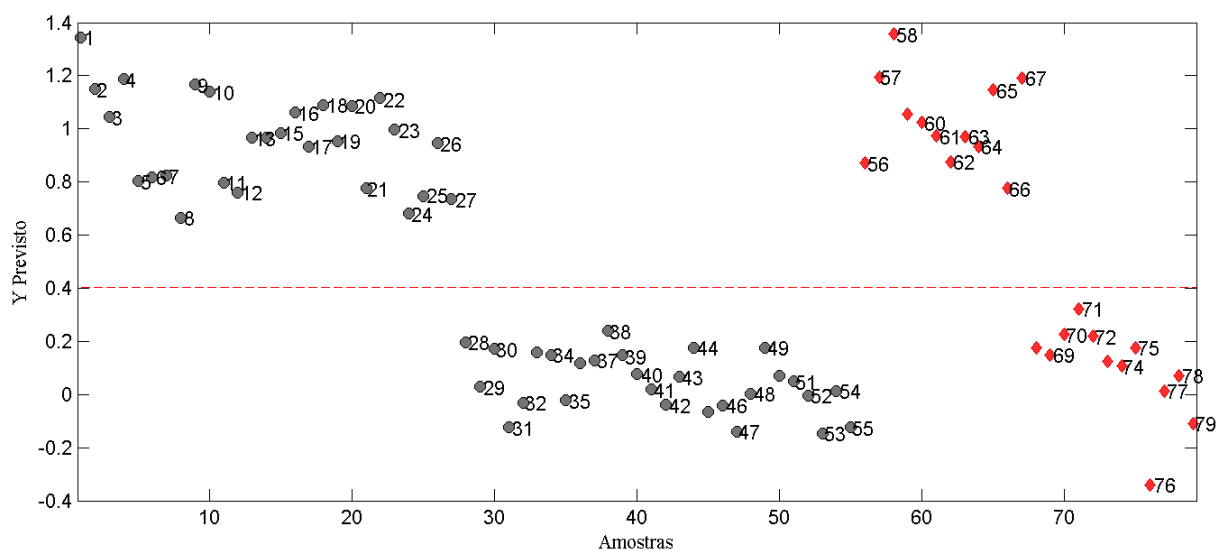
Parâmetros	Valores dos modelos PLS-DA		
	BMMa	BMMo	BMA
RMSECV (% v/v)	0,2013	0,2177	0,1537
RMSEC (% v/v)	0,1511	0,1750	0,1230
RMSEP (% v/v)	0,1781	0,1807	0,1671
Especificidade (Cal, CV, Pred)	1	1	1
Sensibilidade (Cal, CV, Pred)	1	1	1
Taxa de falso positivo (%)	0	0	0
Taxa de falso negativo (%)	0	0	0
Variáveis Latentes	6	6	5
Variância em X (%)	99,94	99,89	99,95
Variância em y (%)	90,86	87,74	93,95

Fonte: O autor.

Portanto, os resultados destes parâmetros mostram a boa eficiência dos modelos PLS-DA em classificar as amostras de B7 e BX de BMMa, BMMo e BMA, como se pode ver nas Figuras 31, 32 e 33, as amostras B7 do conjunto de treinamento (●) e de teste (◆) estão

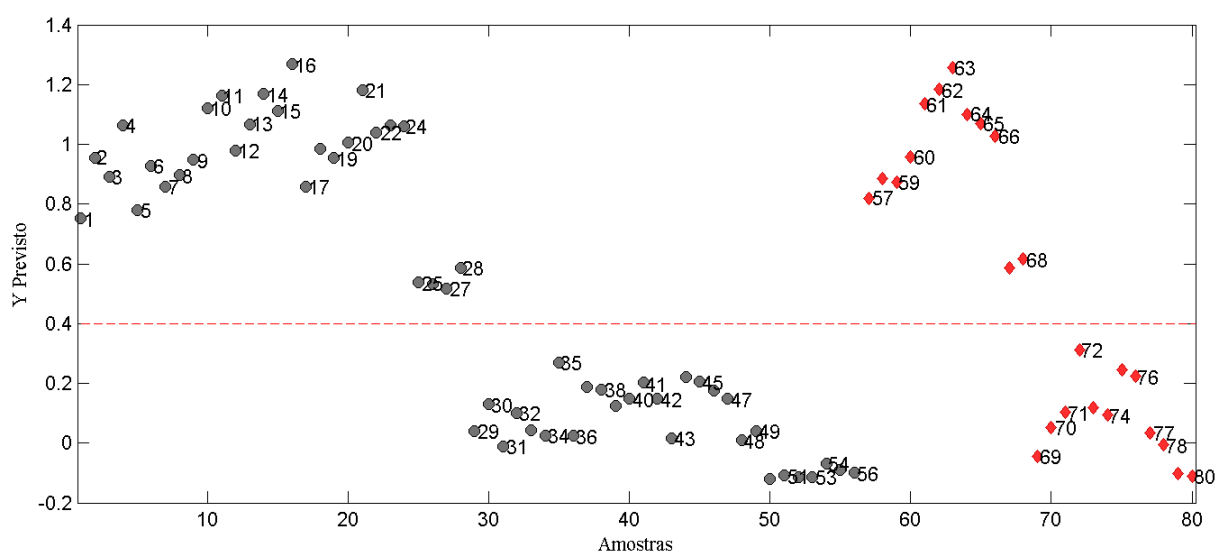
acima do *threshold*, separando-se das amostras BX do conjunto de treinamento (●) e de teste (◆) que estão abaixo do *threshold*.

Figura 31 – Estimativas do modelo PLS-DA para a classificação de amostras de B7 e BX do BMMa.



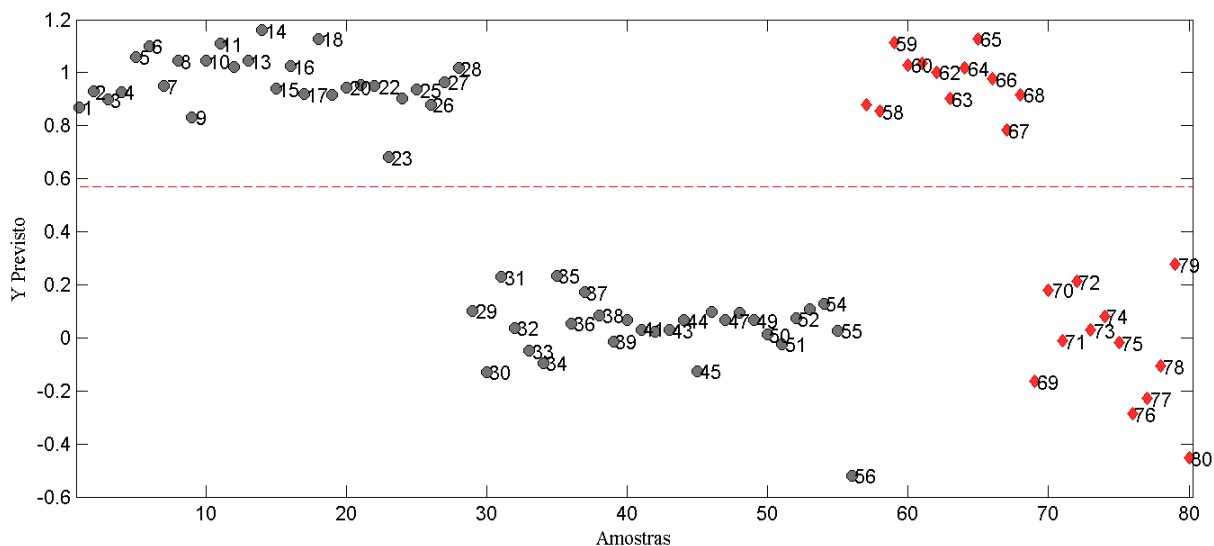
Fonte: O autor.

Figura 32 - Estimativas do modelo PLS-DA para a classificação de amostras de B7 e BX do BMMo.



Fonte: O autor.

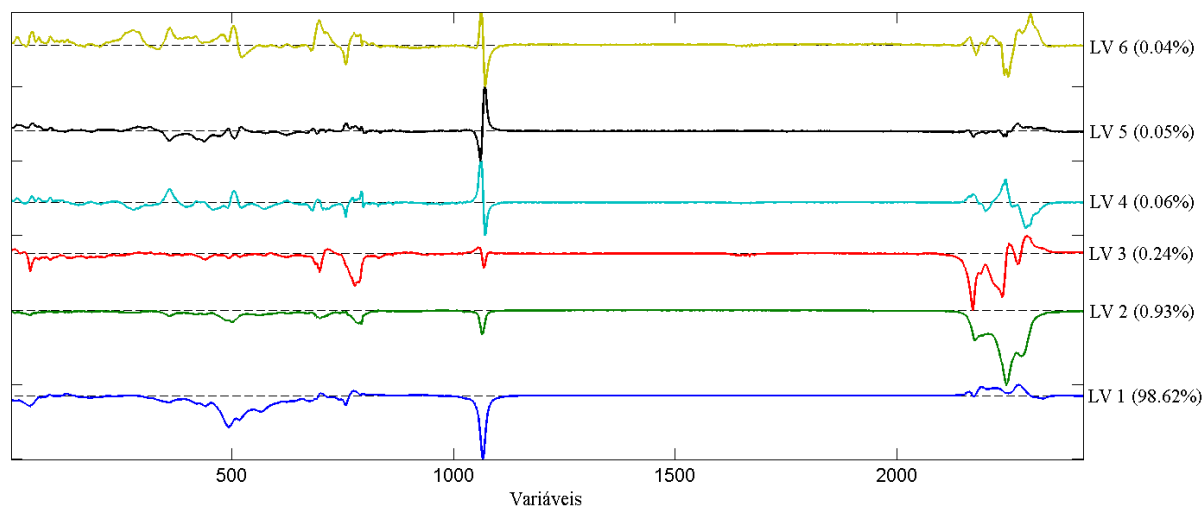
Figura 33 – Estimativas do modelo PLS-DA para a classificação de amostras de B7 e BX do BMA.



Fonte: O autor.

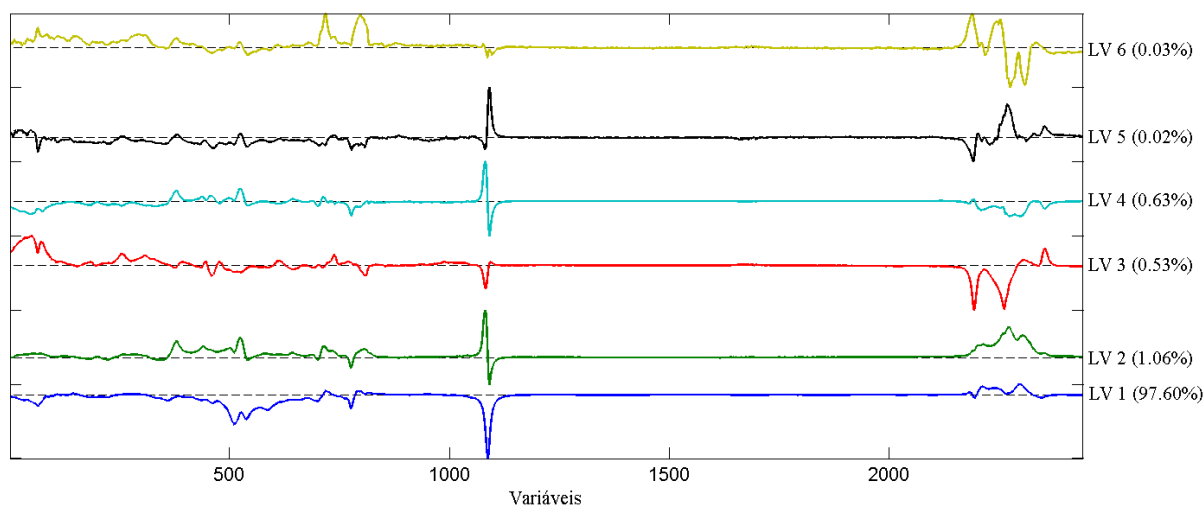
A análise discriminante foi feita com base no perfil dos espectros MIR da mistura B7/BX de BMMA, BMMo e BMA, apresentados nas Figuras 34, 35 e 36. Nestas figuras, nota-se as regiões espectrais que contribuíram para a diferenciação entre as classes. No entanto, ao analisar os pesos das variáveis latentes, verifica-se que as LVs aplicadas para os modelos de BMMA, BMMo e BMA descrevem 99,94, 99,87 e 99,95% das variáveis que influenciaram para a discriminação das classes, respectivamente. Estas variáveis são encontradas nas seguintes regiões espectrais: de 2820-2979 cm^{-1} (2160-2319 variáveis) que são atribuídas às vibrações de estiramento assimétrico da ligação C-H do grupo metila ($-\text{CH}_3$) e às vibrações de estiramento simétrico e assimétrico da ligação C-H do grupo metileno ($-\text{CH}_2$); de 1730-1760 cm^{-1} (1070-1100 variáveis) é atribuído à vC=O do acilgliceróis; de 1360-1475 cm^{-1} (700-815 variáveis) que são atribuídas às deformações angular assimétrica e simétrica da ligação C-H de grupo metila ($-\text{CH}_3$) (SILVERSTEIN; WEBSTER, 1998; BARBOSA, 2008).

Figura 34 – Gráfico dos pesos de LV1, LV2, LV3, LV4, LV5 e LV6 *versus* variáveis do modelo PLS-DA de BMMa.



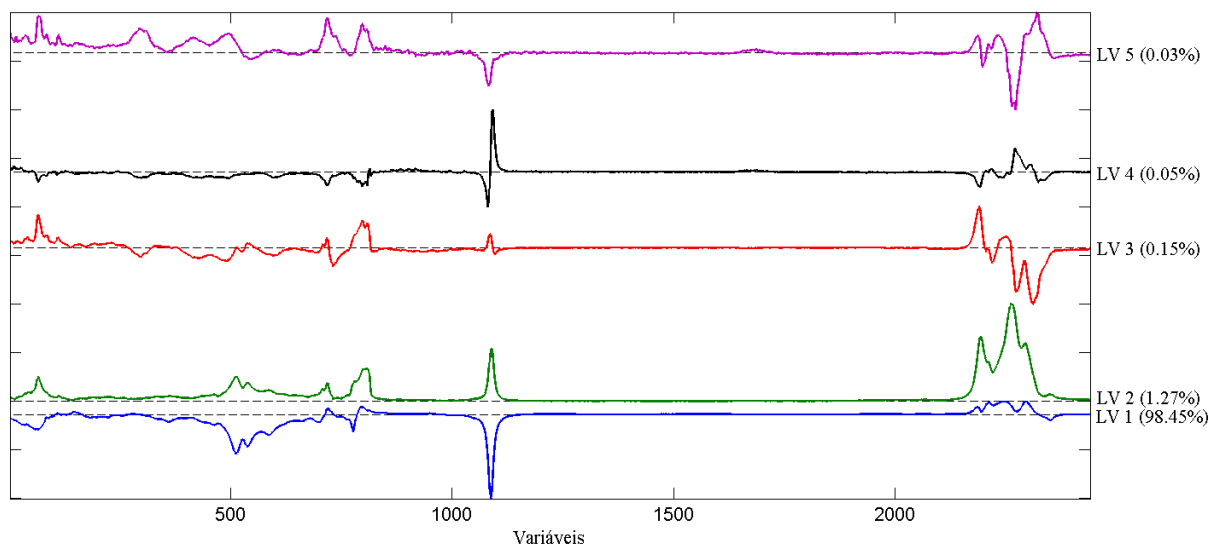
Fonte: O autor.

Figura 35 – Gráfico dos pesos de LV1, LV2, LV3, LV4, LV5 e LV6 *versus* variáveis do modelo PLS-DA de BMMo.



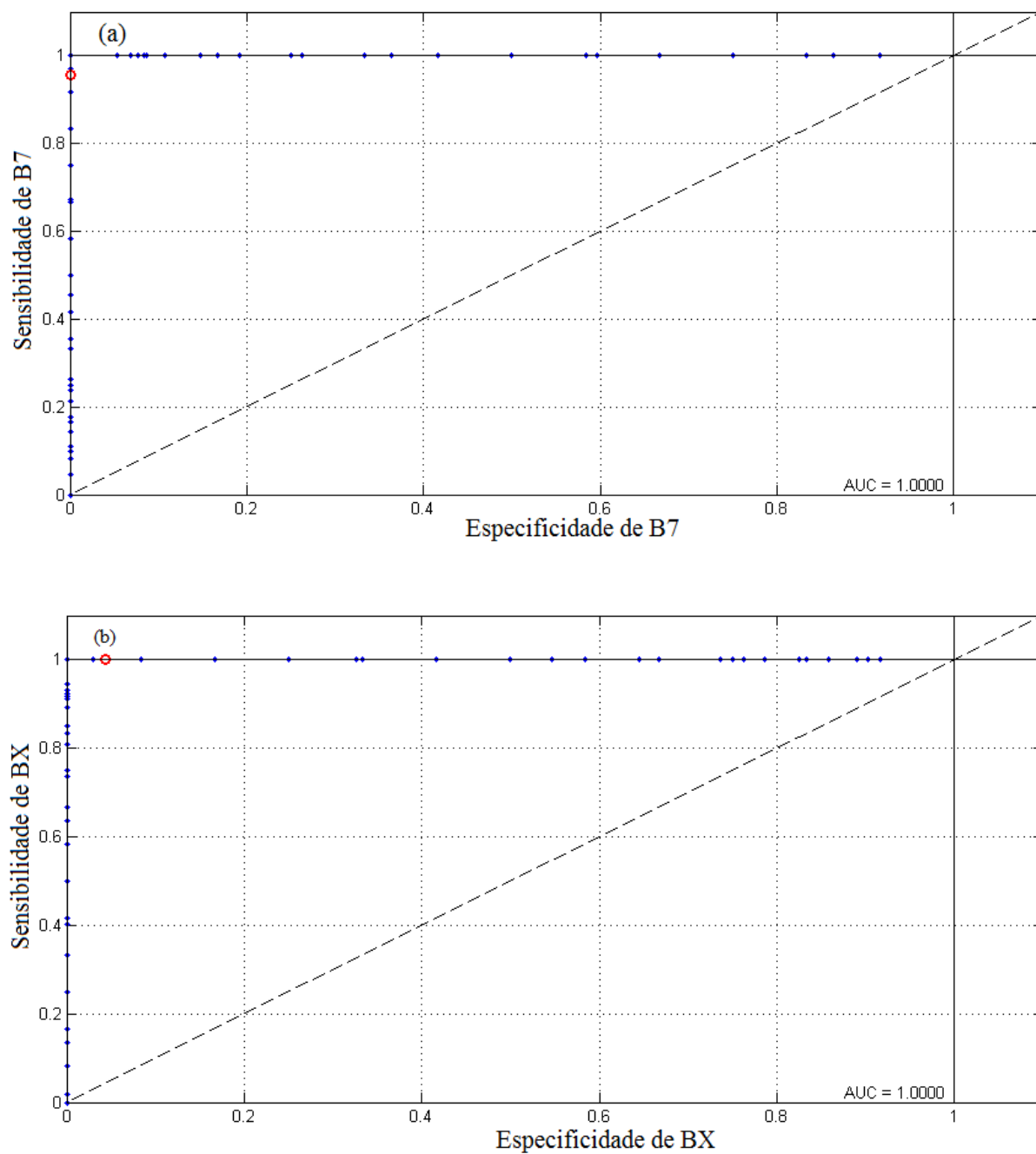
Fonte: O autor.

Figura 36 – Gráfico dos pesos de LV1, LV2, LV3, LV4 e LV5 *versus* variáveis do modelo PLS-DA de BMA.



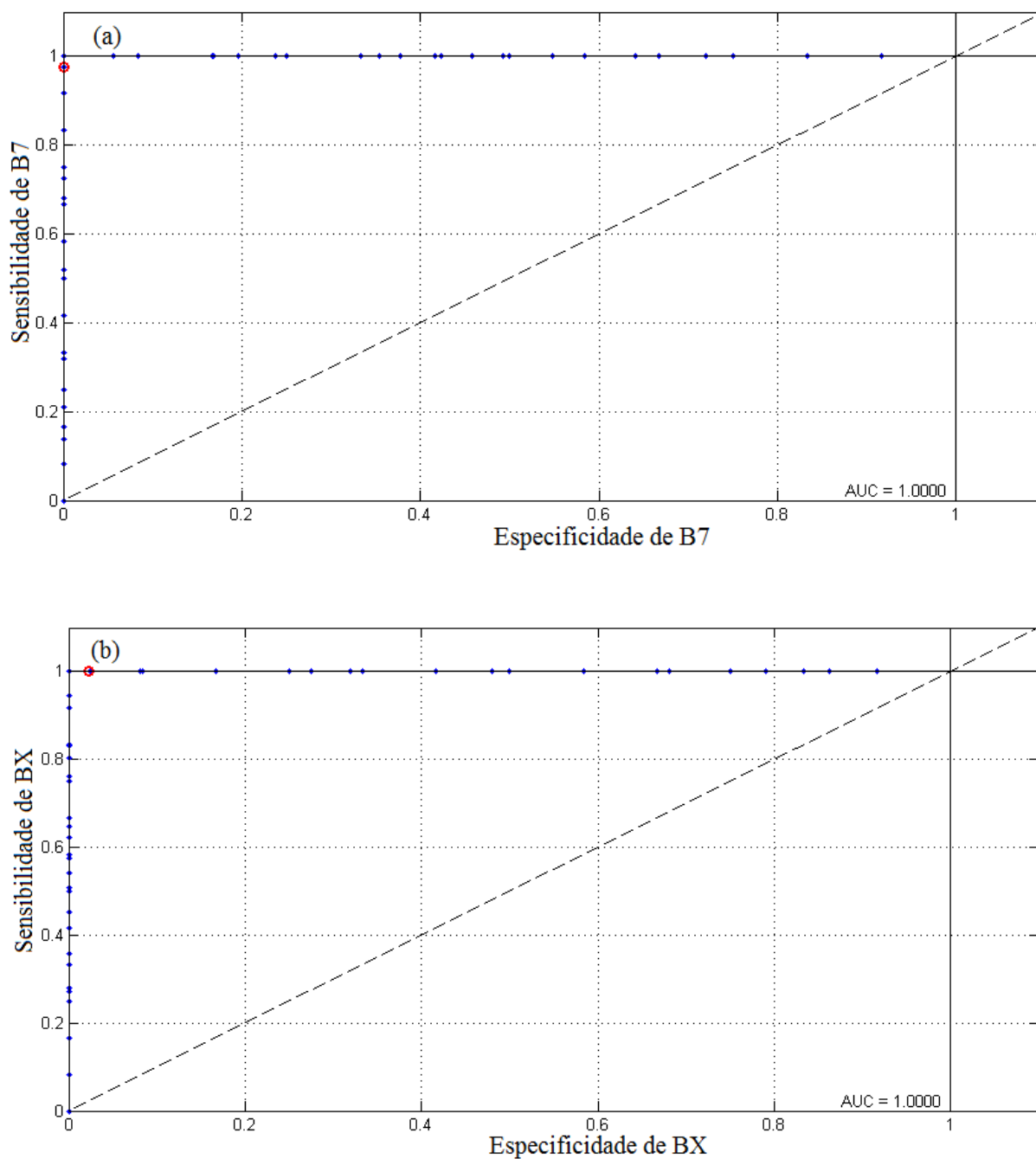
Para avaliar a eficiência dos três modelos PLS-DA (BMMa, BMMo e BMA), foram utilizados os critérios de estatísticas de respostas verdadeiras, a saber: sensibilidade e especificidade (ALMEIDA et al., 2013). Como se pode observar na Tabela 4 e nas Figuras 37, 38 e 39, a sensibilidade e a especificidade dos três modelos foram iguais a 1, o que quer dizer que todas as amostras a serem previstas como pertencentes a classe foram realmente previstas na classe e todas amostras que não pertencem a classe, foram previstas fora da classe.

Figura 37 – Gráfico de sensibilidade *versus* especificidade do modelo PLS-DA de BMMa para amostras de: (a) B7 e (b) BX.



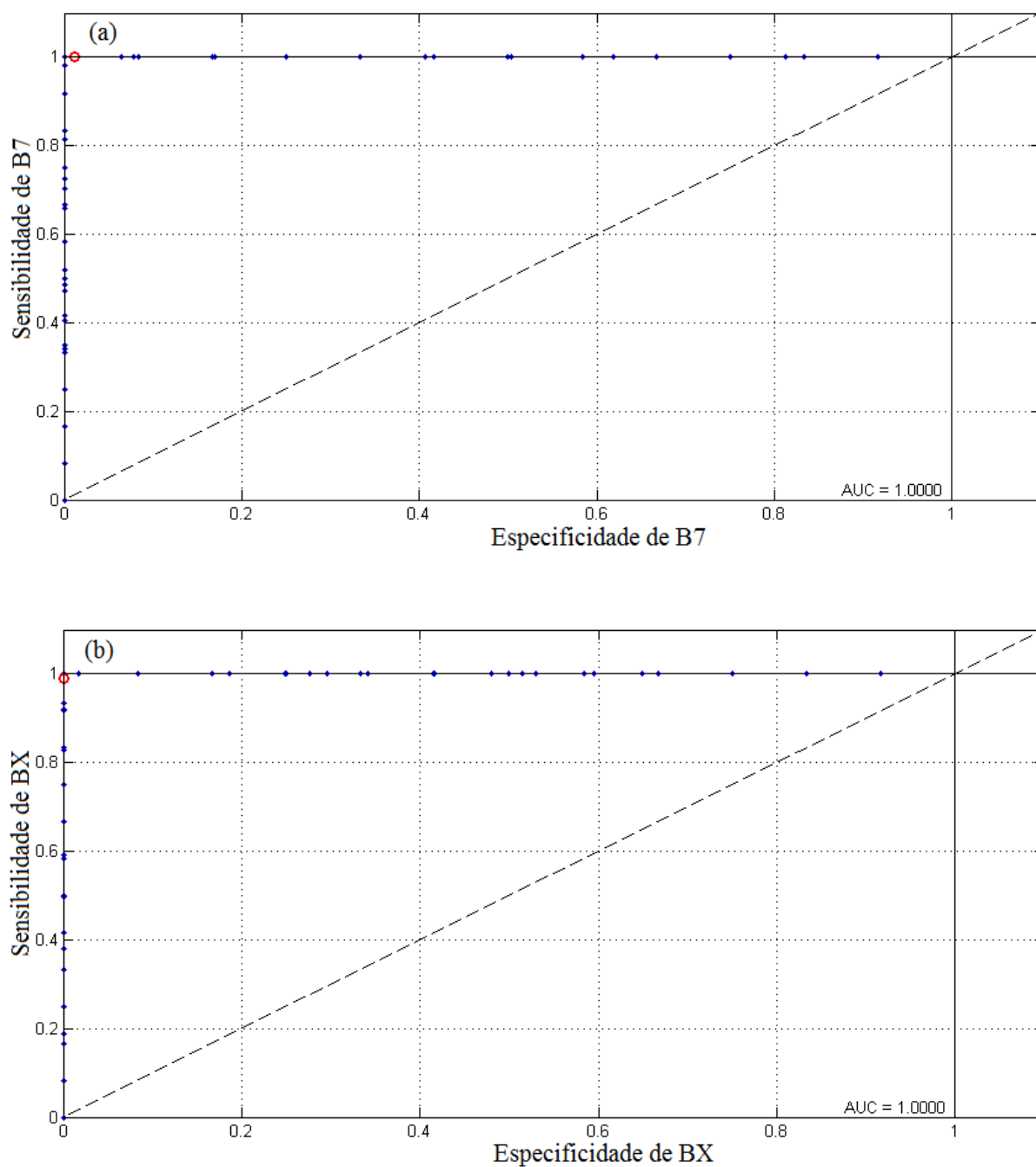
Fonte: O autor.

Figura 38 – Gráfico de sensibilidade *versus* especificidade do modelo PLS-DA de BMMo para amostras de: (a) B7 e (b) BX.



Fonte: O autor.

Figura 39 – Gráfico de sensibilidade *versus* especificidade do modelo PLS-DA de BMA para amostras de: a) B7 e b) BX.



Fonte: O autor.

6.4. Modelo PLS2-DA

O Modelo PLS2-DA desenvolvido para classificar simultaneamente as três classes de misturas de biodiesel/diesel (B7) de acordo com a sua matriz (mafurra, moringa e algoão) foi validado utilizando amostras do conjunto de teste, em que os resultados dos parâmetros de classificação obtidos pelo modelo são apresentados na Tabela 5. As 8 variáveis latentes escolhidas para o modelo, representam 98,52% da variância explicada na matriz **X** e 92,27% da variância na matriz **Y**.

Os valores de RMSEC, RMSECV e RMSEP são relativamente baixos, próximos e considerados aceitáveis de acordo com a literatura (WISE et al., 2006), mostrando que existe uma boa concordância entre os parâmetros RMSEC e RMSEP, ou seja, não há sobreajuste, pois o RMSEC apresenta valor inferior ao RMSEP.

Mazivila et al., (2015) classificaram diferentes tipos de biodiesel de acordo com a matéria prima usada e a rota de produção, utilizando Espectrometria no Infravermelho Médio associado ao método quimiométrico PLS2-DA, em que os parâmetros de classificação obtidos foram iguais com os obtidos neste trabalho.

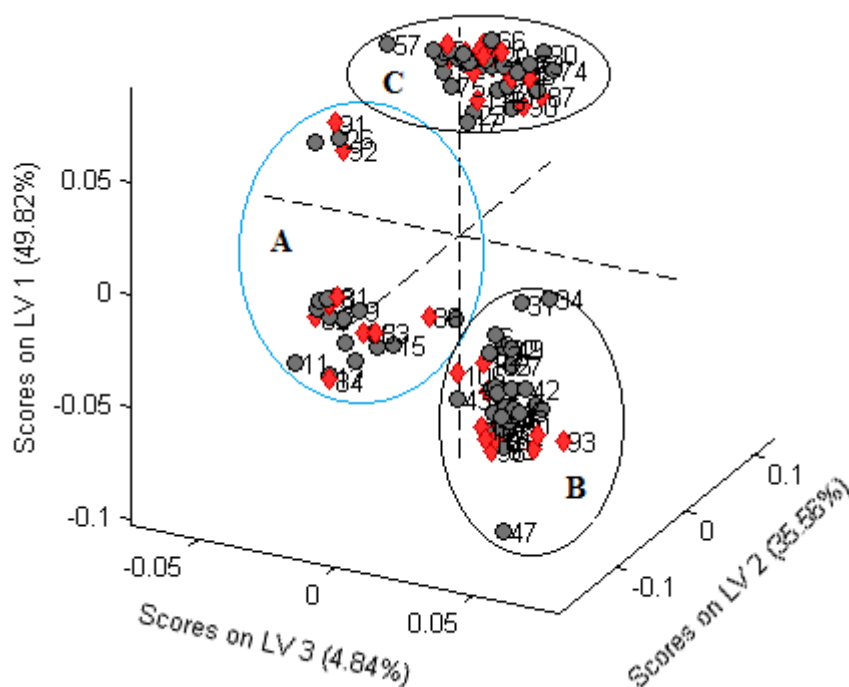
Tabela 5 – Parâmetros de classificação obtidos pelo modelo PLS2-DA.

Parâmetros	Valores dos modelos PLS2-DA		
	BMMa	BMMo	BMA
RMSECV (% v/v)	0,1147	0,1557	0,2202
RMSEC (% v/v)	0,0796	0,1216	0,1744
RMSEP (% v/v)	0,0865	0,1574	0,2120
Especificidade (Cal, CV, Pred)	1	1	1
Sensibilidade (Cal, CV, Pred)	1	1	1
Taxa de falso positivo (%)	0	0	0
Taxa de falso negativo (%)	0	0	0
Variáveis Latentes		8	
Variância em X (%)		98,52	
Variância em y (%)		92,27	

Fonte: O autor.

Os resultados destes parâmetros mostram a boa eficiência do modelo PLS2-DA em classificar simultaneamente a mistura de biodiesel/diesel (B7) de mafurra, moringa e algodão. A Figura 40 mostra a separação das amostras por classes, de acordo com a sua matéria-prima usada para a produção do biodiesel, em que se pode notar que o modelo PLS2-DA conseguiu classificar simultaneamente as três classes: A- mistura de biodiesel/diesel (B7) de moringa, B- mistura de biodiesel/diesel (B7) de mafurra e C- mistura de biodiesel/diesel (B7) de algodão.

Figura 40 – Estimativas do modelo PLS2-DA para a classificação das três classes de B7 de BMMa, BMMo e BMA.



Fonte: O autor.

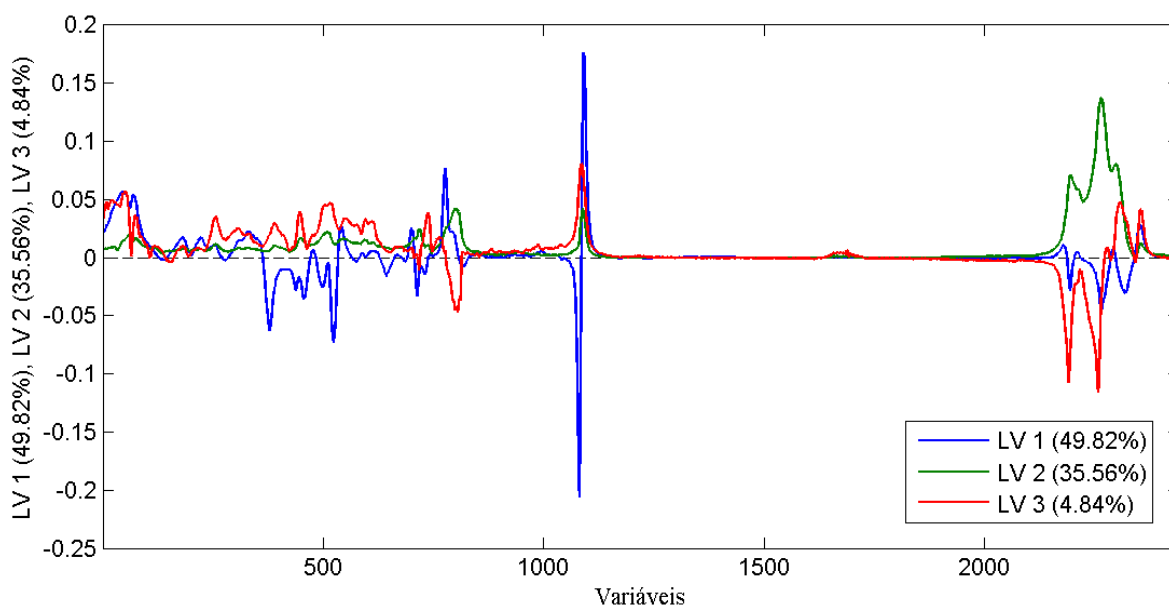
A análise discriminante foi feita com base no perfil dos espectros MIR da mistura B7 de BMMo, BMMa e BMA, apresentados na Figura 41. Nesta figura, é possível notar as bandas que contribuem para a diferenciação entre as classes. Ao analisar os pesos das variáveis latentes, verifica-se que a LV1 descreve 49,82% das variáveis que influenciaram para a discriminação das classes. Essas variáveis são encontradas nas seguintes regiões espectrais: de 1750 à 1747 cm^{-1} que são atribuídas à $\nu\text{C}=\text{O}$; de 1442 à 1432 cm^{-1} atribuídos às vibrações de ligações C-H de olefinas cis – dissustituído; de 1364 à 1357 cm^{-1} atribuído a deformação angular simétrica de tipo tesoura de CH_2 e de 725 à 720 cm^{-1} atribuído à vibração

de deformação angular tipo balanço de C-H do grupo metileno $\rho[\text{CH}_2]_n$, $n \geq 4$.

Outras variáveis que também contribuíram para a discriminação das classes são descritas em 35,56% pela LV2 e, foram encontradas nas seguintes regiões espectrais: de 2957 à 2853 cm^{-1} devido às vibrações de estiramentos simétricos e assimétricos da ligação C-H de grupos metila ($-\text{CH}_3$), metileno ($-\text{CH}_2$); de 1470 à 1450 cm^{-1} atribuído à deformação angular assimétrica de tipo tesoura de CH_3 .

A LV3 descreve 4,84% das variáveis que influenciaram para a discriminação das classes, encontradas nas regiões espectrais de 3015 à 3005 cm^{-1} atribuído às vibrações de estiramentos da ligação C-H de carbonos insaturados; de 1652 à 1645 cm^{-1} atribuído ao alongamento de C=C e de 1293 à 900 cm^{-1} são atribuídas às vibrações de estiramento da ligação do grupo CO em ésteres, a saber: vibrações de acoplamento assimétricos C-C(=O)-O e O-C-C.

Figura 41 – Gráfico dos pesos de LV1, LV2 e LV3 *versus* variáveis do modelo PLS2-DA.

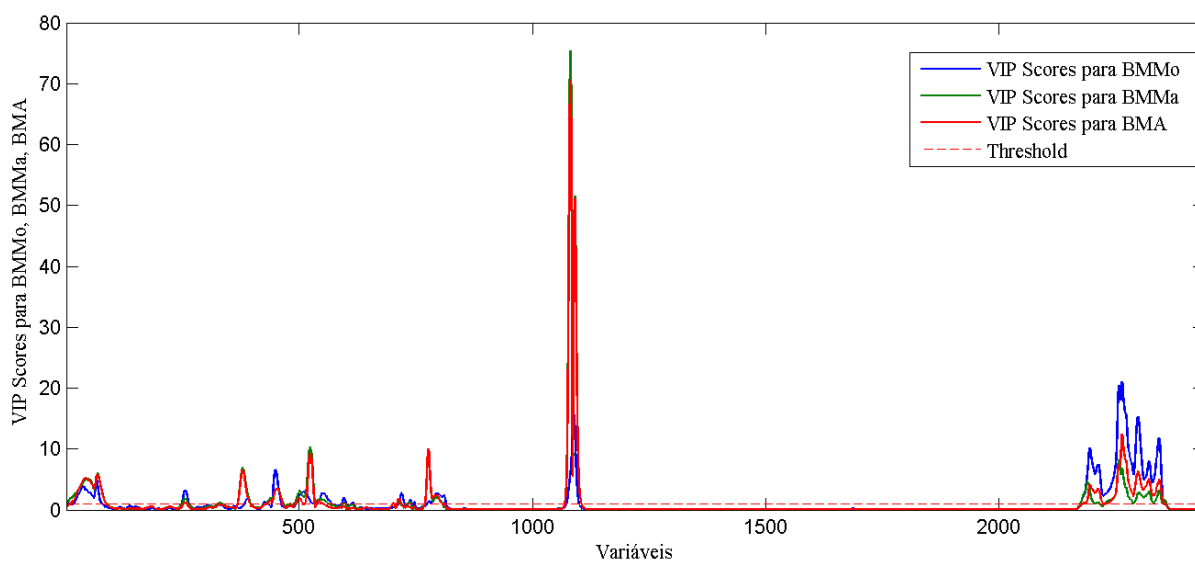


Fonte: O autor.

Para avaliar a importância das variáveis em projeção (VIP), foi analisado o gráfico de VIP scores *versus* variáveis, apresentado na Figura 42, onde os valores de VIP que estão acima da linha tracejada (*threshold*) informam as regiões espectrais mais significativas para a classificação das amostras, isto é, as variáveis que possibilitam um melhor desempenho do

modelo PLS2-DA. Assim sendo, as regiões espectrais de: 3008 à 2848 cm^{-1} , 1484 à 1446 cm^{-1} e 1290 à 1204 cm^{-1} foram importantes para a mistura B7 de BMMo; 1760 à 1733 cm^{-1} e 1190 à 1153 cm^{-1} foram importantes para a mistura B7 de BMMa; 1444 à 1431 cm^{-1} e 1363 à 1359 cm^{-1} foram importantes para a mistura B7 de BMA.

Figura 42 – VIP scores *versus* variáveis.



Fonte: O autor.

Para avaliar a eficiência do modelo de classificação PLS2-DA, foram utilizados os critérios de estatísticas de respostas verdadeiras, a saber: sensibilidade e especificidade (Almeida *et al.*, 2013). Como se pode observar na Tabela 5, a sensibilidade e a especificidade do modelo foram iguais a 1, o que quer dizer que todas as amostras a serem previstas como pertencentes a classe foram realmente previstas na classe e todas amostras que não pertencem a classe, foram previstas fora da classe.

7. CONCLUSÕES

A combinação da técnica de Espectrometria MIR com os métodos quimiométricos PLS-DA e PLS permitiu o desenvolvimento de modelos de classificação e previsão estáveis, robustos e simples de interpretar.

A eficiência dos modelos de previsão (PLS) e classificação (PLS-DA) do teor dos biodieseis metílicos de Mafurra, Moringa e Algodão foram concordantes com os requisitos das figuras de mérito e nos parâmetros de classificação, respectivamente.

No caso dos modelos PLS, foi visível uma alta correlação existente entre valores de concentração real e previsto dos conjuntos de calibração e previsão, com um coeficiente de correlação superior que 0,99 e com erros relativamente baixos.

Para o caso dos modelos PLS-DA, não foi notado a presença de amostras falsos positivos e nem falsos negativos, consequentemente, foi verificado a classificação correta de todas as amostras de calibração e previsão, isto é, os parâmetros de sensibilidade e especificidade foram iguais a 1.

Todas as metodologias desenvolvidas no trabalho foram validadas de acordo com as normas ABNT NBR 15568, ASTM D7371 e ASTM E1655-05

Desta forma, estes resultados demonstram a alta eficiência que os modelos PLS-DA e PLS têm para classificar e prever, respectivamente, o teor dos biodieseis metílicos de Mafurra, Moringa e Algodão na mistura com diesel. Assim, as metodologias analíticas desenvolvidas são viáveis e podem ser usadas para o controle quantitativo e qualitativo destes combustíveis.

REFERÊNCIAS

- ACHTEN, W. M. J. et al. Jatropha bio-diesel production and use. **Biomass and Bioenergy**, v. 32, p. 1063–1084, 2008.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS – ANP. **Boletim mensal do monitoramento da qualidade dos combustíveis líquidos automotivos**. Rio de Janeiro, 2016a. Disponível em: <<http://www.anp.gov.br/?pg=82480&m=&t1=&t2=&t3=&t4=&ar=&ps=&1473367496451>>. Acesso em: 08 set. 2016.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS – ANP. **Boletim mensal do biodiesel**. 2016b. Disponível em: <www.anp.gov.br/?dw=73583>. Acesso em: 07 set. 2016.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS – ANP. **Leilões de biodiesel**. Rio de Janeiro, 2016c. Disponível em: <<http://www.anp.gov.br/?pg=82625&m=&t1=&t2=&t3=&t4=&ar=&ps=&1473369189659>>. Acesso em: 08 set. 2016.
- ALMEIDA, M. R. et al. Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation. **Talanta**, v. 117, p. 305–311, 2013.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **E1655-05: Standard practices for infrared multivariate quantitative analysis**. West Conshohocken, 2012.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D7371: Standard test method for the determination of biodiesel (fatty acid methyl esters) content in diesel fuel oil using mid-infrared spectroscopy (FTIR-ATR-PLS method)**. West Conshohocken, 2007.
- ARNDT, C. et al. **Biofuels, Poverty, and Growth : A Computable General Equilibrium Analysis of Mozambique**. IFPRI Discussion Paper. INSTITUTE, I. F. P. R. Washington, D.C., USA: 00803: 28 p. 2008.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR 15568: Biodiesel – Determinação do teor de biodiesel em óleo diesel por Espectrometria na região do infravermelho médio**. Rio de Janeiro, 2008.
- AVINASH, A.; SUBRAMANIAM, D.; MURUGESAN, A. Bio-diesel - A global scenario **Renewable and Sustainable Energy Reviews**, v. 29, p. 517–527, 2014.
- BAMBO, T. F. **Cooperação Sul-Sul: O acordo Brasil-Moçambique na área de biocombustíveis**. 2014. 80. (Mestrado). Instituto de relações internacionais, Universidade de São Paulo, São Paulo 2014.
- BANKOVIĆ-ILIĆ, I. B.; STAMENKOVIĆ, O. S.; VELJKOVIĆ, V. B. Biodiesel production from non-edible plant oils. **Renewable and Sustainable Energy Reviews**, v. 16, p. 3621–3647, 2012.

- BARBOSA, L. C. D. A. **Espectrometria no Infravermelho- na caracterização de compostos orgânicos**. Viçosa: Editora UFV, 2008. 189 p.
- BARKER, M.; RAYENS, W. Partial least squares for discrimination. **Journal of Chemometrics**, v. 17, p. 166-173, 2003.
- BATIDZIRAI, B.; FAAIJ, A. P. C.; SMEETS, E. Biomass and bioenergy supply from Mozambique. **Energy for Sustainable Development** n. 2006, p. 55-81, 2006.
- BORIN, A.; POPPI, R. J. Multivariate quality control of lubricating oils using Fourier Transform Infrared Spectroscopy. **Journal of the Brazilian Chemical Society**, v. 15, n. 4, p. 570-576, 2004.
- BRASIL. **DECRETO S/Nº, de 02 de Julho de 2003 - DOU 3.7.2003: Institui Grupo de Trabalho Interministerial**. Brasília, 2003a.
- BRASIL. **DECRETO S/Nº, de 23 de Dezembro de 2003 - DOU 24.12.2003: Institui a Comissão Executiva Interministerial**. Brasília, 2003b.
- BRASIL. **Lei Nº 11.097. Introdução da matriz energética na matriz energética brasileira**. Brasil, 2005.
- BRASIL. **Lei nº 13.033, de 24 de Setembro de 2014. Dispõe sobre a adição obrigatória de biodiesel ao óleo diesel comercializado com o consumidor final; altera as Leis nos 9.478, de 6 de agosto de 1997, e 8.723, de 28 de outubro de 1993; revoga dispositivos da Lei no 11.097, de 13 de janeiro de 2005; e dá outras providências**. Brasília, 2014.
- BRASIL. Ministério de Minas e Energia. **Balanço Energético Nacional 2015: Ano base 2014**. Rio de Janeiro, 2015.
- BRASIL. **Resolução ANP Nº. 14**. ANP. RANP 14/2012, 2012.
- BUIATTE, J. E. et al. Qualitative and quantitative monitoring of methyl cotton biodiesel content in biodiesel/diesel blends using MIR Spectroscopy and chemometrics tools. **Journal of the Brazilian Chemical Society**, v. 27, n. 1, p. 84-90, 2016.
- CARDELLO, H. M. A. B. et al. Clarificação convencional de óleo de semente de Algodão (*Gossypium hirsutum*). **Alimentos e Nutrição**, v. 6, p. 77-87, 1995.
- CHANG, C. C.; WAN, S. W. China's motor fuels from tung oil. **Industrial & Engineering Chemistry**, v. 39, p. 1543-1548, 1947.
- CHOWHURY, D. H. et al. Indian vegetable fuel oils for diesel engines. **Gas Oil Power**, v. 37, p. 80-85, 1942.
- CLAYDEN, J. et al. **Organic Chemistry**. Oxford University Press: New York, 2011.