



**UFU**

**Universidade Federal de Uberlândia**

**Instituto de Química**

**Programa de Pós-Graduação em Química**

**LQT- Laboratório de Quimiometria do Triângulo**

**Sarmento Júnior Mazivila**

**Controle de qualidade de biodiesel de soja e pinhão manso usando espectroscopia no infravermelho médio e análise discriminante por quadrados mínimos parciais**

**Uberlândia**

**2015**

Sarmiento Júnior Mazivila

Controle de qualidade de biodieséis de soja e pinhão manso usando espectroscopia no infravermelho médio e análise discriminante por quadrados mínimos parciais

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

Área de Concentração: Química Analítica

Orientador: Prof. Dr. Waldomiro Borges Neto.

**Uberlândia**  
**2015**



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 206/PPQUI.

DATA: 16/1/2015

DISCENTE: Sarmento Júnior Mazivila

MATRÍCULA: 11312QMI011

TÍTULO DO TRABALHO: Controle de qualidade de biodiesel de soja e pinhão manso usando espectroscopia no Infravermelho Médio e análise discriminante por quadrados mínimos parciais.

ÁREA DE CONCENTRAÇÃO: Química

LINHA DE PESQUISA: Espectroanalítica Aplicada

PROJETO DE PESQUISA DE VINCULAÇÃO: Aplicação de espectroscopia no Infravermelho e análise discriminante por quadrados mínimos parciais para controle de qualidade de biocombustíveis.

Às quatorze horas do dia dezesseis de janeiro do ano dois mil e quinze, no Auditório Prof. Manuel Gonzalo Hernández Terrones, Bloco 5I, no Campus Santa Mônica, reuniu-se a Banca Examinadora composta pelos Professores Doutores Luiz Carlos Moutinho Pataca, do Instituto de Geoinformação e Tecnologia de Minas Gerais, Alam Gustavo Trovó 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.

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 40 minutos e lavrada a presente ata, que após lida e aprovada, foi assinada pela Banca Examinadora.

Luiz Carlos Moutinho Pataca  
Dr. Luiz Carlos Moutinho Pataca

IGTEC/MG

Alam Gustavo Trovó  
Prof. Dr. Alam Gustavo Trovó

IQUFU

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



**Este trabalho é dedicado àquele a quem sou eternamente grato por ter me conferido a visão de um quimiometrista, o meu orientador Prof. Dr. Waldomiro Borges Neto.**

## **AGRADECIMENTOS**

A Deus pela essência da vida, por me conceder saúde, em especial, pela possibilidade de realizar este trabalho, na presença dele o Pai Altíssimo, pois sem Ele nada seria possível.

Ao Prof. Dr. Waldomiro Borges Neto, pela oportunidade da realização deste trabalho, pela valiosa orientação, apoio, incentivo e confiança que depositou em mim, desde o dia em que aceitou me orientar estando ainda em Moçambique.

A minha namorada Selda Alberto Fungulane, pelo apoio durante este momento espinhoso da minha formação.

A minha família, meus pais Sarmiento Simione Mazivila e Ana Eduardo Matsombe, meus irmãos Benedito Sarmiento Mazivila, Luís Sarmiento Mazivila, Nocia Sarmiento Mazivila, Paulina Sarmiento Mazivila, Jorge Sarmiento e Cecília Sarmiento Mazivila, obrigado pelo amor, carinho, incentivo e força durante toda minha vida.

Endereço palavras de carinho a Sandra Benedito Mazivila e ao Emídio Eugénio Marrime, por me representar condignamente junto das instituições públicas e privadas na República de Moçambique, neste período da minha formação na República Federativa do Brasil.

Ao Carlos Januário Bié e Vitória Bulandane Paruque, pelo apoio, incentivo e acima de tudo ajuda diferenciada demonstrada durante a tramitação do pedido de autorização pelo Ministro da Educação para minha vinda ao Brasil.

A Universidade Federal de Uberlândia (UFU) por fornecer toda estrutura física e tecnológica para a realização deste trabalho.

Ao Convênio do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil/Ministério de Ciência e Tecnologia de Moçambique (convênio CNPq/MCT-Mz), processo número: 190802/2012-0, pelo suporte financeiro.

A TRANSPETRO S/A (Uberlândia, Estado de Minas Gerais, Brasil), pelo fornecimento de amostras de diesel puro.

Ao grupo de pesquisa do Laboratório de Quimiometria do Triângulo (LQT): Lucas Caixeta Gontijo, Letícia Maria de Souza, Felipe Bachion de Santana, Hery Mitsutake, Eloiza Guimarães, Fábio dos Santos Fonseca, Carlos Alberto Tadeu Zanetti, Eduardo José Buiatte, Edvando Teles, Baltazar Vasco Siteo e ao Flaysner Magayver Portela do Laboratório de Biocombustíveis (LABIO), pelas discussões científicas ao nível da agremiação, viagens em eventos científicos e verdadeira amizade criada, do fundo do coração esses momentos estarão eternamente na minha memória.

A Doutora Hélen Cristine de Rezende, que enquanto colega integrante da linha de pesquisa em Espectroscopia Aplicada do Programa de Pós-Graduação em Química do Instituto de Química da UFU, pelo conhecimento transmitido, amizade.

Aos Professores do Instituto de Química da UFU: Prof. Dr. Reinaldo Ruggiero, Profa. Dra. Nívia Maria Melo Coelho, Prof. Dr. Eduardo Mathias Richter, Profa. Dra. Aline Carlos de Oliveira, Prof. Dr. Alam Gustavo Trovó e ao Prof. Dr. Newton Barbosa do Instituto de Física da UFU, por todo conhecimento transmitido nas disciplinas cursadas.

Um especial agradecimento é endereçado aos Professores que participaram nas minhas bancas de qualificação e defesa, pelas contribuições sábias e que as mesmas foram incorporadas na versão final: Profa. Dra. Nívia Maria Melo Coelho, Prof. Dr. Alam Gustavo Trovó e Dr. Luiz Carlos Moutinho Pataca, Pesquisador no Instituto de Geoinformação e Tecnologia de Minas Gerais.

Ao meu orientador da Graduação, Prof. Dr. André Eduardo Gulube, Docente do Departamento de Química da Faculdade de Ciências Naturais e Matemática da Universidade Pedagógica, Maputo-Moçambique, por tudo aquilo que representa como uma pessoa especial, sendo um exemplo a ser seguido por mim na minha vida e sempre disponível para me facultar as cartas de recomendação na qualidade de orientador na Graduação, com intuito de dar sequências nas etapas da minha formação no ensino superior.

Ao Prof. Dr. Elias Narciso Matos, Docente do Departamento de Química da Faculdade de Ciências Naturais e Matemática da Universidade Pedagógica, Maputo-Moçambique, pelo incentivo demonstrado, através da carta de recomendação facultada e que a mesma foi usada como requisito ao concurso da bolsa do mestrado.

Aos meus eternos amigos: Cremildo Vitorino Zucula, Henriques Reginaldo Maculuve, Edson Fernandes Raso, Anastácio Armando Boane, pelo apoio abnegado empreendido na época do concurso da bolsa assim como nas vésperas da vinda ao Brasil e durante o curso.

Aos meus ex-alunos nos anos letivos de 2010, 2011, 2012 e 2013, da Escola Secundária Josina Machel de Belane, Vilankulo-Inhambane, Moçambique, pela oportunidade de aprendizado, fundamental para minha carreira de docência.

Ao Picardo Miquecene Nhabomba e Lalino Francisco Nhavene, companheiros de docência e de luta na Escola Secundária Josina Machel de Belane, e que foram responsáveis pela digitação e assinatura do parecer favorável, respectivamente, durante a tramitação do pedido de autorização para minha vinda ao Brasil, pelo Ministro da Educação do Governo de Moçambique.

Aos colegas da residência estudantil: Evangelino Aleluia Lopes Sá, cidadão de Guiné-Bissau e Ibrahim Muhammad Abubakar, cidadão da Nigéria, pela amizade e acima de tudo pela convivência mútua e entrosamento da diversidade cultural dos povos da nossa África.

A Secretária do Programa de Pós-Graduação em Química: Mayta Mamede Negreto Peixoto e aos Secretários do Instituto de Química: Ricardo Margonari da Silva, Gabriel Fonseca Rezende, Mário Guimarães Júnior e Henriques Gonçalves Borges, pelo pronto atendimento e profissionalismo durante a minha estadia na UFU.

Ao Coordenador da Divisão de Apoio à Pós-Graduação: João Martins Neto e ao Marcelo José Pereira, técnico da Divisão de Apoio à Pós-Graduação da Pró-Reitoria de Pesquisa e Pós-Graduação da UFU, por terem sido o ponto focal entre o estudante e o convênio CNPq/MCT-Mz.

Enfim, meu agradecimento a todos os amigos, professores e funcionários do Instituto de Química da UFU e a todas as pessoas que, de uma forma ou outra, contribuíram para realização deste trabalho e para esta etapa de minha vida, os meus sinceros agradecimentos.

**“O conhecimento é o processo de acumular dados,  
a sabedoria reside na sua simplificação”.**

*Martin H. Fischer*

## RESUMO

O biodiesel tem-se destacado por ser uma fonte alternativa de energia, visto que pode substituir ou ser adicionado ao óleo diesel. Este combustível renovável é consolidado no Brasil sendo de interesse do Governo de Moçambique, adquirir as tecnologias de produção e controle de qualidade em todo processo até a fase de distribuição. No Brasil, a fiscalização é feita pela Agência Nacional de Petróleo, Gás Natural e Biocombustível (ANP) que estabelece parâmetros de qualidade. Os parâmetros da ANP são baseados em normas das instituições signatárias do Sistema Internacional de Metrologia, por exemplo, as Normas da Sociedade Americana de Teste e Materiais. Dentre as várias técnicas analíticas para controle de qualidade de biodiesel se destaca a espectroscopia no infravermelho médio (MIR), pelo fato de ser rápida, análise não destrutiva, podendo ser aplicada *in situ* em todas as etapas de produção e fiscalização. No entanto, as potencialidades da espectroscopia MIR tornam-se ilimitadas quando combinadas aos métodos quimiométricos para extrair informações relevantes de dados complexos, tornando-os mais simples de serem aplicados em rotinas de laboratórios. Partindo de pressupostos acima apresentados, neste trabalho, foi aplicada a técnica de espectroscopia MIR aliada ao método de análise discriminante por quadrados mínimos parciais (PLS-DA), visando desenvolver metodologias capazes de classificar amostras de biodieseis conforme sua origem em: biodieseis metílicos e etílicos, puros B100 e em misturas com diesel na proporção de 5% (v/v), mistura conhecida por B5, óleo de soja, óleo de fritura usado e óleo de pinhão manso. Além disso, classificar amostras de B5 puras e adulteradas com gasolina, óleo lubrificante automotivo residual, óleo de soja e óleo de fritura usado. Os modelos foram construídos usando PLS1-DA e PLS2-DA, sendo o segundo para demonstrar a vantagem na classificação simultânea de diferentes tipos de biodieseis no mesmo modelo, o que diminui o tempo de análise, fator importante para tomada de decisões na linha de produção e laboratórios. Todos os modelos desenvolvidos apresentaram alta sensibilidade e especificidade com valores iguais a 1 (máxima classificação), o que significa 100% de classificação correta em todas as amostras do conjunto teste. Assim, as metodologias propostas podem ser aplicadas no controle de qualidade destes biodieseis e suas misturas com diesel.

**Palavras-chave:** Biocombustível. Espectroscopia MIR. PLS-DA.

## ABSTRACT

Biodiesel has been highlighted for being an alternative source of energy since it can replace or be added to diesel oil. This renewable fuel is already consolidated in Brazil being of interest to the Government of Mozambique to acquire production technologies and quality control throughout the process to distribution phase. In Brazil, supervision is done by the Brazilian National Agency of Petroleum, Natural Gas and Biofuels (ANP) that establishes quality parameters. The parameters of ANP are based on standards of the signatory institutions of the International Metrology System, for example, the Standards of American Standard for Testing Materials. Among the various analytical techniques for quality control of biodiesel stands out mid-infrared spectroscopy because it is fast, non-destructive analysis, can be applied in situ, at all stages of production and inspection. However, the potentiality of MIR spectroscopy become unlimited when combined with chemometric methods to extract relevant information of complex data, making them simpler to be applied in laboratories routine. Starting from the above assumptions, in this work, MIR spectroscopy technique allied to partial least square discriminant analysis (PLS-DA) was applied, aiming to develop methodologies able to classify samples of biodiesels according its origin the biodiesels methylic and ethylic, pure B100 and mixtures with diesel in proportion of 5% (v/v) mixture known as B5, of soybean oil, used frying oil, jatropha oil. In addition, classify samples of pure B5, and adulterated with gasoline, residual automotive lubricant oil, soybean oil and used frying oil. The models were built using PLS1-DA and PLS2-DA, the second being to demonstrate the advantage of simultaneous classification of different types of biodiesels in the same model, which reduces the analysis time, important factor to take decision in production line and laboratories. All models developed showed high sensitivity and specificity with values equal 1 (maximum classification), meaning 100% correct classification in all sample of the test set. Thus, the methodologies proposed can be applied in quality control of these biofuels and their blends with diesel.

**Keywords:** Biofuel. MIR Spectroscopy. PLS-DA.

## LISTA DE TABELAS

<b>Tabela 1</b>	Regiões do espectro no infravermelho	36
<b>Tabela 2</b>	Resumo sobre princípios, vantagens e limitações dos métodos quimiométricos de classificação citados	59
<b>Tabela 3</b>	Parâmetros de classificação obtidos por PLS1-DA para classificação B100	71
<b>Tabela 4</b>	Parâmetros de classificação obtidos por PLS2-DA para classificação B100	77
<b>Tabela 5</b>	Parâmetros de classificação obtidos por PLS1-DA	80
<b>Tabela 6</b>	Parâmetros de classificação obtidos por PLS1-DA para detectar amostras puras das adulteradas por gasolina	88
<b>Tabela 7</b>	Parâmetros de classificação obtidos por PLS1-DA para detectar amostras puras das adulteradas por OLAR	97
<b>Tabela 8</b>	Parâmetros de classificação obtidos por PLS1-DA para detectar amostras puras das adulteradas por óleo de soja	99
<b>Tabela 9</b>	Parâmetros de classificação obtidos por PLS1-DA para detectar amostras puras das adulteradas por OFU	104

## LISTA DE FIGURAS

Figura 1	Matérias-primas utilizadas na produção do biodiesel no território brasileiro	27
Figura 2	Reação geral da formação de ésteres metílicos ou etílicos de ácidos graxos (Biodiesel)	30
Figura 3	Formação de alcóxido	30
Figura 4	Mecanismo da reação de transesterificação por catálise básica	31
Figura 5	Monoacilglicerol e duas moléculas de ésteres etílicos ou metílicos	31
Figura 6	Glicerol ou glicerina livre e três moléculas de ésteres etílicos ou metílicos	32
Figura 7	Mecanismo de saponificação de triglicerídeos	32
Figura 8	Mecanismo de saponificação de ácidos graxos livres	33
Figura 9	Mecanismo de reação de esterificação (catálise ácida)	33
Figura 10	Espectros MIR de regiões de absorvâncias de grupos funcionais na mistura B5	38
Figura 11	Esquema representativo do funcionamento do ATR	40
Figura 12	Espectrômetro MIR equipado com acessório de ATR de ZnSe	40
Figura 13	Espectros MIR de B100 de BES, BMS, BEFU, BMFU, BEPM e BMPM.	41
Figura 14	Representação da construção da matriz $\mathbf{X}$ para modelagem multivariada	43
Figura 15	Espectros MIR com flutuação na linha de base (a) e corrigidos (b)	44
Figura 16	Dados hipotéticos antes do pré-processamento (a), centrados na média (b) e autoescalados (c)	45
Figura 17	Descrição gráfica do PCA	49
Figura 18	Representação matricial do PCA	50
Figura 19	Correlação de PCA com PLS	55
Figura 20	Exemplo da representação do resultado do modelo PLS-DA	58
Figura 21	Mistura de óleo e alcóxido na chapa de aquecimento com agitação magnética	62

Figura 22	Separação de fases, biodiesel (superior) e glicerina (inferior)	62
Figura 23	Purificação do biodiesel através da lavagem (ésteres metílicos ou etílicos de ácidos graxos)	63
Figura 24	Secagem do biodiesel à pressão reduzida	64
Figura 25	Sistema para a catálise ácida metílica e etílica	65
Figura 26	Mistura heterogênea entre ésteres e álcool, obtida após a reação em meio ácido	66
Figura 27	Espectros MIR de 132 amostras de B5: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM	68
Figura 28	Escores de PC1×PC2×PC3 de análise de PCA de diferentes tipos de B100	69
Figura 29	Gráfico de pesos de PC1, PC2 e PC3 versus número de onda (variáveis) para modelo PCA	70
Figura 30a	Estimativa de (■) BES	73
Figura 30b	Estimativa de (▲) BMPM	73
Figura 30c	Estimativa de (▼) BMFU	74
Figura 30d	Estimativa de (●) BMS	74
Figura 30e	Estimativa de (◆) BEFU	75
Figura 30f	Estimativa de (◀) BEPM	75
Figure 31	Gráfico de resíduos espectral versus leverage em nível de significância de 5 % para modelo PLS2-DA, onde (■ conjunto de treinamento) BEFU, (▲ conjunto de treinamento) BMS, (● conjunto de treinamento) BMFU, (◆ conjunto de treinamento) BES	77
Figura 32	Escores de análise do modelo PLS2-DA de tipos de biodieseis, onde: BEFU (■ conjunto de treinamento e □ conjunto de teste), BMS (▲ conjunto de treinamento e Δ conjunto de teste), BMFU (● conjunto de treinamento e ○ conjunto de teste) e BES (◆ conjunto de treinamento e ◇ conjunto de teste)	78
Figura 33	Espectros MIR de 132 amostras de B5: BES, BMPM, BMFU, BMS, BEFU e BEPM	79

Figura 34a	Estimativa de (■) BES	81
Figura 34b	Estimativa de (▲) BMPM	81
Figura 34c	Estimativa de (▼) BMFU	82
Figura 34d	Estimativa de (●) BMS	82
Figura 34e	Estimativa de (◆) BEFU	83
Figura 34f	Estimativa de (◀) BEPM	83
Figura 35	Variável Importante na Projeção (VIP) escores para modelo PLS1-DA	84
Figura 36	Espectros das amostras das misturas de B5 puras e adulteradas por gasolina: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM	87
Figura 37a	Estimativa de (◀) amostras não adulteradas de BES das (▲) misturas de B5 adulteradas por gasolina	89
Figura 37b	Estimativa de (◀) amostras não adulteradas de BEFU das (▲) misturas de B5 adulteradas por gasolina	89
Figura 37c	Estimativa de (◀) amostras não adulteradas de BEPM das (▲) misturas de B5 adulteradas por gasolina	90
Figura 37d	Estimativa de (■) amostras não adulteradas de BMPM das (▲) misturas de B5 adulteradas por gasolina	90
Figura 37e	Estimativa de (■) amostras não adulteradas de BMFU das (▲) misturas de B5 adulteradas por gasolina	91
Figura 37f	Estimativa de (■) amostras não adulteradas de BMS das (▲) misturas de B5 adulteradas por gasolina	91
Figura 38	Espectros de amostras das misturas de B5 puras e adulteradas por OLAR: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM	92
Figura 39a	Estimativa de (◀) amostras não adulteradas de BES das (▼) misturas de B5 adulteradas por OLAR	93
Figura 39b	Estimativa de (◀) amostras não adulteradas de BEFU das (▼) misturas de B5 adulteradas por OLAR	94
Figura 39c	Estimativa de (◀) amostras não adulteradas de BEPM das (▼) misturas de B5 adulteradas por OLAR	94

Figura 39d	Estimativa de (■) amostras não adulteradas de BMPM das (▼) misturas de B5 adulteradas por OLAR	95
Figura 39e	Estimativa de (■) amostras não adulteradas de BMFU das (▼) misturas de B5 adulteradas por OLAR	95
Figura 39f	Estimativa de (■) amostras não adulteradas de BMS das (▼) misturas de B5 adulteradas por OLAR	96
Figura 40	Espectros de amostras de misturas de B5 puras e adulteradas por óleo de Soja: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM	98
Figura 41a	Estimativa de (◄) amostras não adulteradas de BES das (●) misturas de B5 adulteradas por óleo de soja	100
Figura 41b	Estimativa de (◄) amostras não adulteradas de BEFU das (●) misturas de B5 adulteradas por óleo de soja	100
Figura 41c	Estimativa de (◄) amostras não adulteradas de BEPM das (●) misturas de B5 adulteradas por óleo de soja	101
Figura 41d	Estimativa de (■) amostras não adulteradas de BMPM das (●) misturas de B5 adulteradas por óleo de soja	101
Figura 41e	Estimativa de (■) amostras não adulteradas de BMFU das (●) misturas de B5 adulteradas por óleo de soja	102
Figura 41f	Estimativa de (■) amostras não adulteradas de BMS das (●) misturas de B5 adulteradas por óleo de soja	102
Figura 42	Espectros de amostras das misturas de B5 puras e adulteradas por OFU: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM	103
Figura 43a	Estimativa de (◄) amostras não adulteradas de BES das (◆) misturas de B5 adulteradas por OFU	105
Figura 43b	Estimativa de (◄) amostras não adulteradas de BEFU das (◆) misturas de B5 adulteradas por OFU	105
Figura 43c	Estimativa de (◄) amostras não adulteradas de BEPM das (◆) misturas de B5 adulteradas por OFU	106
Figura 43d	Estimativa de (■) amostras não adulteradas de BMPM das (◆) misturas de B5 adulteradas por OFU	106
Figura 43e	Estimativa de (■) amostras não adulteradas de BMFU das (◆) misturas de B5 adulteradas por OFU	107
Figura 43f	Estimativa de (■) amostras não adulteradas de BMS das (◆) misturas de	107

B5 adulteradas por OFU

## LISTA DE ABREVIATURAS E SIGLAS

<b>ATR</b>	Reflectância Total Atenuada (do inglês, Attenuated Total Reflectance)
<b>B100</b>	Biodiesel não acrescido de óleo diesel de petróleo
<b>B5</b>	Mistura em proporção volumétrica de 5% de biodiesel em 95% de diesel de petróleo
<b>BEFU</b>	Biodiesel Etílico de Óleo de Fritura Usado
<b>BEPM</b>	Biodiesel Etílico de Pinhão Manso
<b>BES</b>	Biodiesel Etílico de Soja
<b>BMFU</b>	Biodiesel Metílico de Óleo de Fritura Usado
<b>BMPM</b>	Biodiesel Metílico de Pinhão Manso
<b>BMS</b>	Biodiesel Metílico de Soja
<b>HCA</b>	Análise de Agrupamentos Hierárquicos (do inglês, Hierarchical Cluster Analysis)
<b>HATR</b>	Reflectância Total Atenuada Horizontal (do inglês, Horizontal Attenuated Total Reflectance)
<b>KNN</b>	K-Vizinhos mais Próximos (do inglês, K- Nearest Neighbor)
<b>MIR</b>	Infravermelho Médio (do inglês, Mid Infrared)
<b>OFU</b>	Óleo de Fritura Usado
<b>OLAR</b>	Óleo Lubrificante Automotivo Residual
<b>PCA</b>	Análise de Componentes Principais (do inglês, Principal Component Analysis)
<b>PC</b>	Componente Principal (do inglês, Principal Component)
<b>PLS</b>	Quadrados Mínimos Parciais (do inglês, Partial Least Squares)
<b>PLS-DA</b>	Análise Discriminante por Quadrados Mínimos Parciais (do inglês, Partial Least Squares-Discriminant Analysis)
<b>RMSEC</b>	Erro Quadrático Médio de Calibração (do inglês, Root Mean Squared Error of Calibration)
<b>RMSECV</b>	Erro Quadrático Médio de Validação Cruzada (do inglês, Root Mean Squared Error of Cross-Validation)
<b>RMSEP</b>	Erro Quadrático Médio de Previsão (do inglês, Root Mean Squared Error of Prediction)
<b>SIMCA</b>	Modelagem Independente e Flexível por Analogia de Classe (do inglês, Soft Independent Modeling of Class Analogy)
<b>VIP</b>	Variável Importante na Projeção

**VL**      Variável Latente

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	21
<b>2 BIODIESEL</b> .....	27
<b>2.1 Especificação do biodiesel</b> .....	29
<b>2.2 Produção de biodiesel</b> .....	29
<b>2.2.1 Catálise básica</b> .....	30
<b>2.3 Adulteração</b> .....	33
<b>3 ESPECTROSCOPIA NO INFRAVERMELHO</b> .....	35
<b>3.1 Fundamentos da espectroscopia</b> .....	36
<b>3.2 Reflexão total atenuada</b> .....	38
<b>4 TRATAMENTO DE DADOS</b> .....	43
<b>4.1 Quimiometria</b> .....	46
<b>4.2 Métodos quimiométricos de classificação</b> .....	48
<b>4.2.1 Métodos de classificação não supervisionada ou conhecimento de padrões</b> .....	48
<b>4.2.1.1 Análise de componentes principais (PCA)</b> .....	48
<b>4.2.1.2 Análise de agrupamentos hierárquicos (HCA)</b> .....	51
<b>4.2.2 Métodos de classificação supervisionada ou reconhecimento de padrões</b> .....	52
<b>4.2.2.1 K-Vizinhos mais Próximos (KNN)</b> .....	52
<b>4.2.2.2 Modelagem independente e flexível por analogia de classe (SIMCA)</b> .....	53
<b>4.2.2.3 Análise discriminante por quadrados mínimos parciais (PLS-DA)</b> .....	53
<b>4.2.2.3.1 Teorema de Bayes</b> .....	56
<b>4.2.2.3.2 Representação do resultado de PLS-DA</b> .....	58
<b>5 CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL</b> .....	60
<b>5.1 Objetivos Específicos</b> .....	60

<b>5.2 Metodologia</b> .....	60
<b>5.2.1 Procedimentos da produção do biodiesel</b> .....	60
5.2.1.1 Procedimento experimental da produção do biodiesel por catálise básica .....	61
5.2.1.2 Procedimento experimental da produção do biodiesel por catálise dupla .....	65
<b>5.2.2 Procedimentos quimiométricos</b> .....	66
<b>5.3 Modelos PLS-DA para classificar diferentes tipos de B100: BES, BMPM, BMFU, BMS, BEFU e BEPM</b> .....	67
<b>5.3.1 Análise exploratória por Análise de Componentes Principais (PCA)</b> .....	69
<b>5.3.2 Classificação de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA)</b> .....	71
<b>5.3.3 Classificação de Análise Discriminante por Quadrados Mínimos Parciais (PLS2-DA)</b> .....	76
<b>5.4 Modelos PLS1-DA para classificar diferentes tipos de B5: BES, BMPM, BMFU, BMS, BEFU e BEPM</b> .....	78
<b>6 MODELOS DE ADULTERAÇÃO EM MISTURAS B5</b> .....	85
<b>6.1 Objetivos Específicos</b> .....	85
<b>6.2 Metodologia</b> .....	85
<b>6.3 Modelos PLS1-DA para detectar adulteração por gasolina em misturas B5</b> .....	86
<b>6.4 Modelos PLS1-DA para detectar adulteração por OLAR em misturas B5</b> .....	92
<b>6.5 Modelos PLS1-DA para detectar adulteração por óleo de soja em misturas B5</b> .....	97
<b>6.6 Modelos PLS1-DA para detectar adulteração por óleo de fritura usado em misturas B5</b> .....	103
<b>7 CONCLUSÕES</b> .....	108
<b>8 REFERÊNCIAS</b> .....	109
<b>ANEXOS</b> .....	118

## 1 INTRODUÇÃO

O aumento da demanda mundial por combustíveis, aliada a vontade política para mitigar a dependência por combustíveis fósseis, incentiva a pesquisa de produção e controle de qualidade de combustíveis renováveis. Na perspectiva da busca de alternativas ao óleo diesel, se destaca o biodiesel.

No entanto, a produção do biodiesel não deve criar concorrência com o preço dos alimentos, em detrimento do uso de óleos usados na alimentação humana para produção do biodiesel.

Nesta perspectiva, a Comissão Européia, aprovou uma diretiva segundo a qual, a partir de 2020 os biocombustíveis produzidos a partir de óleos comestíveis, isto é, utilizados na alimentação humana, como óleo de soja (*Glycine max*), não deverão ser subsidiados de modo a incentivar o uso de matérias-primas que não levam à demanda adicional por terra e que não entram na alimentação humana (EUROPEAN COMMISSION, 2012). Perante este cenário destaca-se, por exemplo, as matérias-primas provenientes de óleo de fritura usado e óleo de pinhão manso (*Jatropha curcas*).

Portanto, o uso de óleo de fritura usado é um potencial promissor para produção de biodiesel (BAPTISTA et al., 2008; CVENGROŠ; CVENGROŠOVÁ, 2004), porque é um produto de descarte após a fritura e evita a competição com óleo vegetal utilizado na alimentação humana (ARAUJO; HAMACHER; SCAVARDA, 2010).

Por sua vez a *Jatropha curcas* é uma fonte promissora pelo fato do seu óleo não ser usado na alimentação humana devido à presença de substâncias tóxicas (MENDONÇA; LAVIOLA, 2009).

O Governo de Moçambique identificou matérias-primas para produção de biodiesel, onde a *Jatropha curcas* é tida como fonte promissora, apesar das tecnologias de produção desta oleaginosa se encontrar na fase inicial (ZÍLIO et al., 2008). Moçambique tem afirmado projetos com Brasil visando à obtenção de tecnologias de produção de biocombustíveis, uma vez que o Governo Brasileiro consolidou a produção de biodiesel, sendo interesse para o Governo Moçambicano. Nesta parceria, entre Moçambique e Brasil, se destaca a entrada em funcionamento da usina de etanol em Moçambique da Empresa Petróleo Brasileiro S. A (PETROBRAS) em parceria com a Empresa Petróleos de Moçambique, S. A. R. L. (PETROMOC). Portanto, na temática de biocombustíveis, Moçambique é o primeiro País Africano a desenvolver um plano nacional para biocombustíveis, plano esse lançado em Dezembro de 2007 (SCHUT et al., 2014).

O biodiesel pode ser usado na sua forma pura denominada por B100 ou misturado com óleo diesel denominado por BX, onde X corresponde à percentagem do biodiesel adicionado no óleo diesel. Em misturas binárias alguns países usam quantidades que variam entre 2% e 35% (v/v), por exemplo, 5-30% na França, 20% nos Estados Unidos da América e Canadá e 30% na República Checa e Eslováquia (ZHANG, 2012).

No entanto, a mistura de biodiesel/diesel comercializada pode ser adulterada com óleos tais como: soja, óleo de fritura usado (OFU), óleo lubrificante automotivo residual (OLAR), entre outros. Tal prática irregular visa à obtenção de lucro motivada pelo elevado preço da mistura do biodiesel/diesel em relação ao custo de adulterantes (BRANDAO; BRAGA; SUAREZ, 2012). Por outro lado, pode ocorrer a contaminação da mistura biodiesel/diesel por gasolina durante o transporte usando gasoduto, assim como pode ocorrer à contaminação por outros produtos derivados de petróleo (gasóleo, querosene) durante a fase de distribuição do combustível usando tanques (ALVES; HENRIQUES; POPPI, 2014).

Uma vez que o biodiesel pode ser produzido de várias fontes de matérias-primas, as quais apresentam características diferentes devido à composição química dos ácidos graxos de cada tipo de óleo usado na produção do biodiesel (BAPTISTA et al., 2008), o produto final apresenta propriedades físico-químicas diferentes.

Nesta perspectiva é necessário o desenvolvimento de metodologias que consigam identificar a fonte de matéria-prima utilizada na produção de biodiesel, de modo a satisfazer as exigências políticas, segundo as quais para subsidiar os biocombustíveis só os produzidos a partir de óleos que não entram na alimentação humana (EUROPEAN COMMISSION, 2012).

No desenvolvimento de metodologias é necessário identificar os tipos de matérias-primas usadas na produção de biodiesel B100, assim como o tipo de biodiesel adicionado na mistura biodiesel/diesel B5. Além disso, a solubilidade da mistura de B5 com OLAR, óleo de soja, OFU, contribui para a adulteração da mistura comercializada, que tem causado grandes prejuízos e afetado o mercado de combustíveis, fazendo-se necessário o desenvolvimento de metodologias acessíveis para identificação de adulterações em misturas B5, a fim de sujeitar os infratores às penalidades estabelecidas por lei.

A cromatografia é a técnica analítica mais utilizada para o controle de qualidade da mistura de biodiesel/diesel. SCHALE; LE; PIERCE (2012) classificaram a mistura de biodiesel/diesel em relação ao tipo de óleo usado na produção de biodiesel, com 100% de classificação correta empregando dados de cromatografia gasosa-espectrometria de massa aliados a K-Vizinhos Próximos (KNN).

BRANDAO et al. (2012) identificaram tipos de óleos vegetais presentes na adulteração da mistura de biodiesel/diesel (B5), onde classificaram corretamente 17 amostras no total de 18 amostras no conjunto teste, usando dados de cromatografia líquida de alta eficiência aliados a K-Vizinhos mais Próximos.

No entanto, no que diz respeito ao sistema de detecção, a cromatografia é uma técnica bastante sensível, porém lenta em relação às técnicas de espectroscopia vibracional (BALABIN; SAFIEVA, 2011). Os métodos de espectroscopia no infravermelho apresentam vantagens, como por exemplo, análise rápida uma vez que o espectro é obtido em menos de um minuto, não destrutiva dado que preserva a integridade física da amostra, não demanda reagentes, exige um mínimo ou não exige tratamento de amostra (DA SILVA et al., 2014), além de possuir equipamentos portáteis permitindo a análise *in situ* (MAZIVILA et al., 2015a). Neste sentido, a técnica de espectroscopia no infravermelho vem ganhando destaque.

PONTES et al. (2011) usaram análise discriminante por quadrados mínimos parciais (PLS-DA) e outras ferramentas quimiométricas de classificação para detectar adulterações por óleo vegetal cru em misturas de biodiesel/diesel usando dados de espectroscopia no infravermelho próximo (NIR), em que obtiveram 100% de classificação correta nas amostras do conjunto teste. MAZIVILA et al. (2015b) desenvolveram uma metodologia associando espectroscopia no infravermelho médio (MIR) e PLS-DA, com a finalidade de detectar adulterações por gasolina, OLAR, óleo de soja e OFU em amostras da mistura biodiesel/diesel B5, na qual obtiveram 100% de classificação correta nas amostras do conjunto teste.

Devido às semelhanças na natureza vibracional presentes tanto nos adulterantes quanto nos componentes da mistura biodiesel/diesel, os espectros da espectroscopia MIR apresentam elevado grau de sobreposição ao longo de toda região espectral e poucas regiões com variação nos valores de absorvância (ANDRADE et al., 1997).

Assim, a maioria dos modos de vibração são observados na região do MIR, com ênfase na região das impressões digitais, as quais que apresentam espectros complexos que incluem muitas vibrações acopladas (GAUGLITZ; VO-DINH, 2003), sendo difícil a distinção das amostras. Conseqüentemente, torna-se necessário o uso de métodos quimiométricos para a classificação dos perfis espectrais das diferentes amostras da mistura de biodiesel/diesel. Nesta perspectiva, para classificação é imprescindível à aplicação de métodos quimiométricos de classificação, como PLS-DA.

PLS-DA é uma variação do algoritmo da regressão por quadrados mínimos parciais (PLS) para análise discriminante (DA SILVA et al., 2014). Portanto, PLS-DA, é um método

de reconhecimento de padrões usado para identificar as semelhanças e diferenças em diferentes tipos de amostras, comparando-as entre si. Este método se fundamenta nas suposições segundo as quais, amostras do mesmo tipo são semelhantes, existem diferenças entre diferentes tipos de amostras, as semelhanças e diferenças são refletidas nas medidas utilizadas para caracterizar as amostras.

Este método após ter definido o perfil das amostras no modelo, durante a etapa de treinamento e posterior validação do modelo, usando amostras diferentes das usadas na etapa de treinamento. Portanto, as amostras do conjunto de treinamento, uma vez que definem o perfil das amostras no modelo, farão com que as amostras desconhecidas com perfil das amostras do conjunto de treinamento sejam reconhecidas ou identificadas pelo modelo (BRERETON, 2003a).

Portanto, esta Dissertação foi dividida em sete capítulos incluindo a introdução como primeiro capítulo, além de referências. O segundo capítulo, intitulado **Biodiesel**, aborda sua importância no cenário brasileiro e mundial, as principais características deste combustível renovável e sua produção através de reações de transesterificação de óleo de soja e óleo de fritura usado e óleo de pinhão manso. Para este último óleo a reação de transesterificação foi antecedida pela reação de esterificação, visando à redução de ácidos graxos livres. Ainda neste capítulo são apresentados os respectivos mecanismos de reação tanto de esterificação assim como de transesterificação.

O terceiro capítulo denominado de **Espectroscopia no Infravermelho** apresenta uma breve descrição dos fundamentos da técnica de absorção molecular da radiação na região do infravermelho com ênfase no MIR. Este capítulo trata ainda de algumas aplicações da técnica de espectroscopia no infravermelho no controle de qualidade de combustíveis.

O quarto capítulo intitulado **Análise Discriminante por Quadrados Mínimos Parciais** trata de forma concisa dos fundamentos quimiométricos para construção de modelos de classificação multivariada, com destaque para PLS-DA, que pode ser PLS1-DA ou PLS2-DA e os cálculos de sensibilidade e especificidade, que são os parâmetros usados para avaliar a eficiência dos modelos de classificação.

A primeira aplicação dos métodos de classificação multivariada aliados aos dados de espectroscopia MIR é discutida no quinto capítulo, **Classificação de Biodieseis e suas misturas com diesel**, através da construção de modelos PLS1-DA, para identificar o tipo de biodiesel de acordo com o tipo de óleo e álcool usado na produção de B100 assim como o tipo de biodiesel usado nas misturas de B5. Neste capítulo foi desenvolvido um modelo PLS2-DA para classificar simultaneamente quatro tipos de B100, numa perspectiva virada ao

desenvolvimento de métodos supervisionados de classificação, para analisar mais de uma propriedade de interesse. Tanto os modelos PLS1-DA quanto PLS2-DA, são avaliados os valores de erros obtidos e realizada a validação dos modelos construídos através da sensibilidade e especificidade nas amostras do conjunto de teste, para cada classe de interesse.

Os **Modelos de adulterações em misturas B5** são apresentados no sexto capítulo, através da construção de modelos PLS1-DA para detectar adulterantes artificiais tais como, gasolina, OLAR, óleo de soja e OFU, em cada tipo de mistura B5, com intuito de classificar amostras puras e adulteradas, e apresentar os valores de sensibilidade e especificidade nas amostras do conjunto de teste, para cada classe de interesse.

Esta Dissertação termina com o sétimo capítulo intitulado **Conclusões**, mostrando as potencialidades do PLS-DA aliada a espectroscopia MIR, na classificação dos tipos de biodieseis de acordo com o tipo de óleo e álcool usado na produção de B100, assim como o tipo de biodiesel usado nas misturas de B5, além de detectar os tipos de adulterantes nas misturas B5. Por fim, segue uma lista de **Referências** em que são apresentados os trabalhos que serviram como suporte científico na elaboração desta Dissertação.

Neste contexto, esta dissertação teve como objetivos os seguintes pontos:

a) Objetivo geral:

- ✓ Desenvolver metodologias analíticas para classificar diferentes tipos de biodieseis puros (B100) e misturas biodiesel/diesel (B5) conforme a matriz (tipo de óleo usado) e rota (tipo do álcool usado) e adulterações em misturas B5.

b) Objetivos específicos:

- ✓ produzir 22 lotes diferentes de B100 de biodiesel etílico de soja (BES), biodiesel metílico de pinhão manso (BMPM), biodiesel metílico de óleo de fritura usado (BMFU), biodiesel metílico de soja (BMS), biodiesel etílico de óleo de fritura usado (BEFU) e biodiesel etílico de pinhão manso (BEPM);
- ✓ construir modelos PLS1-DA e PLS2-DA para classificar os tipos de B100, em relação ao tipo de óleo e rota, usado na produção através de dados de espectroscopia MIR;
- ✓ obter 22 lotes diferentes de B5 partindo de amostras de B100 produzidas de BES, BMPM, BMFU, BMS, BEFU e BEPM;

- ✓ construir modelos PLS-DA para classificar os tipos de biodiesel (B5) em relação ao tipo de óleo e álcool usado na produção de B100 através de dados da espectroscopia MIR;
- ✓ adulterar artificialmente 15 lotes diferentes de amostras B5 na faixa de concentração de 1-20% para cada tipo (BES, BMPM, BMFU, BMS, BEFU e BEPM) usando gasolina, OLAR, soja e OFU;
- ✓ construir modelos PLS-DA para detectar cada tipo de adulterante (gasolina, OLAR, soja, OFU) em cada um dos tipos de B5 (BES, BMPM, BMFU, BMS, BEFU e BEPM) usando dados de espectroscopia MIR.
- ✓ validar os modelos PLS-DA desenvolvidos, através dos valores da sensibilidade e especificidade, para cada parâmetro nas amostras de conjunto teste na classe de interesse.

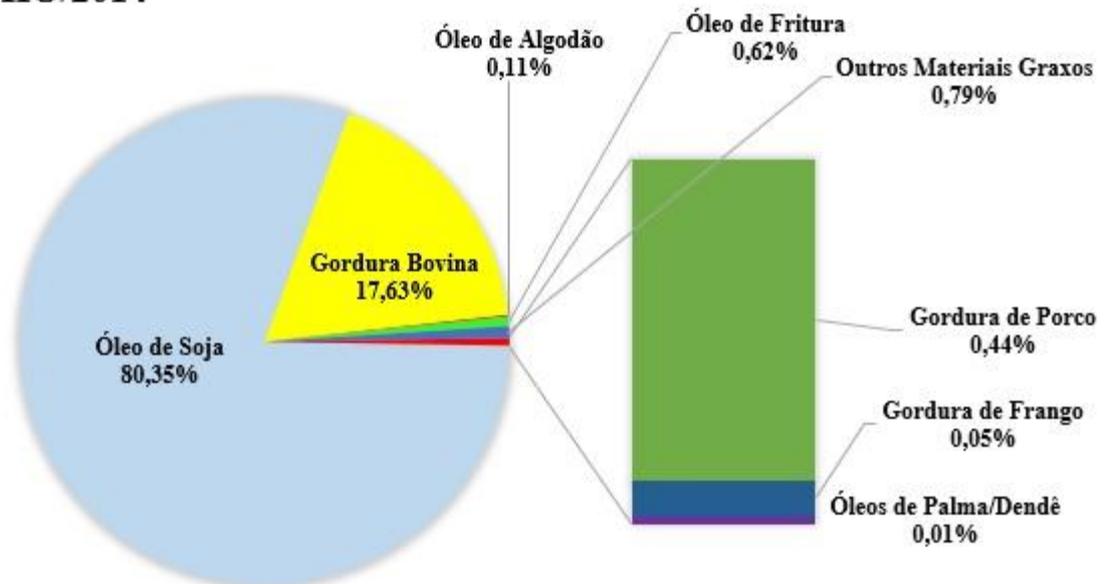
## 2 BIODIESEL

O biodiesel pode ser produzido através da reação química de transesterificação, na qual óleos de diferentes oleaginosas, OFU e gorduras de animais, reagem com álcool, normalmente, metanol ou etanol, na presença de um catalisador (LOPES et al., 2014). Portanto, quimicamente o biodiesel é composto por ésteres etílicos ou metílicos de ácidos graxos (CHEN et al., 2013).

A disponibilidade de diferentes matérias-primas na produção de biodiesel requer muita atenção, pois, a produção de óleos vegetais no Brasil, é baseada em sua grande maioria na soja, com cerca de 80% do biodiesel produzido sendo proveniente de óleo de soja (DE SOUSA et al., 2014). Portanto, a soja é a oleaginosa mais utilizada dentre as fontes de matérias-primas utilizadas na produção do biodiesel com 80,35%, devido às tecnologias consolidadas de produção como modificação genética (Figura 1) (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, 2014).

Figura 1 - Matérias-primas utilizadas na produção do biodiesel no território brasileiro.

**JUNHO/2014**



Fonte: Boletim Mensal de Biodiesel da ANP (Julho, 2014, p.10).

Embora a soja seja a principal matéria-prima utilizada na produção de biodiesel, esta apresenta desvantagem devido à questão de competição alimentar. Assim, depender apenas de uma cultura como principal fornecedora de matéria-prima para produção de biodiesel, pode limitar a produção do combustível para algumas regiões de acordo com as condições climáticas e qualidade de solo (MUELLER et al., 2013).

Com intuito de uma maior diversidade de matérias-primas na produção de biodiesel, de modo a minimizar a concorrência com os produtos alimentícios, há resíduos que podem ser usados na produção do biodiesel como OFU, visando o aproveitamento deste resíduo, uma vez que a utilização de OFU reduz o problema de contaminação do meio ambiente, pois a reutilização destes resíduos graxos pode contribuir na diminuição dos encargos ao governo na remoção destes resíduos nos esgotos (ENCINAR; GONZÁLEZ; RODRÍGUEZ-REINARES, 2007). Assim, o óleo residual proveniente da fritura dos alimentos é potencial promissor na produção de biodiesel (BAPTISTA et al., 2008; CVENGROŠ; CVENGROŠOVÁ, 2004), pelo fato de ser um produto de descarte após a fritura, além de evitar a concorrência com óleo vegetal usado na alimentação humana.

Além do OFU, pode se utilizar o óleo de pinhão manso, uma vez que esta oleaginosa vem se mostrando como uma fonte de óleo, para produção do biodiesel, pelo fato de pinhão manso se desenvolver em terrenos degradados, do semiárido ou em áreas para reflorestamento. O pinhão manso é adequado para a produção de biodiesel, pelo fato de não ser óleo comestível, razão pela qual é uma fonte viável para a produção de biodiesel, sem comprometer a indústria alimentar (DIVAKARA et al., 2010; OM TAPANES et al., 2008). O óleo de pinhão manso não é comestível devido à presença de componentes tóxicos, tais como: ésteres de forbol e curcina (LEUNG; WU; LEUNG, 2010; SHAH; GUPTA, 2007). O óleo de pinhão manso não é usado na alimentação humana uma vez que contêm substâncias tóxicas, em detrimento disso, óleo pode ser destinado à produção do biodiesel. Além disso, há estudos em curso tendentes a destoxificação da torta do pinhão manso. Portanto, a destoxificação da torta irá rentabilizar o cultivo desta oleaginosa, dado que a sua torta será aproveitada para a ração animal, o mesmo que acontece com o farelo proteico da soja (MENDONÇA; LAVIOLA, 2009).

A produção do biodiesel através de OFU e óleo de pinhão manso visa satisfazer as medidas políticas (EUROPEAN COMMISSION, 2012). Nesta perspectiva, nos países como Indonésia, Índia e China, utilizam óleo de pinhão manso como matéria-prima, na produção do biodiesel, pois os óleos comestíveis não estão em oferta excessiva, razão pela qual, não são usados na produção de biodiesel, uma vez que tais óleos são mais necessários como suplemento alimentar (SHAH; GUPTA, 2007; YINGYING et al., 2012).

## 2.1 Especificação do biodiesel

O biodiesel produzido deve atender as especificações estabelecidas, para tal a determinação das características do biodiesel deverá ser feita mediante o emprego das normas das instituições signatárias do Sistema Internacional de Metrologia, como as Normas Internacionais da Sociedade Americana de Teste e Materiais (ASTM), da Organização Internacional para Normalização (ISO), do Comitê Europeu de Normalização (CEN), assim como das instituições que seguem as normas de metrologia internacional como é o caso da Associação Brasileira de Normas Técnicas (ABNT), Regulamento Técnico da ANP.

Assim as adições do biodiesel ao óleo diesel são feitas mediante a legislação em vigor em cada País. Por exemplo, no Brasil de 2012 até meados de 2014, a adição estabelecida de biodiesel era de 5% (v/v) (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, 2012).

## 2.2 Produção de biodiesel

A produção do biodiesel (ésteres metílicos ou etílicos de ácidos graxos) pode ser feita através de várias rotas (KUMAR TIWARI; KUMAR; RAHEMAN, 2007; LAM; LEE; MOHAMED, 2010; LEUNG et al., 2010). Nesta Dissertação foram usadas as rotas metílica e etílica, devido à facilidade e a versatilidade do emprego da catálise alcalina. Por outro, estas rotas não permitem o uso de óleos com elevado teor de água ou ácidos graxos livres, sendo empregados, óleos vegetais refinados ou degomados (LEUNG; WU; LEUNG, 2010).

Os óleos vegetais brutos apresentam diversas impurezas, principalmente os ácidos graxos livres que são removidos no processo de refino, através das etapas de degomagem e neutralização (JIN et al., 2008; PEREZ et al., 2010). Os ácidos graxos livres podem desfavorecer a catálise do biodiesel devido ao consumo de reagentes ou catalisadores através de reações paralelas. Uma alternativa seria uma etapa prévia de refinamento. Contudo, esta etapa é inviável, uma vez que eleva muito o custo deste combustível (JIN et al., 2008).

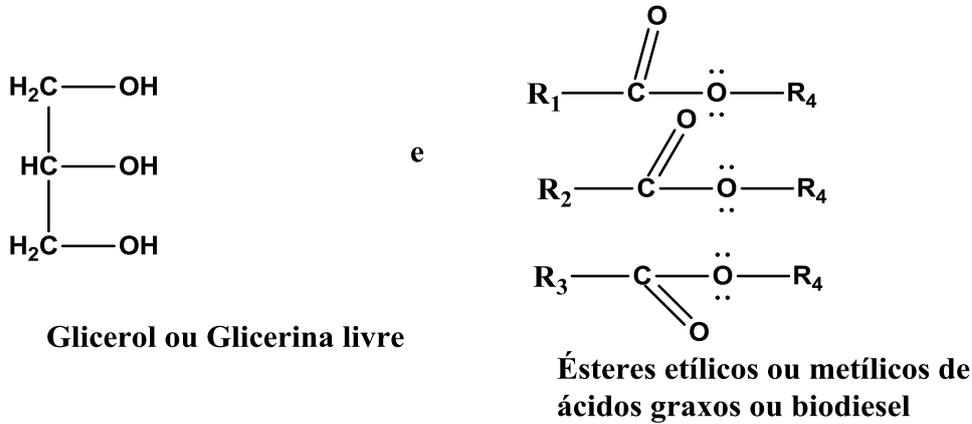
Assim, os óleos vegetais e OFU não são utilizados como combustíveis, pois, apresentam viscosidade, o que pode provocar danos aos motores automotivos, sendo necessário a redução da viscosidade dos óleos, através da reação de transesterificação.





De forma análoga, a terceira molécula do alcóxido atacará a molécula de monoacilglicerol formando glicerol ou glicerina livre e três moléculas de ésteres (Figura 6).

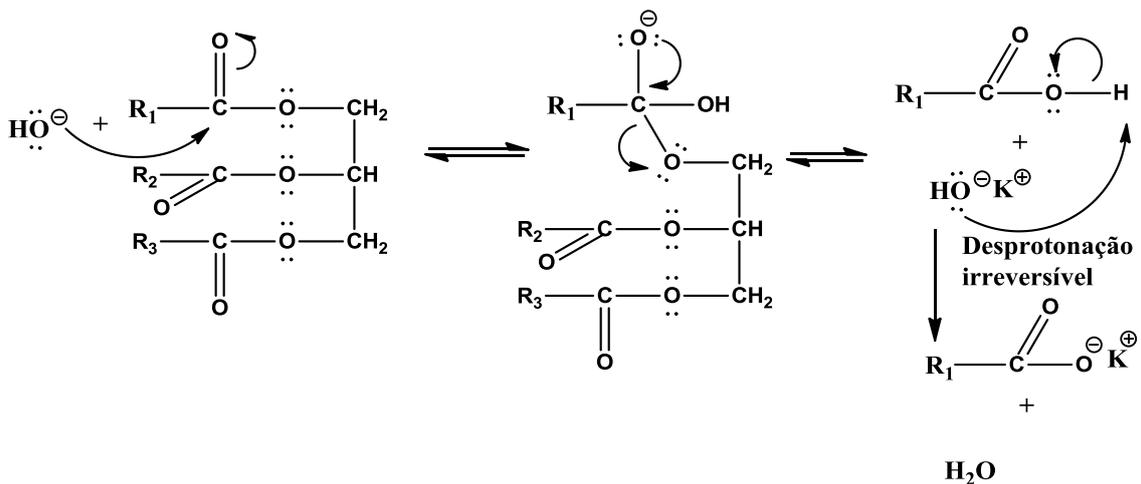
Figura 6 - Glicerol ou glicerina livre e três moléculas de ésteres etílicos ou metílicos.



Fonte: o Autor.

Ao utilizar a catálise básica na presença de ácidos graxos livres nos óleos vegetais, ocasiona a formação de sabões de ácidos graxos, desfavorecendo a reação de transesterificação devido ao consumo do catalisador na reação de saponificação (Figura 7) (LAM et al., 2010; LEUNG et al., 2010).

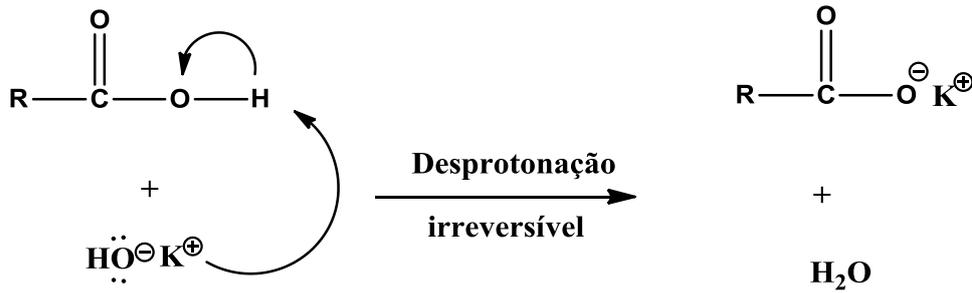
Figura 7 - Mecanismo de saponificação de triglicerídeos.



Fonte: o Autor.

A presença dos ácidos graxos livres e de outras impurezas presentes no óleo vegetal afeta negativamente o rendimento (JIN et al., 2008), além de necessitar de um período longo na etapa de decantação, de modo a ocorrer uma separação satisfatória durante a lavagem dos ésteres devido à formação de sabões (Figura 8).

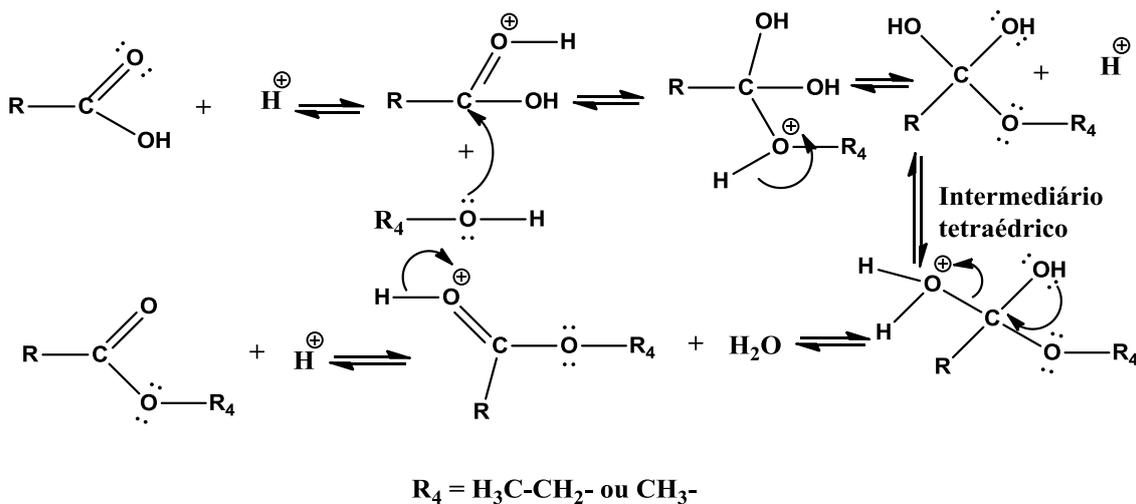
Figura 8 - Mecanismo de saponificação de ácidos graxos livres.



Fonte: o Autor.

Assim sendo, para óleo de pinhão manso é muito aconselhada o uso de catálise ácida, visando a conversão de ácidos graxos livres em ésteres metílicos ou etílicos (Figura 9). Após a catálise ácida, procede-se com a transesterificação da catálise básica (CLAYDEN et al., 2011).

Figura 9 - Mecanismo de reação de esterificação (catálise ácida).



Fonte: o Autor.

### 2.3 Adulteração

A possibilidade de adulteração de combustível aliada à carência de metodologias rápidas e eficientes tem sido investigada de modo a manter a qualidade do produto ao consumidor. Com o uso de combustível adulterado, os veículos movidos a diesel sofrem um desgaste percebido em longo prazo, diferentemente dos veículos movidos a gasolina e álcool que sofrem problemas instantâneos, provocando aquecimento e aceleração do motor, perda de potência, entupimento dos bicos injetores e do filtro do combustível, contaminação do óleo

lubrificante, aparecimento de borra no tanque e corrosão (ALBERICI et al., 2010; CORGOZINHO; PASA; BARBEIRA, 2008).

Deste modo, é fundamental o controle de qualidade de combustível desde o biodiesel puro B100 ou misturas de biodiesel/diesel B5 comercializadas, atestando a conformidade da amostra analisada face às especificações estabelecidas pela ANP (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, 2012). Portanto, visando o monitoramento de qualidade de biodiesel produzido neste trabalho, as análises foram feitas usando espectroscopia MIR.

Na espectroscopia MIR com acessório de refletância total atenuada (ATR), os espectros são obtidos sem preparo de amostras, sendo uma análise rápida uma vez que o espectro é obtido em menos de um minuto; usa menor quantidade da amostra (cerca de 0,5 mL); é um procedimento não destrutivo uma vez que preserva a integridade física da amostra, além de possuir equipamentos portáteis que viabilizam a análise *in situ*, isto é, no local da coleta de amostra.

### 3 ESPECTROSCOPIA NO INFRAVERMELHO

Espectroscopia no infravermelho é uma das técnicas clássicas para identificação de grupos funcionais, devido a sua capacidade em detectar o sinal característico do perfil dos grupos funcionais de acordo com a estrutura das moléculas. Portanto, a espectroscopia no infravermelho é tida como uma poderosa ferramenta na investigação de estruturas das moléculas.

Esta técnica está fundamentada no fenômeno de absorção da radiação eletromagnética, de comprimento de onda de 0,78 até 1000  $\mu\text{m}$ , por um sistema molecular promovendo o aumento de energia vibracional e ou rotacional das espécies químicas (BARTH, 2007).

Do ponto de vista da aplicação e instrumentação, as técnicas espectroscópicas que fazem uso da radiação na região do infravermelho são melhores compreendidas subdividindo-se o espectro em faixas denominadas de infravermelho próximo (NIR) de 0,78  $\mu\text{m}$  até 2,5  $\mu\text{m}$ , infravermelho médio (MIR) de 2,5  $\mu\text{m}$  até 50  $\mu\text{m}$  e infravermelho distante (FIR) de 50  $\mu\text{m}$  até 1000  $\mu\text{m}$ .

O espectro no infravermelho é desenhado em função do número de onda ( $\lambda$ ), o qual é proporcional à transição de energia e tem como unidade  $\text{cm}^{-1}$ . O principal motivo para os químicos preferirem o número de onda como unidade é devido ao fato de ser diretamente proporcional à energia (um número de onda maior corresponde a maior energia).

Tendo o valor de comprimento de onda ( $\lambda$ ) e utilizando a Eq. 1, é possível calcular o valor de número de onda ( $\text{cm}^{-1}$ ).

$$\tilde{\nu}_{(\text{cm}^{-1})} = \frac{1}{\lambda(\text{cm})} \quad (1)$$

Com o valor de comprimento de onda ( $\lambda$ ), é possível calcular a frequência (Eq. 2).

$$\nu_{(\text{Hz})} = \tilde{\nu}c = \frac{c(\frac{\text{cm}}{\text{s}})}{\lambda(\text{cm})} \quad (2)$$

A partir dos valores de comprimento de onda ( $\mu\text{m}$ ), é possível fazer a conversão para número de onda ( $\text{cm}^{-1}$ ) e vice-versa (Eq. 3 e 4).

$$\text{cm}^{-1} = \frac{1}{(\mu\text{m})} \times 10000 \quad (3)$$

$$\mu\text{m} = \frac{1}{(\text{cm}^{-1})} \times 10000 \quad (4)$$

Com o uso das equações acima referidas, podemos resumir as regiões do espectro infravermelho, com as respectivas unidades, segundo a Tabela 1 (BARTH, 2007).

**Tabela 1** - Regiões do espectro no infravermelho

Região	Comprimento de onda (nm)	Número de onda ( $\text{cm}^{-1}$ )	Frequência (Hz)
NIR	780 à 2500	12800 à 4000	$3,8 \times 10^{14}$ à $1,2 \times 10^{14}$
MIR	2500 à 50000	4000 à 200	$1,2 \times 10^{14}$ à $6,0 \times 10^{12}$
FIR	50000 à 1000000	200 à 10	$6,0 \times 10^{12}$ à $3,0 \times 10^{11}$

Fonte: o Autor.

As técnicas baseadas em espectroscopia no infravermelho possuem aplicações diversificadas: na indústria farmacêutica (KRÄMER; EBEL, 2000; ZHANG, C.; SU, 2014), no ramo petroquímico (BALABIN; SYUNYAEV, 2008; CHAKRABORTY et al., 2014), na indústria alimentícia (AOUIDI et al., 2012; BORRAS et al., 2014; COZZOLINO, 2014), dentre outras.

### 3.1 Fundamentos da espectroscopia

As interações da radiação com a matéria são o objeto de estudo da ciência da espectroscopia. Porém, os métodos espectroscópicos de análise são baseados na medida da quantidade de radiação absorvida ou emitida pelas moléculas ou pelas espécies atômicas de interesse. Portanto a absorção ou emissão ocorre geralmente, através das transições eletrônicas, que são situadas na região do ultravioleta ou visível, das transições vibracionais na região do infravermelho e das transições rotacionais na região de microondas, em alguns casos, também na região do infravermelho distante ou longínquo (SKOOG et al., 2008).

Uma molécula ser ativa ou poder absorver a radiação infravermelha é necessário que o momento dipolar varie. Para transferir energia, uma ligação deve apresentar um dipolo elétrico que mude na mesma frequência da radiação que está sendo introduzida. Portanto, o dipolo elétrico oscilante da ligação pode, então, acoplar-se com o campo eletromagnético da radiação incidente, que varia de forma senoidal. Assim, apenas as ligações que têm um momento de dipolo que varie como uma função do tempo são capazes de absorver a radiação infravermelha (PAVIA, 2012).

Espectros de emissão ou absorção são gerados quando a radiação interage com os átomos ou moléculas da matéria. Esta interação pode dar informações tais como a estrutura molecular e a interação de moléculas com seus vizinhos. Portanto a quantidade de informação

que se pode obter de um espectro depende essencialmente do tipo de fonte de radiação empregado frente às amostras, além da resolução temporal ou espectral dos instrumentos (DEMTRÖDER, 2003).

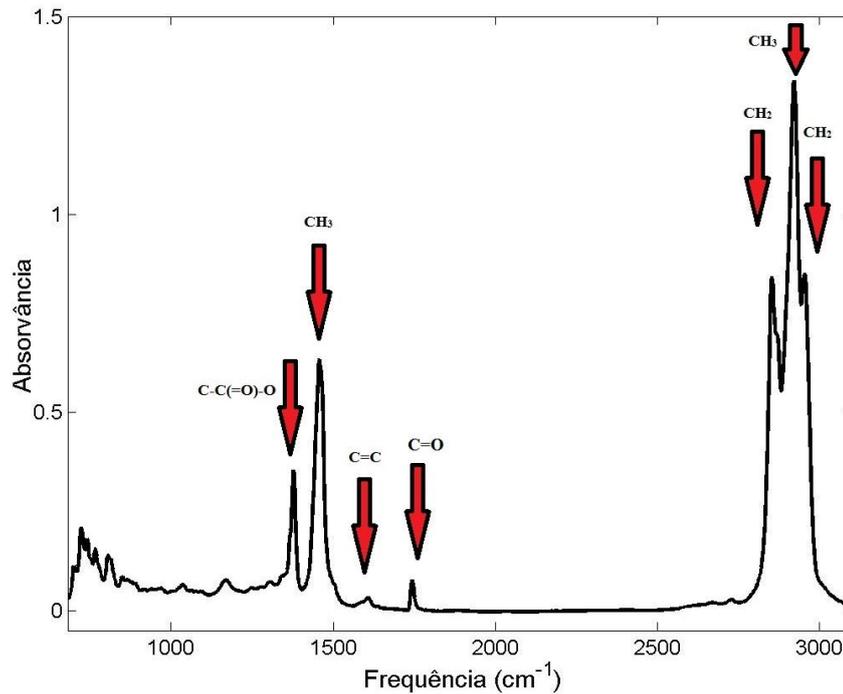
É importante destacar que, a energia de um fóton determina a frequência (e comprimento de onda) de uma onda eletromagnética, assim sendo, o espectro de absorção de qualquer sistema representa seu espectro energético. Diante disso, pode se afirmar que o comprimento de onda, frequência e energia são medidas equivalentes na espectroscopia, resultando em numerosas unidades utilizadas para caracterizar um mesmo parâmetro, isto é, a energia de transição. Portanto, o espectro se refere a um gráfico da radiação emitida em função da frequência ou do comprimento de onda.

Em um espectro de absorção da mistura B5 (Figura 10) observa-se duas bandas de absorção de estiramento das ligações C=O e C-O, características de ésteres metílicos ou etílicos. Na faixa de 1750 a 1730  $\text{cm}^{-1}$  ocorre a absorção do grupo carbonila (ALMEIDA et al., 2012). A região de 1500-900  $\text{cm}^{-1}$  é conhecida como regiões das impressões digitais, de espectros complexos originando a sobreposição espectral, isto é, as bandas particulares nesta região dificilmente são atribuídas às bandas das ligações químicas (GAUGLITZ ; VO-DINH, 2003).

A vibração de estiramento da ligação C-O dos ésteres consiste em dois modos vibracionais assimétricas acoplados: C-C(=O)-O e O-C-C, sendo o primeiro o mais importante. Portanto, estes modos vibracionais ocorrem na região de 1300 a 1100  $\text{cm}^{-1}$ . A banda característica de C-C(=O)-O mostra-se com maior intensidade em 1210 a 1163  $\text{cm}^{-1}$ , comumente mais larga e mais intensa do que a banda de estiramento C=O. A banda O-C-C de ésteres de alcoóis primários ocorre em 1164 a 1031  $\text{cm}^{-1}$  (Figura 16).

As ligações C-H, geram bandas de estiramento na região de 3000 a 2840  $\text{cm}^{-1}$ . A banda proeminente observada em 2920  $\text{cm}^{-1}$  está associada a vibração de  $\text{CH}_3$  e outra vibração é observada em 1450  $\text{cm}^{-1}$ . Por sua vez o grupo metileno ( $\text{CH}_2$ ) apresenta uma banda em 2850  $\text{cm}^{-1}$  associada ao estiramento simétrico e outra banda em 2950  $\text{cm}^{-1}$  associada ao estiramento assimétrico. Em 1600 a 1640  $\text{cm}^{-1}$ , observa-se uma banda ligada à ligação C=C (ALMEIDA et al., 2012).

Figura 10 - Espectro MIR de regiões de absorvâncias de grupos funcionais na mistura de B5.



Fonte: o Autor.

### 3.2 Reflexão total atenuada

Os principais métodos de reflexão no infravermelho são o método por reflexão difusa, o método por reflexão especular (ou externa), método por reflexão interna ou reflexão total atenuada.

A reflectância difusa é observada quando uma luz incide em uma matriz descontínua, penetra na amostra (amostras do tipo pó, sólido, placa ou filme) e reflete trazendo informações espectrais. O caminho percorrido pela luz no interior da matriz pode ser considerado aleatório devido a múltiplas reflexões, algumas das quais após percorrer o interior de algumas partículas que constituem a amostra. Desta forma a luz refletida pode ser atenuada por absorção.

É importante salientar que num experimento de reflectância difusa será também observado a reflexão especular, sendo mais importante àquela que ocorre na interface ar/superfície da matriz, mas também ocorrendo nas faces das partículas, mas cujo efeito pode ser minimizado pela redução do tamanho da partícula. A reflexão especular é de maior intensidade na região onde a amostra apresenta forte absorção e neste caso pode ocorrer severas distorções no espectro obtido.

Reflexão total atenuada foi usada na aquisição dos dados espectrais neste trabalho. Com uso da técnica de ATR, os espectros de absorção são obtidos de forma rápida para amostras sólidas ou líquidas com um mínimo de preparação prévia das amostras (DA SILVA et al., 2014), sendo aplicável tanto para análises qualitativas quanto quantitativas.

O princípio deste tipo de espectroscopia com acessório de ATR baseia-se no fato de que quando um feixe de radiação passa de um meio mais denso (cristal de ATR) para um meio menos denso (amostra), ocorre reflexão. A fração do feixe de luz incidente que é refletida aumenta conforme aumenta o ângulo de incidência, e quando excede um determinado ângulo crítico ( $45^\circ$ ) a reflexão é completa (Figura 11) (COZZOLINO; ROUMELIOTIS; EGLINTON, 2014b).

Assim, o mecanismo de absorção na espectroscopia do infravermelho equipado com ATR, se deve a dois fenômenos (COZZOLINO et al., 2014b):

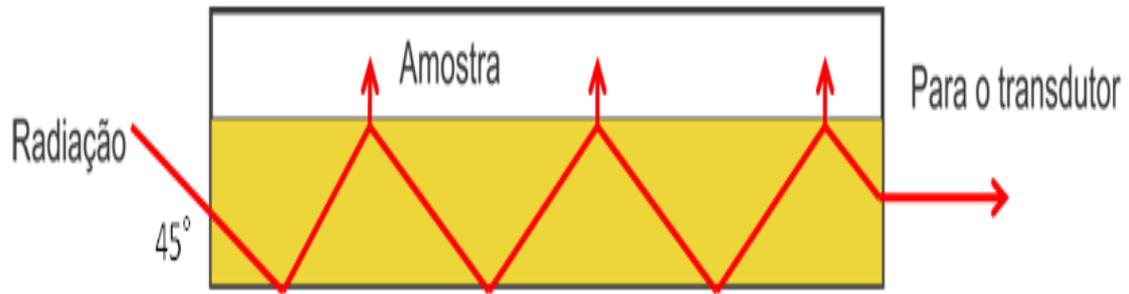
- ✓ rápida mudança no índice de refração da amostra através de um pico de absorção da radiação infravermelha e
- ✓ o coeficiente de absorção molar atribuído ao modo vibracional.

As principais vantagens da utilização da espectroscopia MIR equipada com acessório de ATR é a disponibilidade simples, fácil de usar, utilizada como ferramenta para triagem de alto rendimento em pesquisas (COZZOLINO et al., 2014a).

Essa radiação denominada onda evanescente pode ser parcialmente absorvida, colocando-se uma amostra em contato com meio mais denso. Se o meio menos denso absorve essa onda evanescente, ocorre atenuação do feixe nos comprimentos de onda das bandas de absorção. Esse meio mais denso é denominado prisma ou elemento de reflexão interna.

O cristal de seleneto de zinco (ZnSe) é um dos materiais utilizados para reflectância total atenuada, pois é insolúvel em água, resistente a choques térmicos e pode ser aplicado a soluções ácidas, básicas e a solventes orgânicos, olhando as recomendações do fabricante para a faixa de pH em que as amostras devem apresentar afim de não danificar o cristal.

Figura 11 - Esquema representativo do funcionamento do ATR.



Fonte: o Autor.

Os espectros usados foram obtidos por Espectrômetro MIR equipado com acessório de ATR (Figura 12).

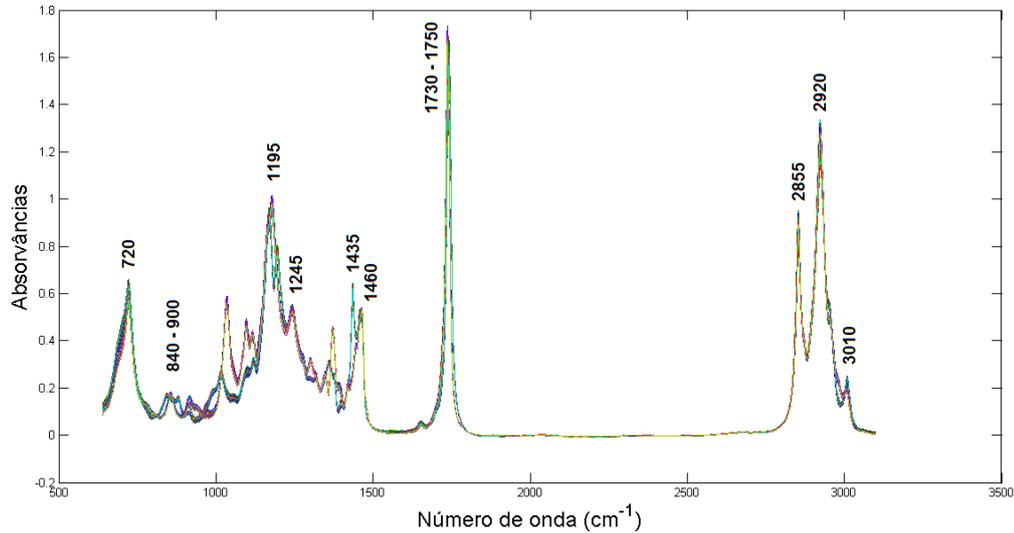
Figura 12 - Espectrômetro MIR equipado com acessório de ATR de ZnSe.



Fonte: o Autor.

Observando os espectros na Figura 13, constata-se uma elevada semelhança não sendo possível distinguir os componentes da mistura associados aos tipos de biodiesel.

Figura 13 - Espectros MIR de B100 de BES, BMS, BEFU, BMFU, BEPM e BMPM.



Fonte: o Autor.

A sobreposição dos perfis espectrais (Figura 13), na região das impressões digitais (1500-900  $\text{cm}^{-1}$ ), está ligada ao fato da maioria das moléculas darem sinal característico nesta região, associado a cada tipo de biodiesel em relação ao tipo de óleo usado (óleo de soja, OFU e óleo de pinhão manso) e rota (álcool etílico e álcool metílico).

As posições dos picos no espectro MIR, estão correlacionadas com as estruturas moleculares, isto é, força da ligação e massa dos átomos da ligação (Eq. 5), derivada da Lei de Hooke para molas em vibração (PAVIA, 2012).

$$\tilde{\nu} = \frac{1}{2\pi c} \sqrt{\frac{K}{\mu}} \quad (5)$$

onde:

$\tilde{\nu}$  é a frequência natural de vibração de uma ligação em  $\text{cm}^{-1}$ ;

$c$  é a velocidade da luz;

$K$  é a constante da força em dinas/s;

$\mu = \frac{m_1 m_2}{m_1 + m_2}$ , massas de átomos em gramas. Assim, ligações mais fortes têm constante de

força  $K$  maior, vibram em frequências mais altas do que ligações mais fracas envolvendo as mesmas massas enquanto que ligações entre átomos de massas maiores (massas reduzidas

maiores,  $\mu$ ) vibram em frequências mais baixas do que ligações entre átomos mais leves envolvendo o mesmo tipo de ligação.

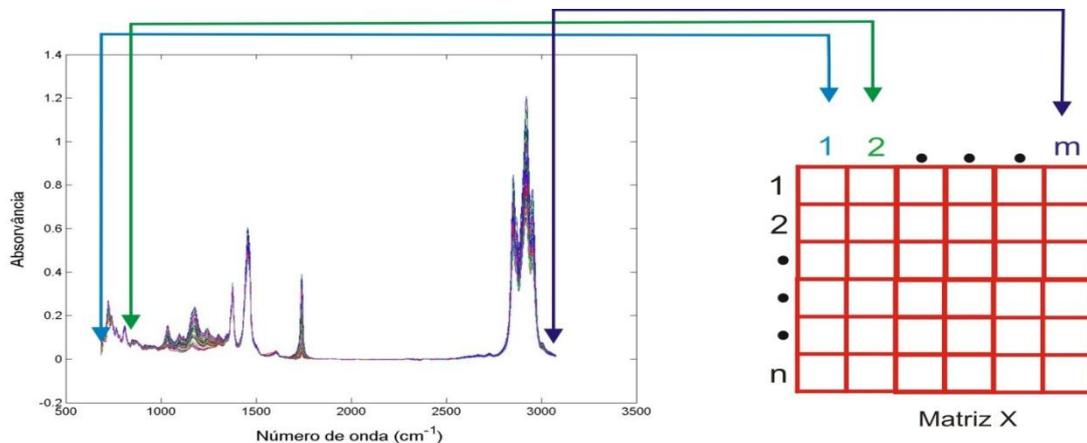
Portanto, as sobreposições observadas estão relacionadas com o fato de muitas moléculas apresentarem fortes absorções nesta região para além dos espectros MIR estarem munidas de informação rica, no que diz respeito ao perfil dos grupos funcionais (SMITH, 2011).

Com análise visual não é possível extrair informação devido à sobreposição dos perfis espectrais. Portanto, é necessário o uso de métodos quimiométricos para classificar os componentes da mistura.

#### 4 TRATAMENTO DE DADOS

No uso de métodos quimiométricos é necessário um conjunto de medidas (ou variáveis), podendo assim extrair informações químicas relevantes do sistema em estudo. Para dados de primeira ordem, como os tratados nesta Dissertação, o conjunto de medidas deve ser organizado na forma de uma matriz de dados  $\mathbf{X}$ , denominada matriz de dados originais. Nesta matriz, as linhas representam as  $n$ -amostras e as colunas as  $m$ -variáveis, sendo cada variável representada por um vetor coluna e cada amostra por um vetor linha (Figura 14) (KOWALSKI, 1975).

Figura 14 - Representação da construção da matriz  $\mathbf{X}$  para modelagem multivariada



Fonte: o Autor.

Na construção do conjunto de dados, isto é, transformação dos dados (amostras) é necessário ter em conta que o sinal medido consiste em duas contribuições, a saber:

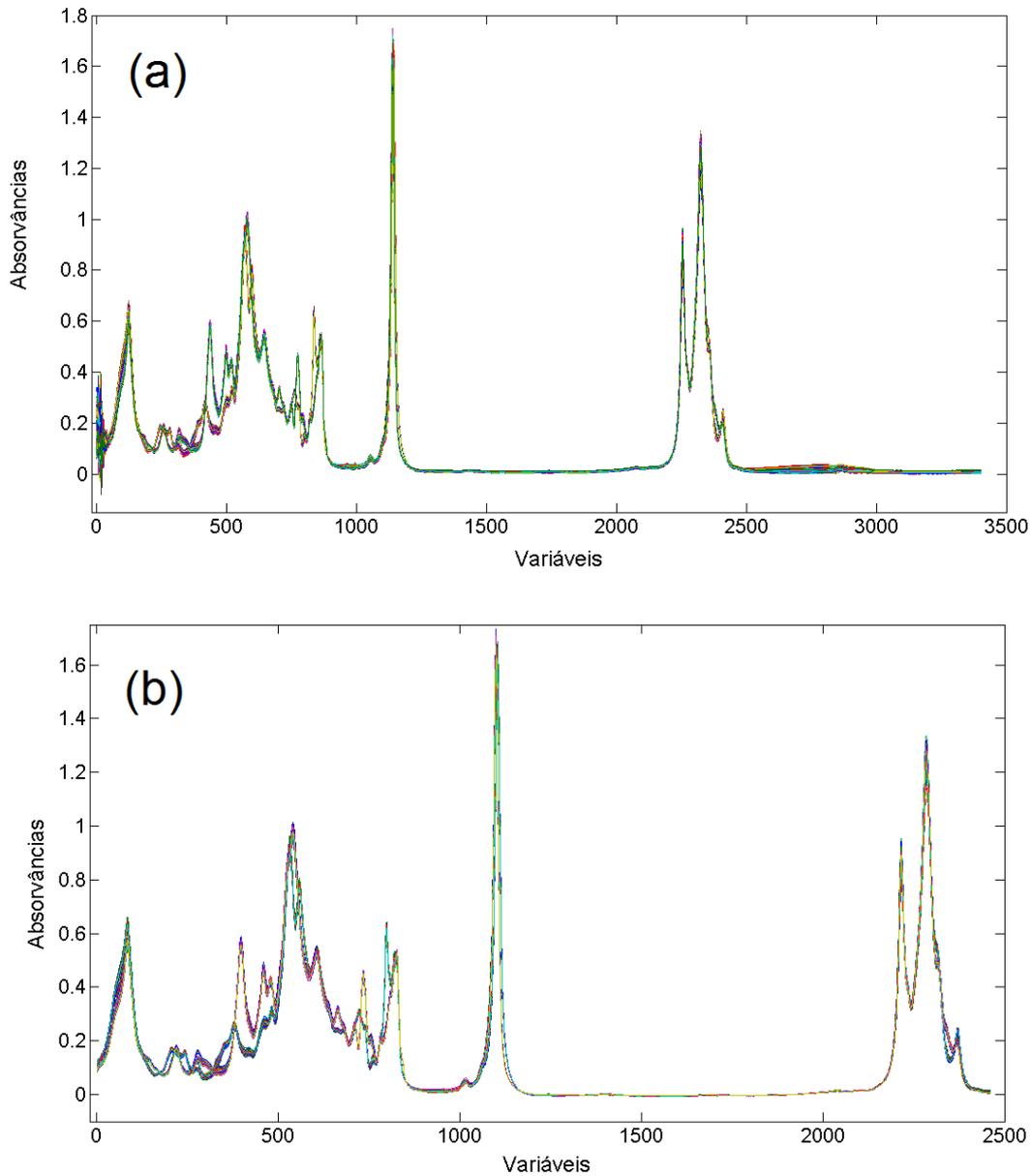
- ✓ contribuição determinística: sinal verdadeiro (informação relevante + indesejada) e
- ✓ contribuição estocástica: ruído aleatório.

Portanto as variações sistemáticas indesejadas e aleatórias devem ser reduzidas ou removidas. As variações aleatórias (ruído experimental) são reduzidas por meio de técnicas de alisamento para aumentar a razão sinal/ruído enquanto que as variações indesejadas são reduzidas ou eliminadas por meio de correções da linha de base no caso de dados de espectroscopia MIR.

Em medidas baseadas em espectros MIR, geralmente é possível observar ruído aleatório e variação da linha de base. O método mais simples para superar tais variações não informativas é a correção da linha de base, através do algoritmo *baseline* (Figura 15).

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

Figura 15 - Espectros MIR com flutuação na linha de base (a) e corrigidos (b).

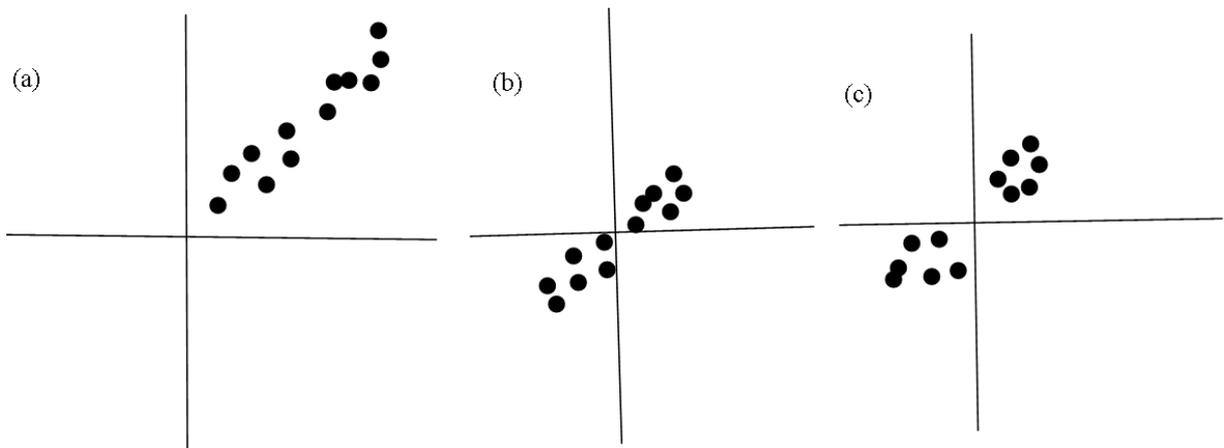


Fonte: o Autor.

Após a transformação das amostras no conjunto de dados é necessário avaliar a necessidade de pré-processar os dados, num processo comumente chamado de pré-processamento de dados (aplicado as variáveis), portanto antes de serem analisados pelo método quimiométrico adequado de acordo com o objetivo da pesquisa. Os dois tipos de pré-processamentos mais conhecidos na literatura são a centralização dos dados na média e o autoescalamto (OTTO, 1999). Na perspectiva de OTTO, (1999) a Figura 16 apresenta dados hipotéticos sem pré-processamento, centrados na média e autoescalados.

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

Figura 16 - Dados hipotéticos antes do pré-processamento (a), centrados na média (b) e autoescalados (c)



Fonte: o Autor.

Centrar os dados na média, didaticamente é transladar o eixo para o centro dos dados, isto é, a origem fica no ponto médio e a projeção das amostras no sistema n-dimensional é feita em função do novo centro (Figura 16b). Este pré-processamento é comumente usado para dados de espectros contínuos, oriundos de técnicas como espectroscopia no infravermelho ou no ultravioleta.

Por sua vez autoescalamamento significa centrar todos dados (grandeza maior e menor) na média, seguida pela normalização dos mesmos ao valor do desvio padrão do conjunto de dados de modo a conferir mesmo peso (diminuir a influência do dominante), isto é, além da subtração da média, os dados são divididos pelo desvio padrão de cada subconjunto de dados. No autoescalamamento ocorre uma adaptação do conjunto de dados de forma a uniformizar a variância de cada variável e tem como resultado dados com média zero e variância um (KOWALSKI, 1975).

O autoescalamamento é usado geralmente para grandezas diferentes, como valor de pH, temperatura, concentração de metais, havendo a necessidade de centrar os dados na média e normalizar em função do desvio para todos parâmetros terem mesmo peso. Em dados de espectroscopia MIR, NIR e Raman, não é recomendável empregar autoescalamamento, uma vez que pode ampliar o ruído (dar peso ao sinal menor incluindo ruído).

Assim, o pré-tratamento de dados, processo que engloba a transformação e pré-processamento de dados, visa remover as fontes de variação indesejadas que não são removidas naturalmente durante a análise de dados (FERREIRA; ANTUNES; MELGO,

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

1999). Portanto após o pré-tratamento, a matriz de dados pode ser tratada pelos métodos quimiométricos.

### 4.1 Quimiometria

A terminologia *chemometrics* foi introduzida em 1972, pelo sueco Svante Wold e americano Bruce R. Kolwalski. A partir do surgimento de grupos de química analítica iniciou-se o tratamento de dados de experimentos analíticos com métodos matemáticos, estatísticos e de lógica, surgindo então o termo Quimiometria (BRERETON, 2003a).

A quimiometria pode ser definida como a aplicação de métodos matemáticos, estatísticos e de lógica formal para o tratamento de dados químicos, de forma a extrair uma maior quantidade de informações e melhores resultados analíticos (BRUNS; FAIGLE, 1985). Com o surgimento da quimiometria, tornou-se possível o processamento e a interpretação de dados complexos. Assim, a quimiometria tem como grande contribuição tornar métodos matemáticos complexos práticos para utilização em laboratórios industriais e para análises de rotina (OTTO, 1999).

A quimiometria pode ser usada com a finalidade de:

- ✓ classificação de amostras: a partir de uma vasta gama de informações (medidas químicas ou espectrais) sobre uma série de objetos pretende-se encontrar agrupamentos de amostras (objetos) que são similares entre si e desta forma detectar tendências nos dados;
- ✓ calibração multivariada: procura-se estabelecer um modelo que relaciona uma série de medidas (químicas ou espectrais) realizadas em amostras com determinada propriedade (concentração);
- ✓ planejamento e otimização de experimentos: visa encontrar variáveis que mais afetam um determinado processo, assim como a interação entre elas.

Para estes fins podem ser usada à redução de dados, no qual o sistema estudado é representado da melhor forma possível, mantendo as informações iniciais contidas nas variáveis e a verificação da dependência entre variáveis, no qual se devem avaliar quais variáveis são dependentes entre si.

A calibração é definida como sendo uma série de operações que estabelecem, sob condições específicas, uma relação entre medidas instrumentais e valores para uma propriedade de interesse correspondente (BRERETON, 2000b).

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

Um modelo de calibração é uma função matemática ( $f$ ) que relaciona dois grupos de variáveis, uma delas denominada dependente ( $\mathbf{Y}$ ) e a outra denominada independente ( $\mathbf{X}$ ), (Eq. 6):

$$\mathbf{Y} = f(\mathbf{X}) = \mathbf{Xb} \quad (6)$$

Os dados usados nesta etapa de calibração são chamados de conjunto de calibração. Os parâmetros do modelo são denominados de coeficientes de regressão ( $\mathbf{b}$ ) determinados matematicamente a partir dos dados experimentais (HELLAND, 1998; MARTENS, 1996). Após a calibração é feita a validação.

Na etapa de validação, as variáveis independentes obtidas, para um outro conjunto de amostras, são utilizadas em conjunto com o coeficiente de regressão para calcular os valores previstos para a variável dependente. No conjunto de validação utilizam-se amostras cujas variáveis dependentes sejam conhecidas para que seja possível estabelecer uma comparação entre os valores previstos pelo modelo e os valores conhecidos previamente através de uma metodologia padrão, o que permitirá a avaliação sobre o desempenho do modelo de calibração proposto (MARTENS, 1996).

Para a construção de modelos de calibração, existem vários métodos que podem ser empregados, sendo que a função que ajusta as variáveis dependentes e independentes pode ser linear ou não, dependendo da complexidade do sistema em estudo (SEKULIC, 1993).

Dentre os métodos de calibração existentes, os mais difundidos são ainda os métodos de calibração univariada que também são conhecidos como calibração de ordem zero, ou seja, tem-se apenas uma medida instrumental para cada uma das amostras de calibração, isto é, para cada amostra tem-se apenas um escalar.

Todavia, a aplicação da calibração univariada é restrita, para amostras contendo interferentes em que medida é realizada diretamente por uma metodologia instrumental, uma vez que provocará desvios na determinação da propriedade de interesse, o que torna este método de calibração inviável (CHUI; ZUCCHINI; LICHTIG, 2001).

Na calibração multivariada, mais de uma resposta instrumental é relacionada com a propriedade de interesse. Os métodos de calibração multivariada possibilitam a análise mesmo na presença de interferentes, desde que esses interferentes estejam presentes nas amostras utilizadas para a construção do modelo de calibração dado que é onde se define o perfil das amostras a serem previstas pelo mesmo.

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

Outras possibilidades apresentadas por calibração multivariada são determinações simultâneas como PLS2-DA, a ser tratado nesta dissertação e análises de dados mesmo sem resolução que é feita através de uso de métodos de correção de sinal. Isso faz com que os modelos de calibração multivariada sejam uma alternativa quando os métodos univariados não são aplicáveis.

Na calibração multivariada de primeira ordem, a resposta instrumental é representada na forma de matriz, para métodos enquanto a propriedade de interesse, determinada por uma metodologia padrão, é representada por um vetor (MARTENS, 1996).

### 4.2 Métodos quimiométricos de classificação

Os métodos quimiométricos de classificação são agrupados normalmente em dois tipos:

- ✓ métodos de classificação não supervisionada ou conhecimento de padrões e
- ✓ métodos de classificação supervisionada ou reconhecimento de padrões.

#### 4.2.1 Métodos de classificação não supervisionada ou conhecimento de padrões

Estes métodos são usados para análise exploratória de dados. Os métodos de análise exploratória mais usados são: análise de componentes principais (PCA) e análise de agrupamentos hierárquicos (HCA).

##### 4.2.1.1 Análise de componentes principais (PCA)

A análise exploratória permite identificar tendências dos padrões, em formar agrupamentos de acordo com as semelhanças e diferenças entre amostras, partindo do conjunto de medidas realizadas. Análise de componentes principais (PCA) é um dos métodos bastante usado na análise exploratória.

A PCA é um método que permite a redução da dimensionalidade através da representação do conjunto de dados em um novo sistema de eixos, denominados componentes principais (PC), permitindo a visualização da natureza multivariada dos dados em poucas dimensões.

No espaço original, as amostras são pontos localizados em um espaço n-dimensional, sendo n igual ao número de variáveis. Com a redução de dimensionalidade proporcionada pela PCA, as amostras passam a serem pontos localizados em espaços de dimensões reduzidos definidos pelas PCs, por exemplo, bi- ou tridimensionais, porém, matematicamente

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

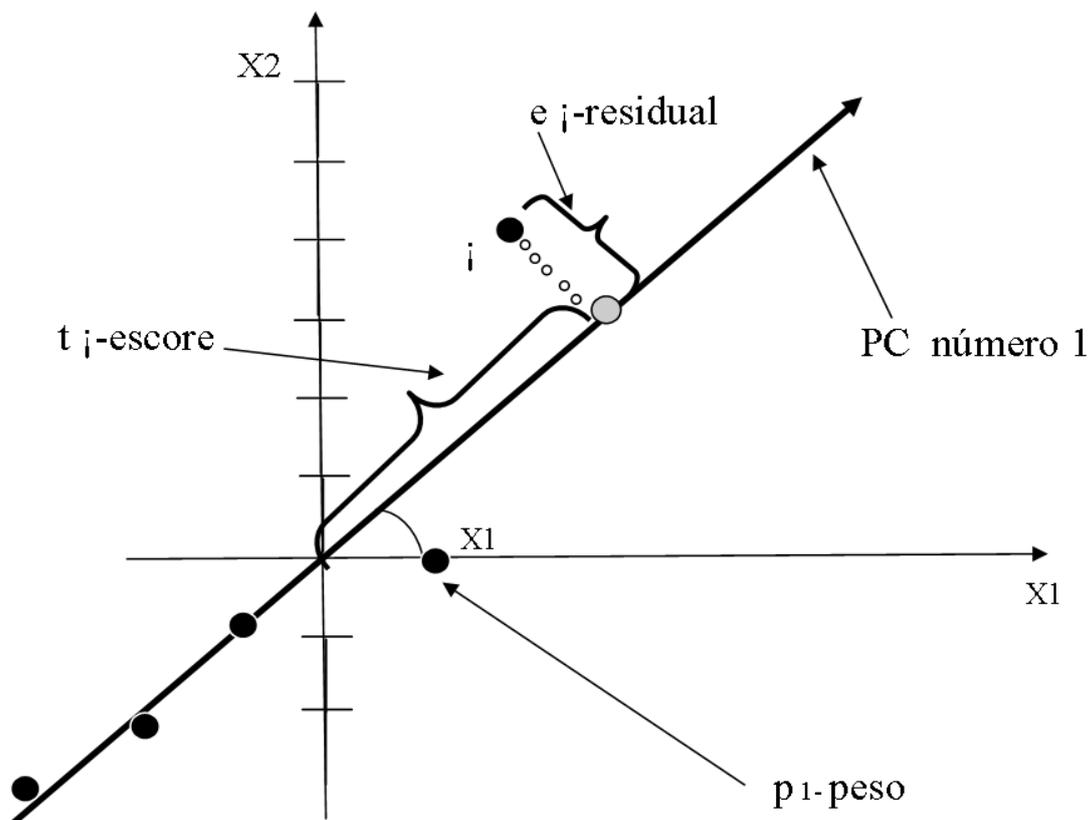
PCA a matriz de dados multivariados (matriz  $\mathbf{X}$ ) é decomposta em escores (T), pesos (P) e matriz de erros, isto é, resíduos (SAVORANI et al., 2013).

Por definição matemática pesos (P) são os cossenos dos ângulos formados entre o eixo da variável original e a PC, representando, portanto o quanto cada variável original contribui para uma determinada PC (Figura 17).

A primeira componente principal (PC1) é traçada no sentido da maior variância no conjunto de dados, a PC2 é traçada ortogonalmente à PC1, com intuito de descrever a maior percentagem da variância não explicada pela PC1 e assim por diante. Por sua vez, escores são as distâncias das projeções dos pontos (amostras) na PC até a sua origem, isto é, coordenadas das amostras no novo sistema de eixos (as PCs), onde quimiometricamente falando representam as relações de similaridade entre as amostras enquanto resíduo é a distância do ponto (amostra) em relação a PC (Figura 17).

Do ponto de vista quimiométrico, a avaliação dos pesos permite entender quais as variáveis que mais contribuem para os agrupamentos observados no gráfico dos escores, portanto, através da análise conjunta do gráfico de escores e pesos, é possível verificar quais variáveis são responsáveis pelas diferenças observadas entre amostras.

Figura 17 - Descrição gráfica do PCA



Fonte: o Autor.

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

A equação básica para PCA é representada pelas equações 7 ou 8 (WOLD; ESBENSEN; GELADI, 1987):

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + \dots + \mathbf{t}_F\mathbf{p}_F^T + \mathbf{E} \quad (7)$$

ou

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (8)$$

onde:

$\mathbf{X}$  é a matriz de dados

$\mathbf{t}$  são os escores

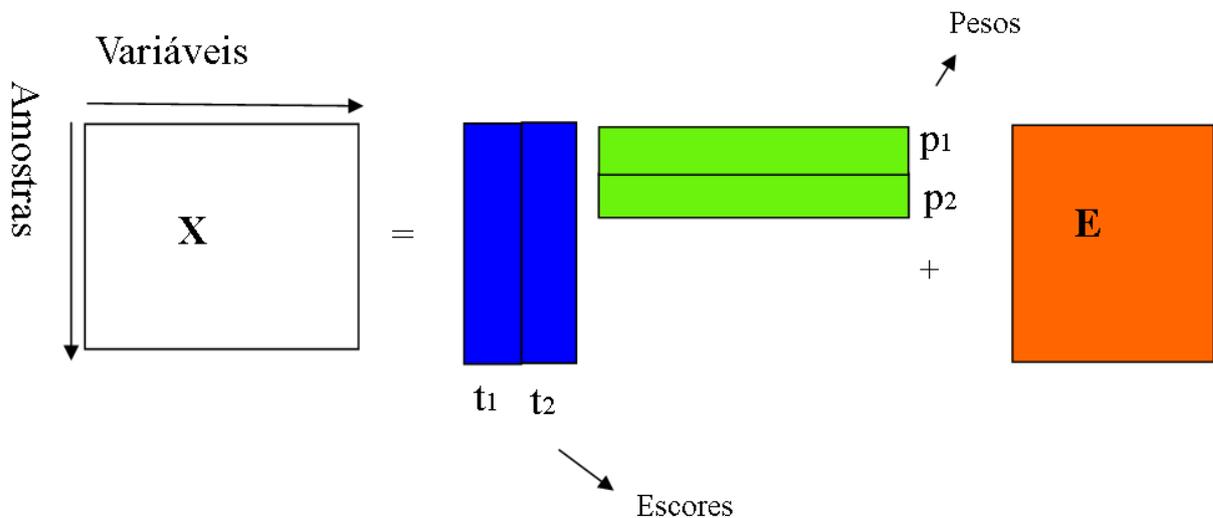
$\mathbf{p}$  são os pesos

$\mathbf{E}$  é a matriz de resíduos

$F$  é o número de componentes principais usados para descrever a matriz de dados.

Portanto as equações 7 e 8, podem ser descritas conforme a Figura 18:

Figura 18 - Representação matricial do PCA.



Fonte: o Autor.

O número de componentes principais a ser utilizado no modelo PCA é determinado pela percentagem da variância explicada. Assim, seleciona-se um número de componentes de tal maneira que a maior porcentagem da variância presente no conjunto de dados originais seja capturada (WOLD; ESBENSEN; GELADI, 1987).

A análise de PCA é feita geralmente usando o algoritmo *Nonlinear Iterative Partial Least Squares* (NIPALS). NIPALS é um método comumente utilizado para o cálculo das componentes principais de um conjunto de dados, onde os vetores dos pesos e dos escores são

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

calculados iterativamente, um de cada vez, em ordem decrescente de importância. O processo iterativo, para a PC1, é inicializado com uma primeira estimativa de escores, que pode ser a coluna de  $\mathbf{X}$  que tem maior variância.

Utilizando estes escores, calculam-se os pesos (Eq. 9):

$$\mathbf{p} = \frac{\mathbf{tX}}{\mathbf{t}^T \mathbf{t}} \quad (9)$$

e os pesos são normalizados para comprimento igual a 1. Após isso, os escores são calculados (Eq. 10):

$$\mathbf{t} = \frac{\mathbf{Xp}}{\mathbf{p}^T \mathbf{p}} \quad (10)$$

Esses valores de escores são comparados com os anteriores e se forem diferentes (dentro de um critério pré-estabelecido), os pesos são novamente calculados como mostrado acima. Esse processo continua até que os escores sejam semelhantes ou certos números de iterações tenham sido realizadas. Após a convergência, o produto  $\mathbf{tp}^T$  calculado a partir da equação 8 é subtraído de  $\mathbf{X}$ , obtendo-se o resíduo  $\mathbf{E}$  (equação 11).

$$\mathbf{E} = \mathbf{X} - \mathbf{tp}^T \quad (11)$$

O processo continua para o próximo PC, substituindo  $\mathbf{E}$  por  $\mathbf{X}$ .

### 4.2.1.2 Análise de agrupamentos hierárquicos (HCA)

A análise de agrupamentos por métodos hierárquicos, HCA é um método de classificação não supervisionada baseada na distância, que visa agrupar as amostras em classes de acordo com a similaridade dos participantes de uma mesma classe e nas diferenças entre os membros de classes diferentes (CORREIA; FERREIRA, 2007).

HCA representa graficamente os dados multidimensionais em esquema bidimensional. Os resultados são apresentados na forma de uma árvore hierárquica, chamada Dendrograma, onde o comprimento dos ramos da árvore representa o grau de similaridade entre os objetos.

Portanto, há duas maneiras de agrupar hierarquicamente os objetos:

- ✓ aglomerativa: considera inicialmente cada objeto como um grupo unitário e segue agrupando-os sistematicamente por ordem de similaridade;

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

- ✓ divisiva: considera inicialmente um único grupo de objetos e divide-o em dois ou mais grupos, começando pelos objetos que são mais dissimilares.

Assim, HCA é um método aglomerativo, onde se assume que as amostras próximas entre si no espaço multidimensional, sejam semelhantes em relação às variáveis consideradas. HCA e PCA são ferramentas quimiométricas amplamente utilizadas na triagem de dados analíticos, tornando explícitas as inter-relações entre variáveis e permitindo a visualização das informações latentes que não seriam observadas através de um tratamento univariado (CORREIA; FERREIRA, 2007).

PCA assim como HCA são métodos de classificação não supervisionada, portanto, não conseguem classificar novas amostras desconhecidas pelo modelo, sendo necessário o uso de métodos de classificação supervisionada de modo a reconhecer novas amostras.

### ***4.2.2 Métodos de classificação supervisionada ou reconhecimento de padrões***

As técnicas são ditas serem supervisionadas, pois supervisionam o desenvolvimento dos critérios de discriminação das amostras. Neste grupo de métodos se destacam: Modelagem Independente e Flexível por Analogia de Classe (SIMCA), K-Vizinhos mais Próximos (KNN) e Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA).

#### ***4.2.2.1 K-Vizinhos mais Próximos (KNN)***

Este método de classificação supervisionada calcula a distância de uma amostra a todas outras do conjunto de treinamento. A atribuição de uma classe é feita por votos usando a classe de amostras mais próximas como critério (BEEBE; PELL; SEASHOLTZ, 1998; FERREIRA, 2002).

Calcula-se a distância entre todas as amostras do conjunto de treinamento no espaço n-dimensional, portanto, os K- Vizinhos mais Próximos de cada amostra são escolhidos para votar, em que cada um deles, dá um voto para sua classe (BEEBE; PELL; SEASHOLTZ, 1998; FERREIRA, 2002).

A classe que recebe mais votos ganha a amostra, em caso de empate, a classe com menor distância acumulada fica com amostra. O processo é repetido para diferentes valores de K. O número de vizinhos selecionados no modelo final é escolhido com base no número de erros de classificação.

#### 4.2.2.2 Modelagem independente e flexível por analogia de classe (SIMCA)

Este método modela a localização e distribuição das classes, através do uso de componentes principais (PCs) e assume-se que os valores medidos para um grupo de amostras parecidas tenderão para uma distribuição uniforme e modelável (MASSART et al., 1997).

O SIMCA determina o número de PCs necessário para descrever a classe e constrói-se uma hipercaixa envolvendo as amostras de cada classe, onde os limites das mesmas são definidos com um dado nível de confiança. Após obter as fronteiras de cada classe, é avaliada se há sobreposição entre as mesmas.

A atribuição de uma amostra teste a uma dada classe é baseada na projeção da mesma no espaço dos escores e a sua distância às fronteiras da classe em questão, isto é, repetido para todas as classes e no final, se os modelos não tiverem tido o poder de discriminação suficiente, a amostra pode ser membro de mais de uma classe. Assim, pode acontecer que uma amostra possa ser anômala em relação ao conjunto de treinamento (suficientemente diferente) e neste caso, não pertencer a nenhuma das classes (MASSART et al., 1997).

Portanto, KNN e SIMCA são métodos baseados na suposição de que quanto mais próximas as amostras se assemelham entre si, mais próximas elas estarão no espaço multidimensional gerado pelas variáveis.

#### 4.2.2.3 Análise discriminante por quadrados mínimos parciais (PLS-DA)

Nos métodos de reconhecimento de padrões esta dissertação irá destacar o PLS-DA, tanto PLS1-DA quanto o PLS2-DA. PLS-DA é uma variação do algoritmo da regressão por PLS (WOLD; SJÖSTRÖM; ERIKSSON, 2001) para análise discriminante, que pode ser PLS1-DA e PLS2-DA.

O PLS-DA é um método multivariado utilizado para classificação de amostras onde é necessária a redução de variáveis e a variabilidade dentro do grupo é maior que a variabilidade entre grupos.

O que diferencia o PLS-DA do PLS é que o bloco  $\mathbf{Y}$  em um modelo de PLS-DA indica as classes as quais as amostras pertencem. Uma amostra tem valor igual a 1, se for um membro de uma dada classe de interesse e 0 se não pertencer a esta classe de interesse. Pode-se utilizar o PLS1-DA quando se tem apenas duas classes para distinguir e o PLS2-DA quando se tem mais de duas classes por fazer a classificação simultânea.

No PLS1-DA as variáveis contidas no bloco  $\mathbf{X}$  (dados espectrais) são relacionadas com as classes contidas em única coluna no vetor  $\mathbf{y}$ , enquanto que no PLS2-DA as variáveis

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

contidas no bloco  $\mathbf{X}$  (dados espectrais) são relacionadas com as classes contidas na matriz  $\mathbf{Y}$ , na qual cada coluna representa a classe de interesse.

Valores inteiros de classe são usados na única coluna no PLS1-DA ou são usados em cada coluna da matriz  $\mathbf{Y}$  para PLS2-DA, onde 1 é usado para classe de interesse e 0 para classe não de interesse, isto é, em ambos os casos existe uma variável dependente  $\mathbf{y}$  que pode assumir os valores 0 ou 1 (SERRANO-LOURIDO et al., 2012).

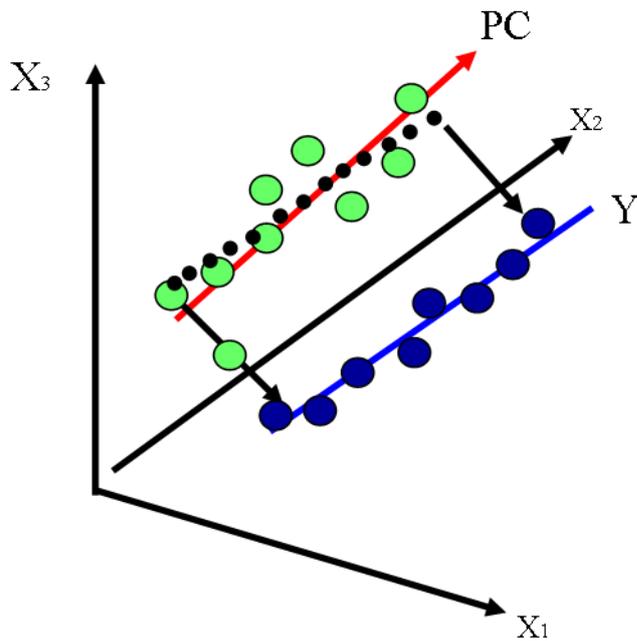
O valor de *threshold* é previsto entre 0 e 1, baseado no teorema de Bayes, visando a minimização de erros de previsão de classe através de um número apropriado de variáveis latentes (VLs) (ROGGO et al., 2007). Portanto o número de VLs escolhido para modelos PLS-DA, seguindo o critério de menor erro de previsão na validação cruzada *leave-one-out* e avaliação da variância explicada em  $\mathbf{X}$  e  $\mathbf{Y}$  bloco.

Assim, o PLS-DA procura encontrar VLs que descrevem a covariância nas amostras das futuras matrizes, que tem a correlação máxima com a classe de valores conhecida, dando menos peso para a classe-irrelevante ou variância do ruído (BARKER; RAYENS, 2003).

É importante frisar que os métodos supervisionados, uma vez que a análise é feita correlacionando o bloco  $\mathbf{X}$  (dados espectrais) com o bloco  $\mathbf{Y}$  (que pode ser um vetor ou matriz conforme o tipo do método), no PLS-DA existe um compromisso entre a explicação da variância em  $\mathbf{X}$  e encontrar a correlação com  $\mathbf{Y}$ . A correlação entre os dois blocos é feita através de uma leve rotação no eixo do PC para aumentar a correlação com  $\mathbf{Y}$  (Figura 19). Portanto a componente principal passa a se chamar variável latente (VL).

VL são as novas variáveis provenientes da redução do espaço multivariado através da criação de um novo conjunto de variáveis.

Figura 19 - Correlação de PCA com PLS.



Fonte: o Autor.

Quando se tem um modelo de classificação, dois parâmetros estatísticos são indispensáveis para avaliar o seu desempenho: a sensibilidade e a especificidade. Sensibilidade (SEN) é taxa de verdadeiros positivos que mede quão bem o modelo é capaz de classificar corretamente as amostras da sua classe. Especificidade (ESP) é a habilidade do modelo em identificar amostras que não são desta classe e que foram corretamente classificadas como não sendo (WESTERHUIS et al., 2008). A sensibilidade e especificidade são calculadas de acordo com as equações 12 e 13, respectivamente (XU et al., 2012):

$$\mathbf{SEN} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}} \quad (12)$$

$$\mathbf{ESP} = \frac{\mathbf{TN}}{\mathbf{TN} + \mathbf{FP}} \quad (13)$$

onde:

TP é número de verdadeiros positivos,

FN é número de falsos negativos,

TN é número de verdadeiros negativos e

FP é o número de falsos positivos.

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

Estes dois parâmetros são os mais importantes nos métodos de classificação porque refletem a eficiência do modelo em detrimento da precisão que mascara a má eficiência do modelo, dado que a precisão é calculada através da equação 14:

$$\text{Precisão} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (14)$$

A variável importante na projeção (VIP) é analisada com intuito de estudar as variáveis que foram mais relevantes na separação de cada classe de amostras em relação a outras classes de amostras. VIP escores tem valor maior que 1 (CHONG; JUN, 2005). VIP é um parâmetro que acumula a importância de cada variável  $j$ , sendo refletida por  $w$  variável latente, portanto a medida de VIP  $v_j$  é definida conforme a equação 15 (MEHMOOD et al., 2012):

$$v_j = \sqrt{\frac{\sum_{a=1}^A [SSa(w_{aj}/\|w_a\|)^2]}{\sum_{a=1}^A (SSa)}} \quad (15)$$

onde :

$SSa$  é a soma dos quadrados explicados pela VL. Uma vez que os pesos  $v_j$  é uma medida da contribuição de cada uma das variáveis de acordo com a variância explicada por VL de PLS, em que  $(w_{aj}/\|w_a\|)^2$  representa a importância da variável  $j$ .

$A$  é igual ao número de componentes para previsão, logo  $a = 1, 2, \dots, A$ . Em que “a” é índice de componentes (dimensão do modelo).

Por sua vez a variância explicada por cada VL pode ser calculado pela expressão  $(q^2 at'ata)$ ,  $v_j$  pode, alternativamente ser expresso como a equação 16 (MEHMOOD et al., 2012):

$$v_j = \sqrt{\frac{\sum_{a=1}^A [(q^2 at'ata)(w_{aj}/\|w_a\|)^2]}{\sum_{a=1}^A (q^2 at'ata)}} \quad (16)$$

onde:

$q$  é pesos do  $Y$  e

$t$  é escores.

### 4.2.2.3.1 Teorema de Bayes

O Teorema de Bayes utiliza informação prévia ou priori, de um dado conjunto de modo a obter informações de incerteza sobre as estimativas, uma vez que o intervalo de estimativa pode ser facilmente construído com base na distribuição posterior (CHEN; BAKSHI; GOEL,

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

2007). Porém, isto significa que ao estabelecer o modelo através das amostras do conjunto treinamento, estima-se um intervalo de probabilidade para o parâmetro das amostras desconhecidas a serem reconhecidas pelo modelo no caso do PLS-DA.

No entanto, os referidos intervalos tem um significado simples no contexto Bayesiano, uma vez que representam a região mais provável dos parâmetros do modelo com base em todas as informações contidas nas amostras usadas durante o treinamento do modelo, dado que estas amostras definem o perfil das amostras a serem reconhecidas pelo modelo.

Em métodos baseados em Teorema de Bayes, os parâmetros do modelo ( $\theta$ ) são considerados ser estocástica em vez de uma quantidade desconhecida determinista no quadro tradicional. Importa realçar que padrões estocásticos são aqueles que têm origem em processos não determinísticos, com origem em eventos aleatórios.

Antes de qualquer medição é obtido um valor prévio acerca da incerteza nos parâmetros do modelo o qual é representado por uma probabilidade de distribuição ( $P(\theta)$ ), conhecida como distribuição a priori. Considerando os parâmetros ( $\theta$ ), a distribuição condicional dos dados medidos ( $D$ ) é chamada por probabilidade dos dados  $P(D/\theta)$ .

O conhecimento prévio acerca dos parâmetros desconhecidos é combinado com a informação dos dados observados ( $D$ ) numa rigorosa probabilidade de maneira a usar o Teorema de Bayes. A informação atualizada é representada pela distribuição posterior (equação 17) (CHEN; BAKSHI; GOEL, 2007):

$$P(\theta/D) = \frac{P(D/\theta)P(\theta)}{P(D)} \quad (17)$$

onde a expressão  $P(D)$  representa a distribuição de probabilidade marginal dos dados ( $D$ ), tida como evidência.

A distribuição posterior combina todas as informações disponíveis, após a realização da modelagem com novo dado. Esta operação resume a incerteza atual sobre os parâmetros e são usados para todos os problemas de estimação de parâmetros.

Para encontrar uma estimativa do ponto ideal de ( $\theta$ ), é necessário especificar uma função de perda ( $\theta, \hat{\theta}$ ) para medir as consequências de discrepância entre o parâmetro verdadeiro ( $\theta$ ) e o seu valor estimado ( $\hat{\theta}$ ). Assim, o processo de estimativa Bayesiana pode ser resumido em três passos fundamentais:

- i. Especifica a distribuição prévia ou a priori e a probabilidade dos dados;

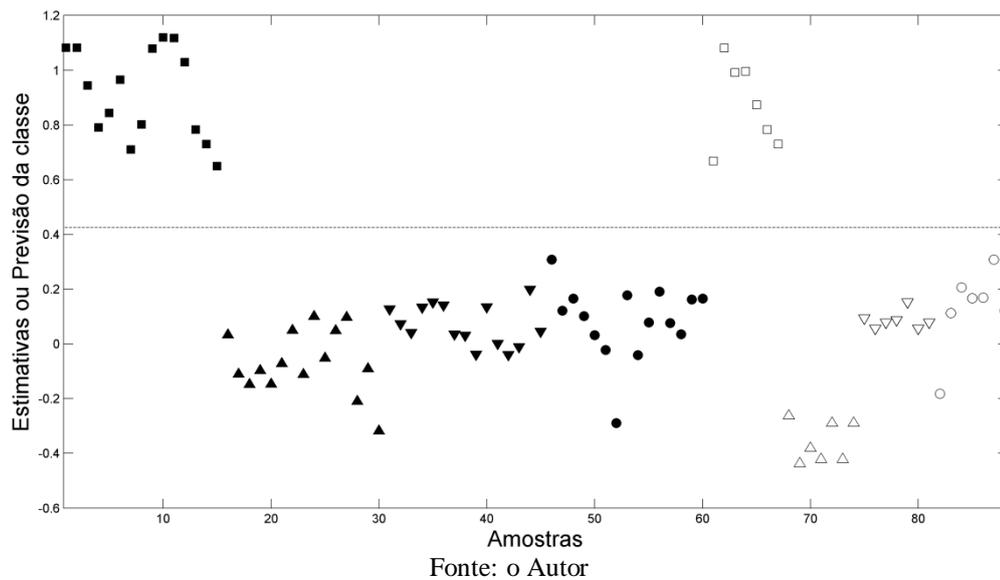
## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

- ii. Calcula a distribuição posterior pela regra de Bayes, onde geralmente apenas o numerador do lado direito da equação (17) é de interesse, porque o denominador é apenas uma normalização constante;
- iii. Especifica uma função de perda e obtém a estimativa ideal com base na função de perda e distribuição posterior.

### 4.2.2.3.2 Representação do resultado de PLS-DA

O resultado de PLS1-DA pode ser expresso através de um gráfico de estimativas, onde as amostras acima do valor de *threshold* indicam que foram classificadas como pertencentes a classe de interesse. As amostras abaixo do valor de *threshold* indicam que pertencem a classe que não seja de interesse. A Figura 20 mostra resultado do modelo PLS1-DA.

Figura 20 – Exemplo da representação do resultado modelo PLS-DA.



Para terminar, apresenta-se de forma resumida o princípio, vantagens e limitações de cada método de classificação (Tabela 2).

## CAPÍTULO 4- ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS

**Tabela 2** - Resumo sobre princípio, vantagens e limitações dos métodos quimiométricos de classificação citados.

<b>Métodos quimiométricos de classificação citados</b>				
<i>Métodos de classificação não supervisionada</i>		<i>Métodos de classificação supervisionada</i>		
HCA	PCA	SIMCA	KNN	PLS-DA
Baseia-se na distância.	Baseia-se nas PCs.	Baseia-se nas PCs.	Baseia-se na distância.	Baseia-se nas VLs.
Vantagem: não requer conhecimento <i>a priori</i> sobre a classificação ou informações sobre a existência de agrupamentos entre as amostras ou variáveis.	Vantagem: não requer conhecimento <i>a priori</i> sobre a classificação ou informações sobre a existência de agrupamentos entre as amostras ou variáveis.	Vantagem: consegue prever novas amostras desconhecidas pelo modelo.	Vantagens: consegue prever novas amostras desconhecidas pelo modelo; as amostras pertencem somente a uma classe, pelo fato de ser um método discriminante.	Vantagens: consegue prever novas amostras desconhecidas pelo modelo; pode classificar uma ou mais classes de interesse no mesmo modelo; as amostras pertencem somente a uma classe, pelo fato de ser um método discriminante.
Limitação: não prevê novas amostras desconhecidas pelo modelo.	Limitação: não prevê novas amostras desconhecidas pelo modelo.	Limitações: necessita de conhecimento <i>a priori</i> das classes; as amostras podem não pertencer a nenhuma classe, pertencer a uma classe ou várias classes ao mesmo tempo, uma vez que é um método modelativo.	Limitação: necessita de conhecimento <i>a priori</i> das classes.	Limitação: necessita de conhecimento <i>a priori</i> das classes.

Fonte: o Autor.

## **CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL**

### **5 CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL**

#### **5.1 Objetivos Específicos**

Este capítulo teve como objetivo os seguintes pontos:

- ✓ produzir 22 lotes diferentes de B100 de BES, BMPM, BMFU, BMS, BEFU e BEPM;
- ✓ construir modelos PLS1-DA e PLS2-DA para classificar os tipos de biodiesel (B100) em relação ao tipo de óleo e rota (tipo de álcool), usado na produção através de dados de espectroscopia MIR;
- ✓ obter 22 lotes diferentes de B5 partindo de amostras de B100 produzidas de BES, BMPM, BMFU, BMS, BEFU e BEPM;
- ✓ construir modelos PLS-DA para classificar os tipos de biodiesel (B5) em relação ao tipo de óleo e álcool usado na produção de B100 através de dados da espectroscopia MIR;
- ✓ validar os modelos PLS-DA desenvolvidos, através dos valores da sensibilidade e especificidade, para cada parâmetro nas amostras de conjunto teste na classe de interesse.

#### **5.2 Metodologia**

##### ***5.2.1 Procedimentos da produção do biodiesel***

Como a maioria dos processos químicos, a produção de biodiesel pode ser generalizada em três principais etapas sequenciais: pré-tratamento, reação e purificação. Na etapa de pré-tratamento, os resíduos, partículas que possam estar em OFU são removidos por meio de filtração. Caso o teor de ácidos graxos livres seja maior que 4% (m/m), é considerado alto nível de acidez, o que pode induzir a reação de saponificação, gerando sabão em vez de biodiesel (RINCÓN; JARAMILLO; CARDONA, 2014).

Assim, constatou-se que os óleos de soja e de OFU, neste trabalho, estavam com acidez aceitável para fazer a transesterificação por catálise básica. No entanto, para biodiesel de

## **CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL**

pinhão manso, o teor de ácidos graxos livres estava acima de 4% (m/m). Assim, o óleo de pinhão manso, foi submetido à pré-tratamento através da esterificação de ácidos graxos livres no processo chamado catálise ácida, seguida da transesterificação por catálise básica, Visando a conversão de ácidos graxos livres e a parte dos triglicerídeos em ésteres metílicos ou etílicos.

A esterificação como método de pré-tratamento em produção do biodiesel, pode ser combinado com a transesterificação por catálise básica visando a obtenção de uma conversão completa de biodiesel (RINCÓN et al., 2014).

A transesterificação é o processo amplamente utilizado na conversão de triglicerídeos em ésteres monoalquílicos derivados de ácidos graxos de cadeia longa. Este processo consiste basicamente em promover a mistura entre álcool de cadeia curta e óleo vegetal, auxiliado pelo emprego de um catalisador, sob determinadas condições de temperatura, tempo e agitação (DIAS et al., 2014).

A produção do biodiesel, foi feita seguindo a metodologia otimizada no Laboratório de Biocombustíveis do Instituto de Química da Universidade Federal de Uberlândia. Os biodieseis de soja e de OFU, tanto metílico assim como etílico, foram produzidos através da transesterificação por catálise básica. Por sua vez os biodieseis de pinhão manso tanto metílico quanto etílico, foram obtidos através da catálise dupla, onde em primeiro lugar foi usada a catálise ácida com intuito de converter os ácidos graxos livres em ésteres metílicos e etílicos, seguida da transesterificação por catálise básica.

### *5.2.1.1 Procedimento experimental da produção do biodiesel por catálise básica*

Os biodieseis de soja e de OFU, tanto metílico assim como etílico, foram produzidos à temperatura ambiente, na razão de 10:3 (m/m) do óleo em relação ao álcool e 1% em massa do hidróxido de potássio em relação à massa do óleo.

Em todas as misturas para produção de cada tipo de biodiesel, adicionou-se em primeiro lugar, o álcool no hidróxido de potássio até homogeneizar a nível visual. Após homogeneizar o alcóxido foi misturado com óleo. Cada mistura permaneceu sob agitação por barra magnética, na chapa de aquecimento, durante 1 hora em frasco de vidro (Figura 21).

## CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL

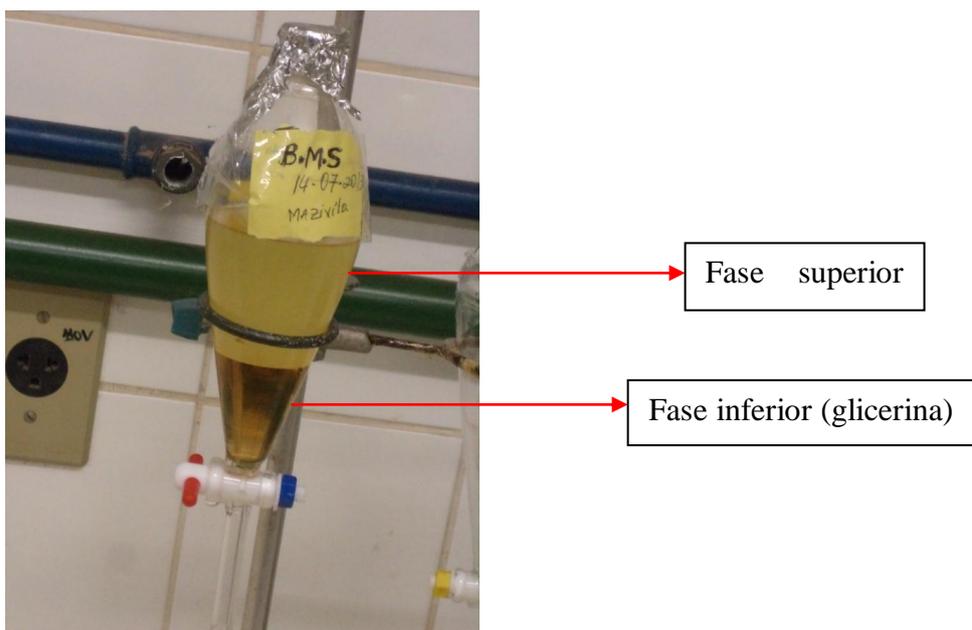
Figura 21 - Mistura do óleo e alcóxido na chapa de aquecimento com agitação magnética.



Fonte: o Autor.

Em seguida, a mistura foi transferida para um funil de decantação e foi mantida em repouso durante 1 hora até a separação da fase rica em glicerina (fase inferior) e da fase rica em ésteres metílicos de ácidos graxos (fase superior), (Figura 22).

Figura 22 - Separação de fases, biodiesel (superior) e glicerina (inferior).



Fonte: o Autor.

## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Foi retirada a fase inferior do funil de decantação e em seguida foi feita a purificação do produto obtido através da lavagem com água destilada quente (temperatura aproximada de 85 °C), com intuito de remover os resíduos, sub produtos e excedentes do álcool ou catalisador. A água de lavagem foi drenada juntamente com outros resíduos e colocados em frasco para posterior tratamento (Figura 23).

Figura 23 - Purificação do biodiesel através da lavagem (ésteres metílicos ou etílicos de ácidos graxos).



Fonte: o Autor.

Após a etapa de purificação, os ésteres metílicos assim como etílicos de ácidos graxos, foram introduzidos num balão de fundo redondo, onde foram levados a um sistema de destilação a pressão reduzida para um processo de secagem que foi realizado a 100 mmHg, 90 rpm e 80 °C por um período de 1 hora (Figura 24).

## CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL

Figura 24 - Secagem do biodiesel à pressão reduzida.



Fonte: o Autor.

O produto final foi transferido para um frasco de vidro envolto em papel alumínio e mantido ao abrigo da luz a temperatura ambiente para posterior análise. No entanto, quando o processo de purificação do biodiesel não for eficaz, a presença de glicerina livre pode provocar a emissão da acroleína, substância altamente tóxica para o ambiente. A acroleína se forma quando o glicerol sofre a oxidação junto ao biodiesel. Além disso, a presença da glicerina no biodiesel pode gerar baixo desempenho nos motores atuais (MITTELBAACH, 1996).

Ai a glicerina pode ser encontrada dispersa na forma de gotículas, a presença de sabões residuais pode interferir, aumentando a concentração da glicerina no biodiesel devido à formação de emulsões (SHARMA; SINGH; UPADHYAY, 2008).

## CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL

### 5.2.1.2 Procedimento experimental da produção do biodiesel por catálise dupla

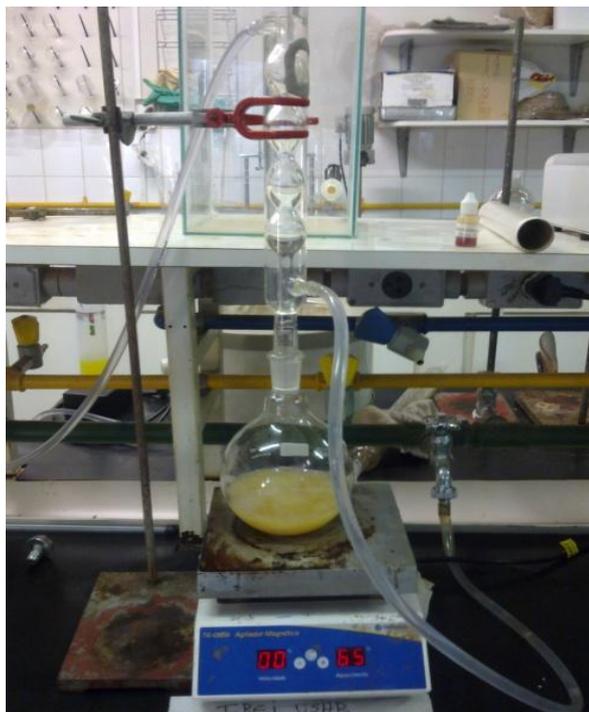
Os biodieseis de pinhão manso foram produzidos através da catálise dupla (esterificação/transesterificação). No entanto, a catálise dupla consiste basicamente em duas etapas. Em primeiro lugar realiza-se a catálise ácida, com intuito de converter os ácidos graxos livres e parte dos triglicerídeos em ésteres metílicos ou etílicos.

Após a redução dos ácidos graxos livres da matéria-prima, inicia-se a segunda etapa que é a catálise básica, responsável pela maior conversão dos triglicerídeos em ésteres metílicos assim como etílicos.

A metodologia para catálise ácida (primeira etapa) também foi otimizada no próprio laboratório, que consistiu na proporção de 1:1 (m/m) de óleo em relação ao álcool e 3% em massa do ácido sulfúrico (catalisador) em relação à massa do óleo.

Adicionou-se o ácido sulfúrico ao álcool e posteriormente a mistura foi adicionada ao óleo num balão de fundo redondo de vidro. Após a mistura formar única fase, permaneceu sob refluxo em ebulição por um período de 1 hora (Figura 25).

Figura 25 - Sistema para a catálise ácida metílica e etílica.

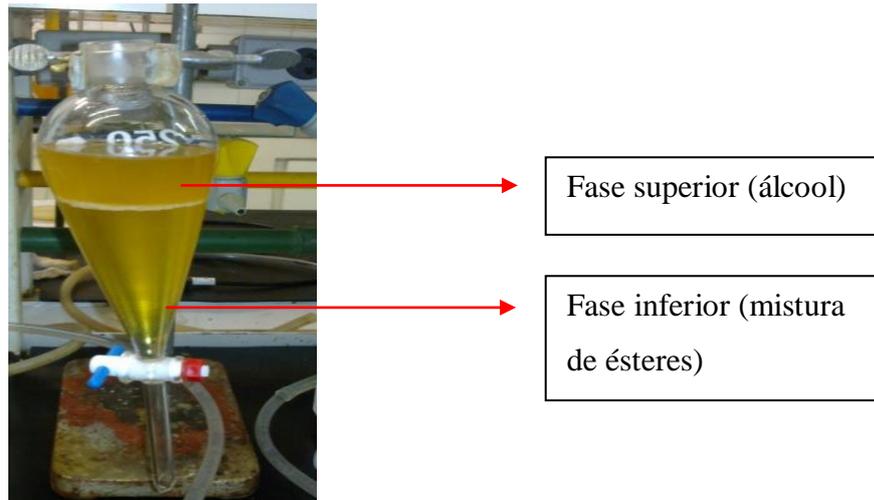


Fonte: o Autor.

## CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL

Decorrido o tempo, transferiu-se a mistura para um funil de decantação durante 1 hora para separação de fases (Figura 26).

Figura 26 - Mistura heterogênea entre ésteres e álcool, obtida após a reação em meio ácido.



Fonte: o Autor.

No caso de ésteres metílicos a fase inferior foi submetida à catálise básica enquanto que para ésteres etílicos a fase inferior foi colocada no balão de fundo redondo para destilação à vácuo, processo esse realizado a 100 mmHg, 90 rpm e 80 °C por um período de 1 hora, com intuito de remover álcool etílico em excesso. Depois desta etapa seguiu a catálise básica, conforme procedimento acima descrito.

### 5.2.2 Procedimentos quimiométricos

Os dados das amostras de B100 foram analisadas em primeiro lugar usando PCA. Após a análise exploratória, foram desenvolvidos seis modelos de classificação usando PLS1-DA para classificar 6 classes de B100, nomeados como: biodiesel etílico de soja (BES), biodiesel metílico de pinhão manso (BMPM), biodiesel metílico de óleo de fritura usado (BMFU), biodiesel metílico de soja (BMS), biodiesel etílico de óleo de fritura usado (BEFU) e biodiesel etílico de pinhão manso (BEPM).

Além dos seis modelos PLS1-DA, foi desenvolvido um modelo PLS2-DA para classificar simultaneamente 4 classes de B100.

## **CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL**

Os B100 usados foram produzidos no Laboratório de Biocombustíveis do Instituto de Química da UFU.

Na construção dos modelos B100 usando PLS1-DA, 22 amostras diferentes foram usadas (onde 15 amostras foram usadas para o conjunto de treinamento e 7 amostras para conjunto de teste) para cada uma das 6 classes discriminadas, perfazendo um total de 132 amostras, enquanto que para o modelo PLS2-DA uma vez que foram 4 classes o total de amostras foi de 88.

Após a classificação das classes de amostras de B100, em cada uma das 22 amostras de lotes diferentes, nas 6 classes foram preparadas as misturas de B5, onde o diesel puro foi fornecido pela TRANSPETRO S/A (Uberlândia, Estado de Minas Gerais, Brasil).

Depois de preparadas as 22 amostras de lotes diferentes de B5, das quais 15 amostras foram usadas para o conjunto de treinamento e 7 amostras para conjunto de teste, foram usadas no desenvolvimentos dos 6 modelos PLS1-DA para classificar as seis classes de B5, perfazendo um total de 132 amostras.

Os espectros foram obtidos usando um Espectrômetro FT-MIR SpectrumTwo (Perkin Elmer) utilizando um acessório Horizontal de Reflectância Total Atenuada (HATR) de Seleneto de Zinco (ZnSe) com as seguintes configurações: região de 4000 a 600  $\text{cm}^{-1}$ , resolução de 4  $\text{cm}^{-1}$  e 16 varreduras por espectro. Os espectros MIR, foram obtidos em quintuplicata para cada amostra, para a construção dos modelos PLS-DA, tanto para B100 quanto para B5.

Os espectros obtidos tiveram suas linhas de base corrigidas utilizando o método de baseline nas regiões espectrais de 1850 a 2500  $\text{cm}^{-1}$  e de 3150 a 4000  $\text{cm}^{-1}$  e foram analisados com PLS\_Toolbox 7.5 (Eigenvector Research) e em ambiente MATLAB 7.5 (MathWorks Inc.). Nos modelos construídos, o conjunto de dados foi dividido em um grupo para treinamento e outro para teste, de forma independente. Os grupos de amostras foram centrados na média e empregou-se validação cruzada por *leave one out*.

### **5.3 Modelos PLS-DA para classificar diferentes tipos de B100: BES, BMPM, BMFU, BMS, BEFU e BEPM**

Os espectros MIR das 132 amostras de B100 (Figura 27) apresentam perfis espectrais semelhantes para inspeção visual. De acordo com ALMEIDA et al., 2012, de 3000-2850  $\text{cm}^{-1}$

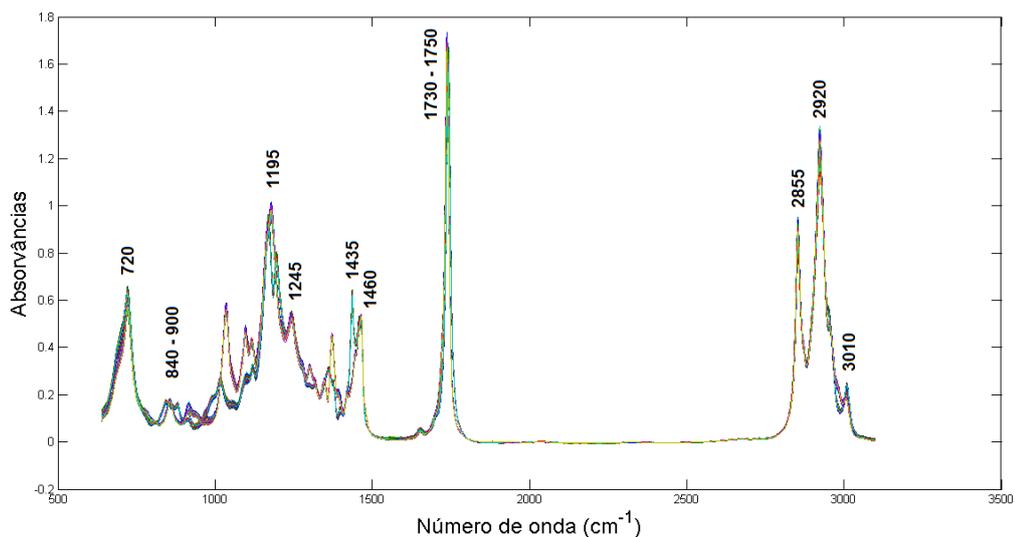
## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

é característica das vibrações dos estiramentos simétrico e assimétrico de CH, dos grupos metileno e metílicos terminais. Assim, a proeminente banda em  $2920\text{ cm}^{-1}$  é atribuída a vibração simétrica da banda  $\text{CH}_3$  e perto de  $2855$  e  $3010\text{ cm}^{-1}$  são atribuídas à vibração do estiramento simétrico e assimétrico do metileno, respectivamente; entre  $1730$ -  $1750\text{ cm}^{-1}$  é observada uma intensa banda associada ao estiramento da banda C=O do éster.

Picos em volta de  $1460\text{ cm}^{-1}$  são devidos à vibração no plano  $\text{CH}_2$  e  $\text{CH}_3$ ; perto de  $1435\text{ cm}^{-1}$  é atribuída a vibração da deformação angular no plano da banda CH de olefinas *cis*-dissubstituídas (SÁNCHEZ-ALONSO; CARMONA; CARECHE, 2012); perto de  $1245\text{ cm}^{-1}$  pode ser atribuído a vibração do estiramento axial antissimétrico de  $\text{CC}(=\text{O})\text{-O}$  de bandas ésteres, enquanto que picos perto de  $1195\text{ cm}^{-1}$  pode ser atribuído a vibração do estiramento axial assimétrico de bandas de O-C-C (SOARES et al., 2008).

Na região de baixa frequência a partir  $900$ - $700\text{ cm}^{-1}$ , as vibrações significantes foram atribuídas a deformação fora do plano de grupos  $\text{C}=\text{CH}-\tau$  (olefinico CH) e  $\omega$  (olefinico CH) de derivados insaturados (ALMEIDA et al., 2012).

Figura 27 - Espectros MIR de 132 amostras de B100: BES, BMPM, BMFU, BMS, BEFU e BEPM.



Fonte: o Autor.

A grande similaridade entre os perfis espectrais das amostras de B100 (Figura 27), torna-se necessário o uso de métodos quimiométricos de classificação para discriminar estes biodieseis.

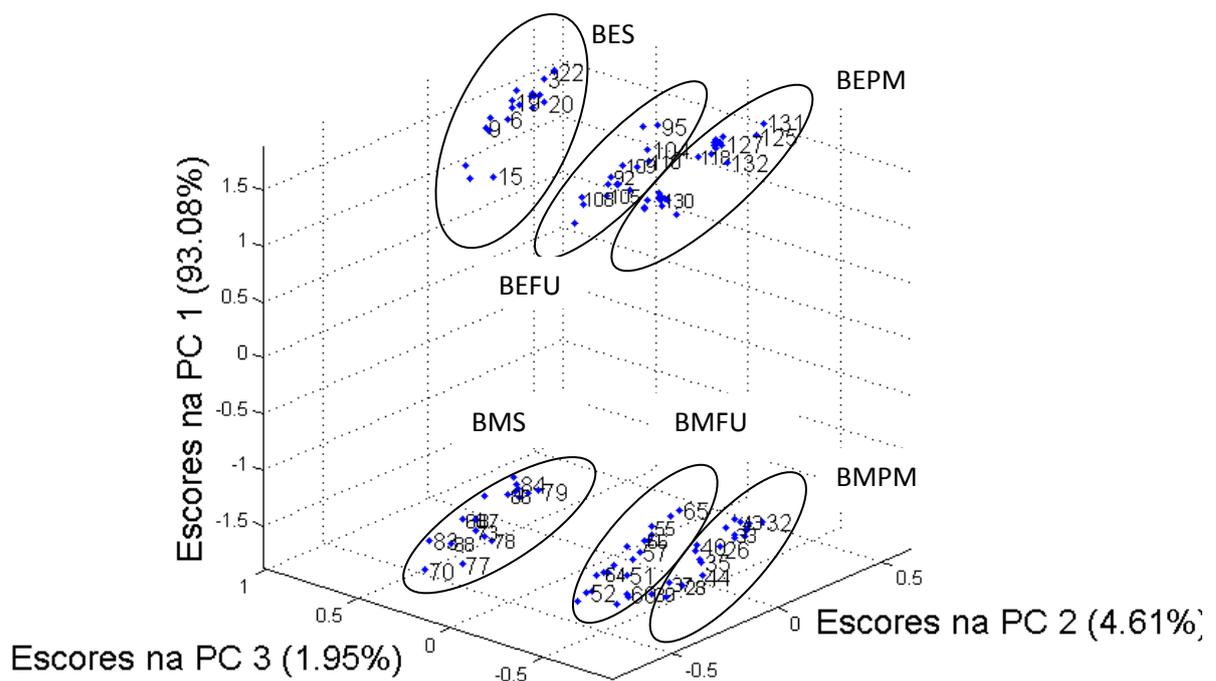
## CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL

### 5.3.1 Análise exploratória por Análise de Componentes Principais (PCA)

Análise exploratória das 2461 variáveis de cada uma das 132 amostras de B100 foi feita pelo PCA, com intuito de conhecer a tendência dos agrupamentos nos dados. Para este efeito 3 PCs foram selecionadas, os quais explicaram 99,64% de variância das amostras.

Na Figura 28, é possível visualizar escores do modelo PCA, com a projeção de  $PC1 \times PC2 \times PC3$ , onde pode ser observada a formação de seis agrupamentos diferentes das amostras de B100, na qual as amostras 1-22 pertencem a BES, de 23-44 BMPM, de 45-66 pertencem a BMFU, de 67-88 são BMS, de 89-110 são amostras BEFU e finalmente de 111 até 132 pertencem a BEPM.

Figura 28 - Escores de  $PC1 \times PC2 \times PC3$  de análise de PCA de diferentes tipos de B100.



Fonte: o Autor.

Era de esperar que os biodieseis de OFU, fossem semelhante com biodieseis de soja pelo fato de OFU ser proveniente do óleo de soja. Analisando a Figura 28, observa-se que tanto para biodieseis metílicos assim como etílicos, os biodieseis de pinhão e de OFU são semelhantes, isso se deve a composição química dos óleos. Óleo de pinhão e OFU apresentam

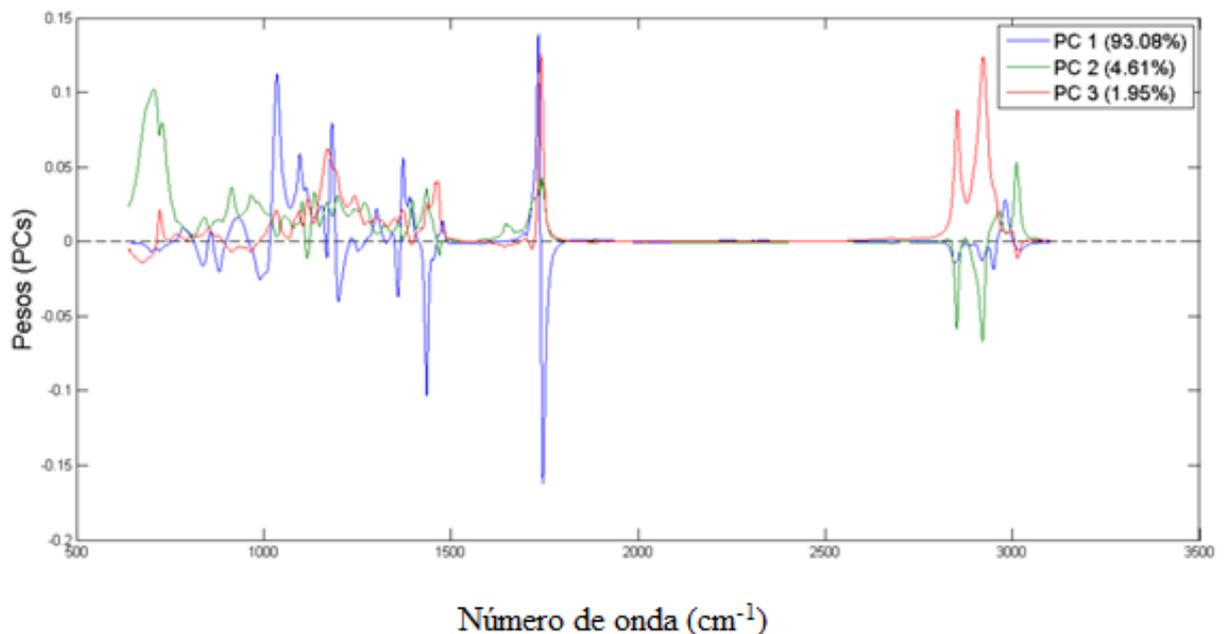
## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

mais ligações saturadas na sua estrutura, por sua vez o óleo de soja apresenta mais ligações insaturadas, logo ao sofrer a oxidação durante a fritura dos alimentos ocorre a quebra das ligações o que faz com que o OFU apresente mais ligações saturadas.

Importa referir que na Figura 28, observa-se que há maior variabilidade dentro do grupo em relação à variabilidade entre grupos, sendo mais indicado para realizar a análise quimiométrica usando PLS-DA.

As variáveis que foram mais determinantes na formação de agrupamentos são observadas por inspeção de pesos de cada variável para primeira, segunda e terceira PC (Figura 29).

Figura 29 - Gráfico de pesos de PC1, PC2 e PC3 versus número de onda (variáveis) para modelo PCA.



Fonte: o Autor.

Os pesos de PC1 mostram que as variáveis com maior contribuição estão em  $1740\text{ cm}^{-1}$ , atribuídas ao estiramento da banda C=O do éster,  $1435\text{ cm}^{-1}$  associados à vibração da deformação angular no plano da banda CH de olefinas *cis*-dissubstituídas e as bandas dos espectros MIR entre  $1030\text{-}1040\text{ cm}^{-1}$ .

No entanto as variáveis que tem grande contribuição para PC2 corresponde para regiões de espectros MIR em  $2920\text{ cm}^{-1}$ , atribuída a vibração simétrica da banda CH<sub>3</sub>,  $1730\text{-}1750$

## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

$\text{cm}^{-1}$  atribuídas ao estiramento da banda C=O e 700-720  $\text{cm}^{-1}$  atribuídas a deformação fora do plano de grupos C=CH- $\tau$  de derivados insaturados.

Finalmente, as variáveis com maior contribuição em PC3 estão relacionadas com a banda perto de 2855  $\text{cm}^{-1}$  referente à vibração do estiramento simétrico do metileno e em 2920  $\text{cm}^{-1}$ , atribuída à vibração simétrica da banda CH<sub>3</sub>.

A análise exploratória empregando PCA e espectros MIR de B100 permitiu a separação das amostras de acordo com o tipo de matérias-primas utilizadas. No entanto, o uso de PCA visava promover uma análise visual do comportamento das amostras bem como realizar uma exploração das variáveis importantes na distinção destas amostras.

### 5.3.2 Classificação de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA)

Depois de encontrar a similaridade entre amostras através de uma análise preliminar de extração de características nos dados de B100, usando PCA, foram desenvolvidos seis modelos PLS1-DA de modo a prever novas amostras desconhecidas de cada classe.

A Tabela 3 mostra os resultados de parâmetros de classificação obtidos por cada modelo PLS1-DA para classificação de B100. Os valores de Erro Quadrático Médio de Calibração (RMSEC), de Validação Cruzada (RMSECV) e de Previsão (RMSEP), mostram uma excelente concordância, o que significa que o erro de calibração foi uma boa estimativa do erro padrão de previsão, observada em amostras do conjunto de teste (ALMEIDA et al., 2013). Além disso, verifica-se que os modelos PLS1-DA desenvolvidos são eficientes na classificação de B100 empregando dados de espectroscopia MIR (Tabela 3).

**Tabela 3** - Parâmetros de classificação obtidos por PLS1-DA para classificação B100

Tipos de B100 /Parâmetros	BES	BMPM	BMFU	BMS	BEFU	BEPM
VL	7	6	8	8	9	7
RMSEC	0,1167	0,1241	0,1448	0,1049	0,1289	0,1244
RMSECV	0,1444	0,1574	0,2003	0,1380	0,1829	0,1546
RMSEP	0,1507	0,1894	0,2039	0,1499	0,1844	0,1563
<b>Sensibilidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b>Especificidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<i>Threshold</i>	0,4920	0,5056	0,5440	0,4600	0,4550	0,4855

Fonte: o Autor.

## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

A Figura 30 apresenta os valores de classe estimada para o conjunto de treinamento e de teste. Amostras que se encontram acima do valor de *threshold* são classificadas como pertencente à classe B100 e abaixo do valor de *threshold* como amostras B100 de classe que não seja de interesse.

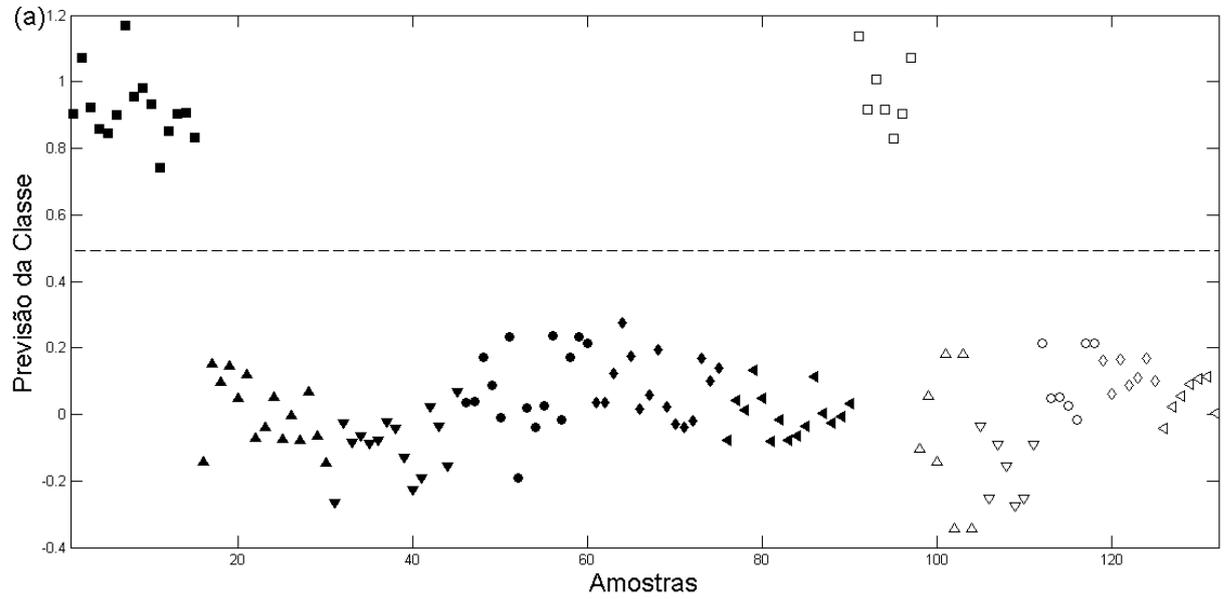
Nos modelos construídos, observa-se que as amostras do conjunto teste foram classificadas dentro das respectivas classes (Figura 30) conseqüentemente, todos os modelos mostram excelentes níveis de sensibilidade (capacidade do modelo em classificar corretamente as amostras da classe de interesse) e especificidade (habilidade do modelo em identificar as amostras que não são de interesse e que foram corretamente classificadas como não sendo) (Tabela 3) na qual 100% de amostras do conjunto de teste foram classificadas corretamente nos respectivos grupos.

Diante destes resultados obtidos, os seis modelos PLS1-DA desenvolvidos com intuito de classificar os tipos de biodieseis (B100) de acordo com o tipo de matéria-prima usado no seu processo de produção (óleo e álcool), podem ser usados pelos órgãos distribuidores e fiscalizadores, no sentido de satisfazer as medidas políticas segundo as quais os biocombustíveis produzidos na base de óleos usados na alimentação humana não devem ser subsidiados.

Analisando a Figura 30, é observada que a estimativa dos valores da classe para treinamento (símbolos cheios) e conjunto de teste (símbolos vazios) para discriminação de classe pelo modelo PLS1-DA: (a) BES dos outros biodieseis, (b) BMPM dos outros biodieseis, (c) BMFU dos outros biodieseis, (d) BMS dos outros biodieseis, (e) BEFU dos outros biodieseis e (f) BEPM dos outros biodieseis (■) BES, (▲) BMPM, (▼) BMFU, (●) BMS, (◆) BEFU, (◄) BEPM e (---) *threshold* de discriminação.

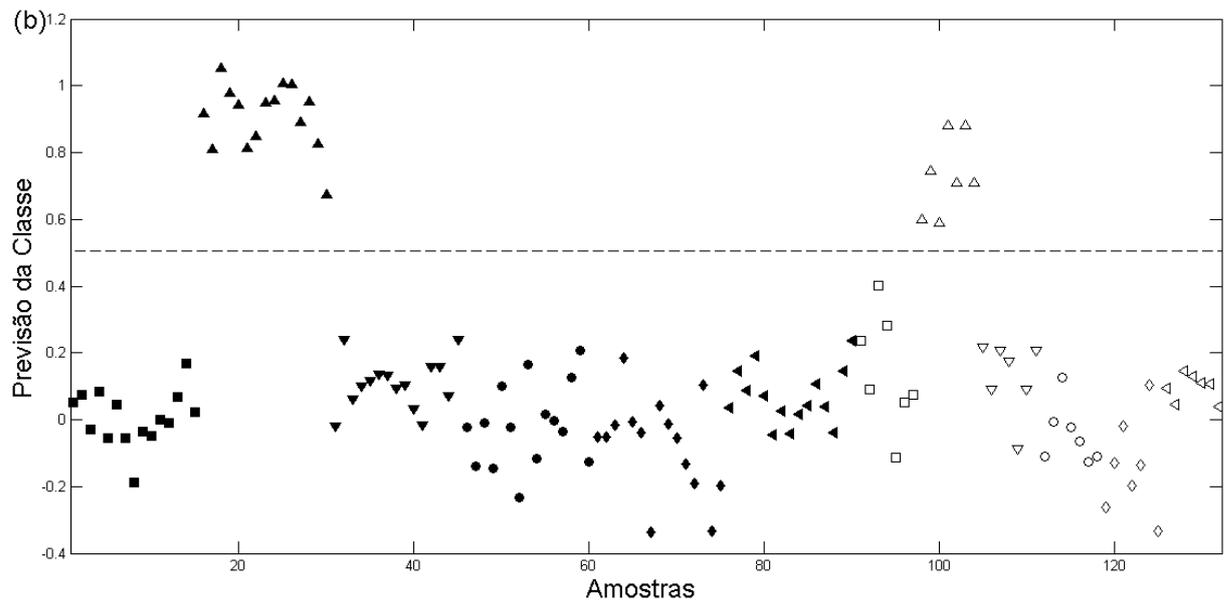
## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 30a – Estimativa de (■) BES.



Fonte: o Autor.

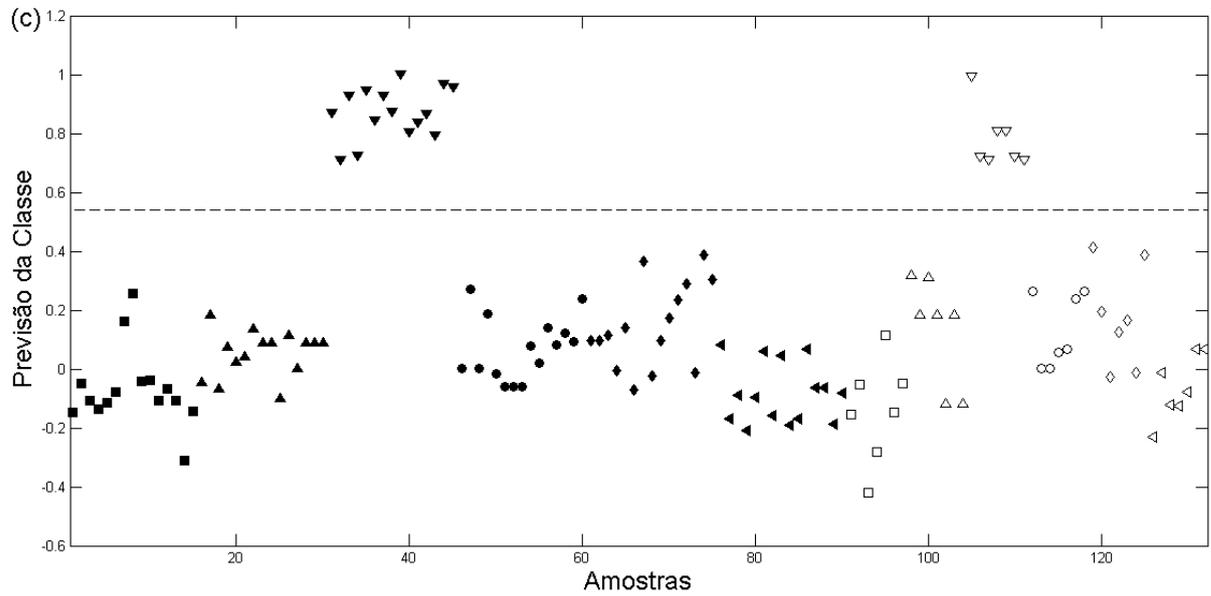
Figura 30b – Estimativa de (▲) BPPM.



Fonte: o Autor.

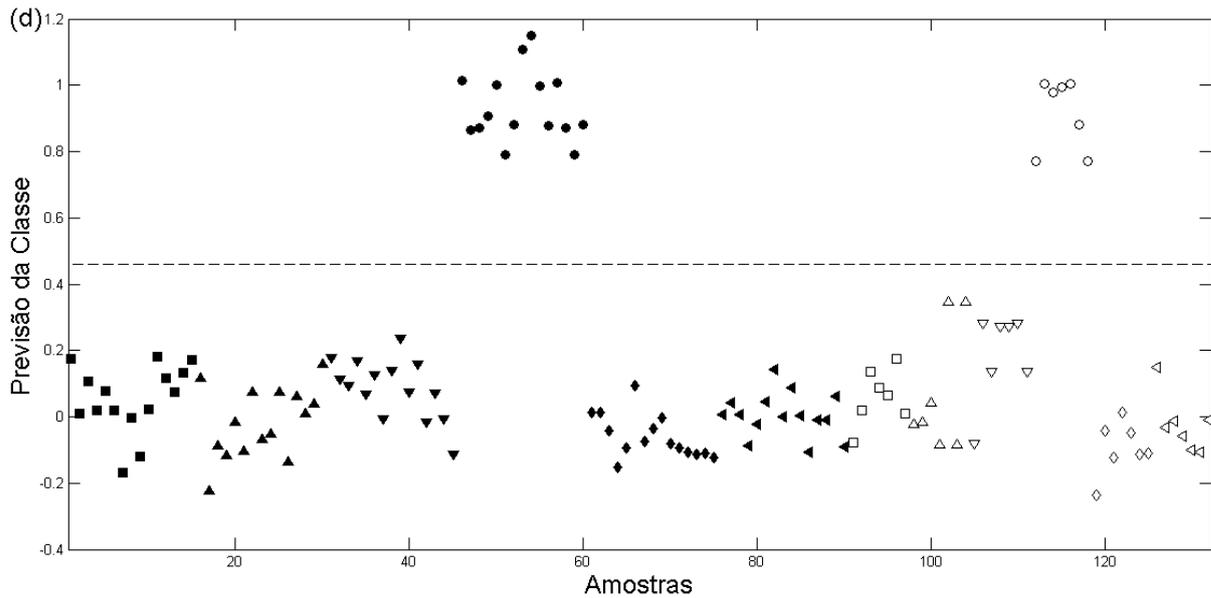
## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 30c – Estimativa de (▼) BMFU.



Fonte: o Autor.

Figura 30d – Estimativa de (●) BMS.



Fonte: o Autor.

## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 30e – Estimativa de (♦) BEFU.

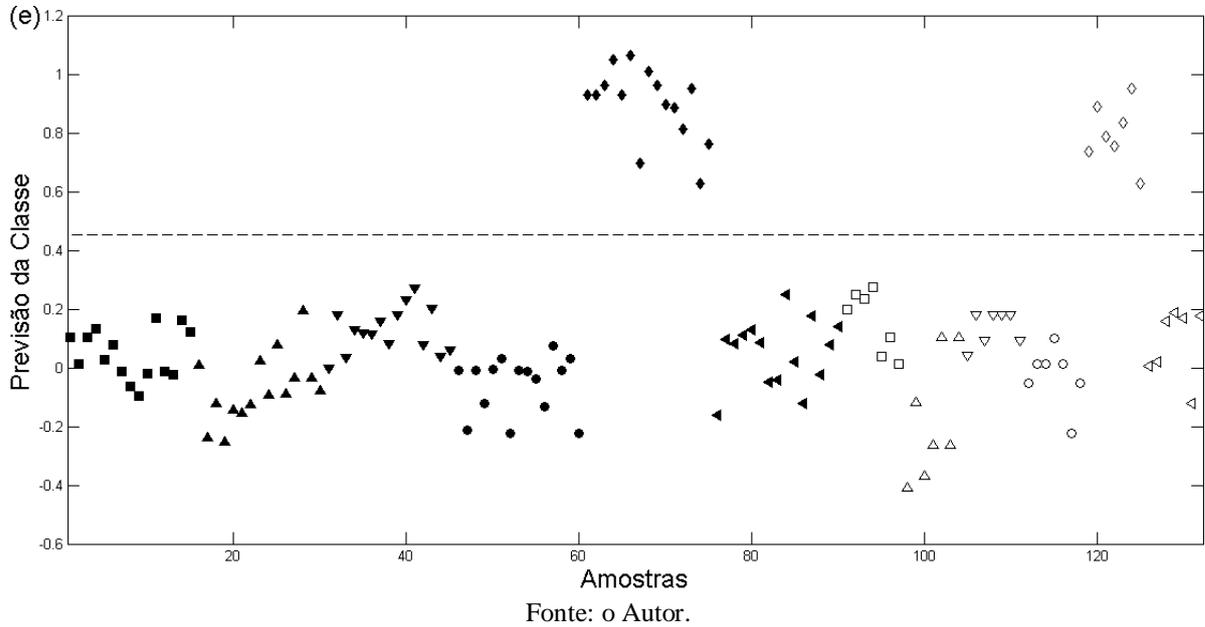
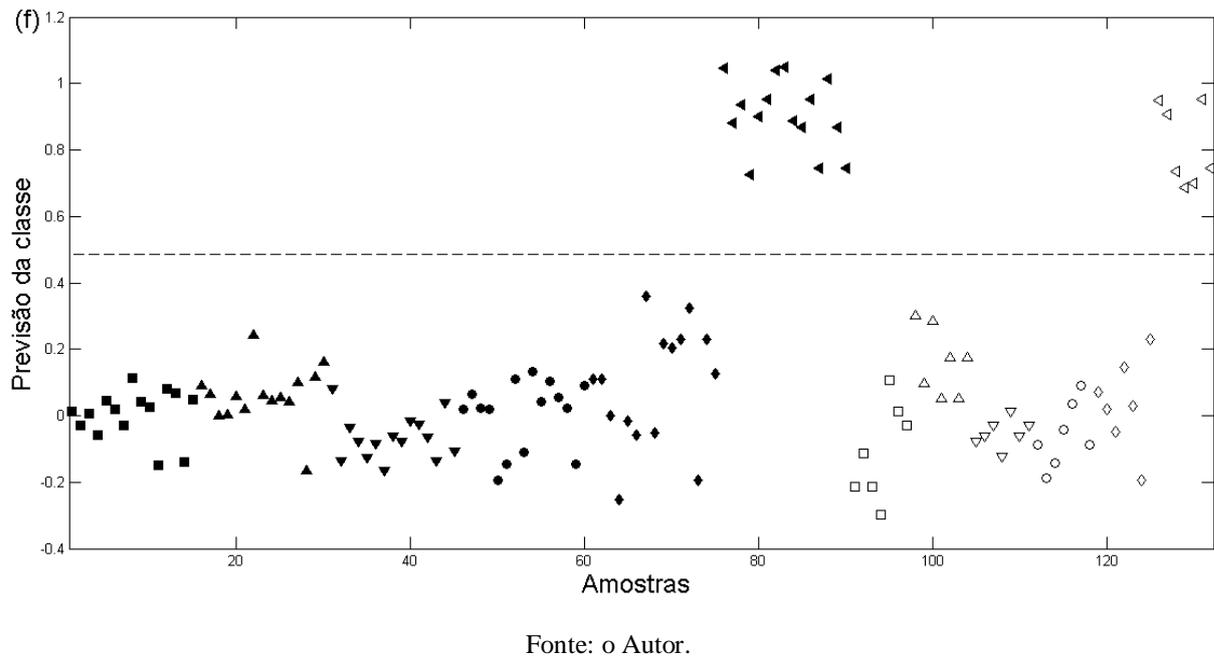


Figura 30f – Estimativa (◄) BEPM.



## **CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL**

### ***5.3.3 Classificação de Análise Discriminante por Quadrados Mínimos Parciais (PLS2-DA)***

Após o desenvolvimento de seis modelos PLS1-DA para classificação de seis classes de B100, foi desenvolvido um modelo PLS2-DA, para classificar quatro classes de B100, isto é, quatro tipos diferentes de B100, nomeadamente: BEFU, BMS, BMFU e BES. Este modelo visa responder a demanda da procura de um método capaz de identificar as amostras de biodieseis obtidas na mistura entre Óleo de Soja e OFU para além de classificar mais de 2 tipos de biodieseis no mesmo modelo.

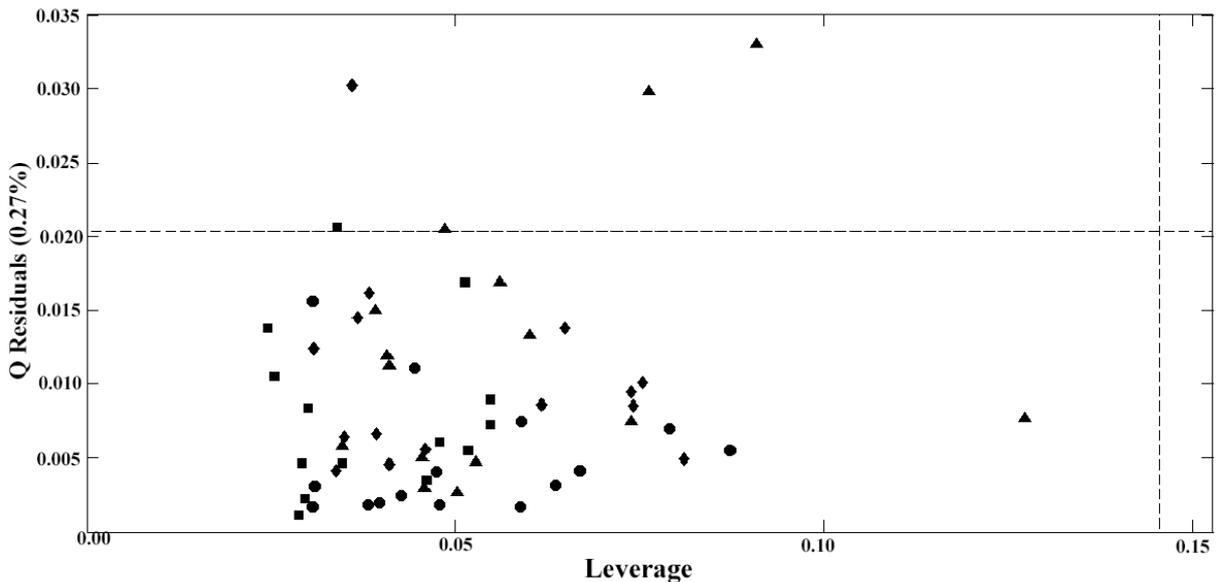
O emprego de PLS2-DA aliada à MIR como método analítico para classificação simultânea de biodieseis provenientes de óleos (óleo de soja e OFU) e rotas (metílica e etílica) diferentes, com intuito de demonstrar a grande aplicabilidade da calibração multivarida no que concerne a calibração simultânea de mais de uma propriedade de interesse.

A classificação simultânea usando PLS2-DA é vantajosa, para análise de várias propriedades de interesse numa única calibração (GALTIER et al., 2011) que pode ser usada na indústria com intuito de ganho de tempo.

Para verificar a presença de amostras anômalas no modelo PLS2-DA, foi usada a análise do gráfico dos valores de resíduos não modelados versus valores de leverage para cada amostra (Figura 31). Nenhuma amostra foi considerada anômala para um limite de confiança de 95,0%.

## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 31 - Gráfico de resíduos espectral versus leverage em nível de significância de 5 % para modelo PLS2-DA, onde (■ conjunto de treinamento) BEFU, (▲ conjunto de treinamento) BMS, (● conjunto de treinamento) BMFU, (◆ conjunto de treinamento) BES. Gráfico obtido com 3 VLs.



Fonte: o Autor.

A Tabela 4 mostra os resultados dos parâmetros de classificação obtidos para cada classe do modelo PLS2-DA.

**Tabela 4** - Parâmetros de classificação obtidos por PLS2-DA para classificação B100

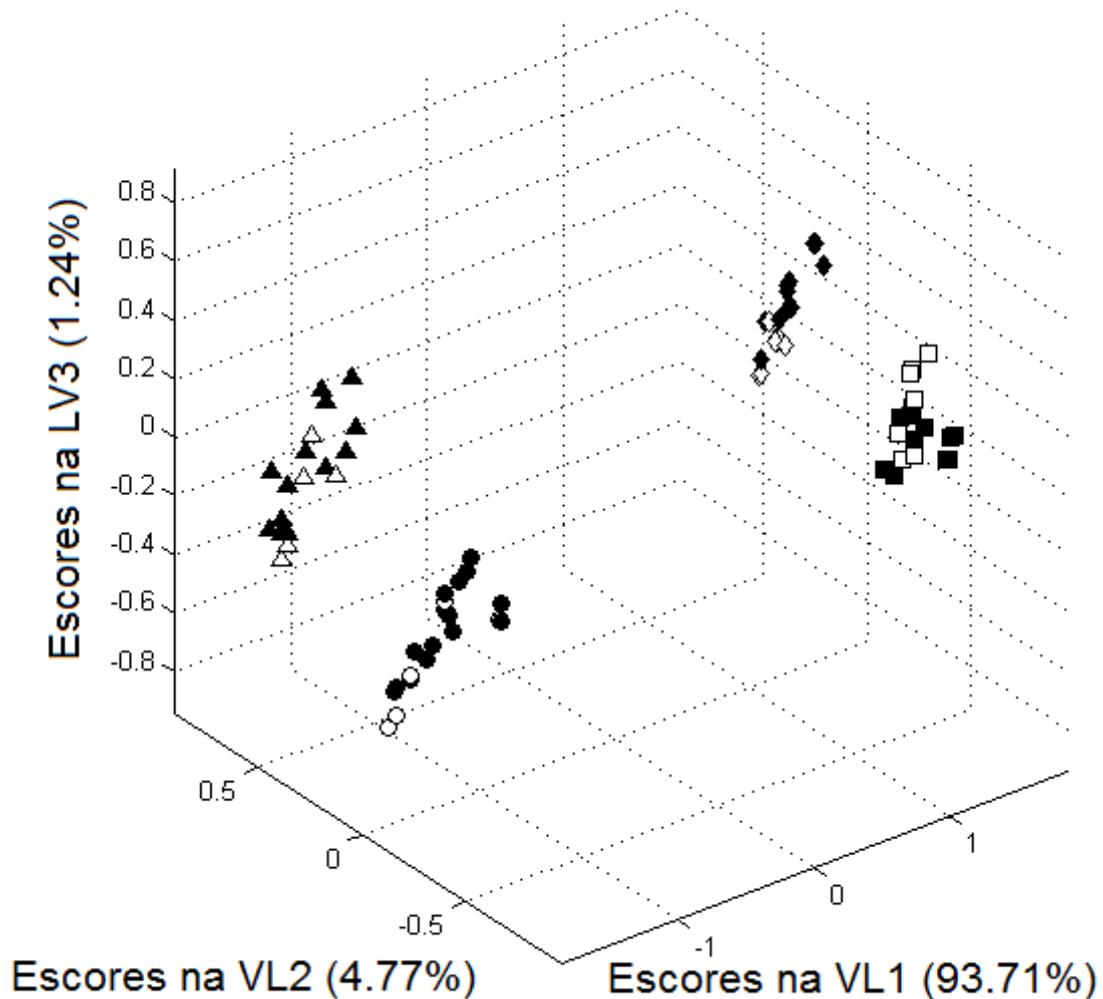
Classe	Número de VL	Sensibilidade	Especificidade
<b>BEFU</b>	3	<b>1,00</b>	<b>1,00</b>
<b>BMS</b>	3	<b>1,00</b>	<b>1,00</b>
<b>BMFU</b>	3	<b>1,00</b>	<b>1,00</b>
<b>BES</b>	3	<b>1,00</b>	<b>1,00</b>

Fonte: o Autor.

De acordo com a Figura 32, observa-se que o modelo PLS2-DA desenvolvido, as amostras foram classificadas nas respectivas classes, demonstrando a vantagem da classificação simultânea dos quatros tipos de B100.

## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 32 - Escores de análise do modelo PLS2-DA de tipos de biodieseis, onde: BEFU (■ conjunto de treinamento e □ conjunto de teste), BMS (▲ conjunto de treinamento e △ conjunto de teste), BMFU (● conjunto de treinamento e ○ conjunto de teste) e BES (◆ conjunto de treinamento e ◇ conjunto de teste).



Fonte: o Autor.

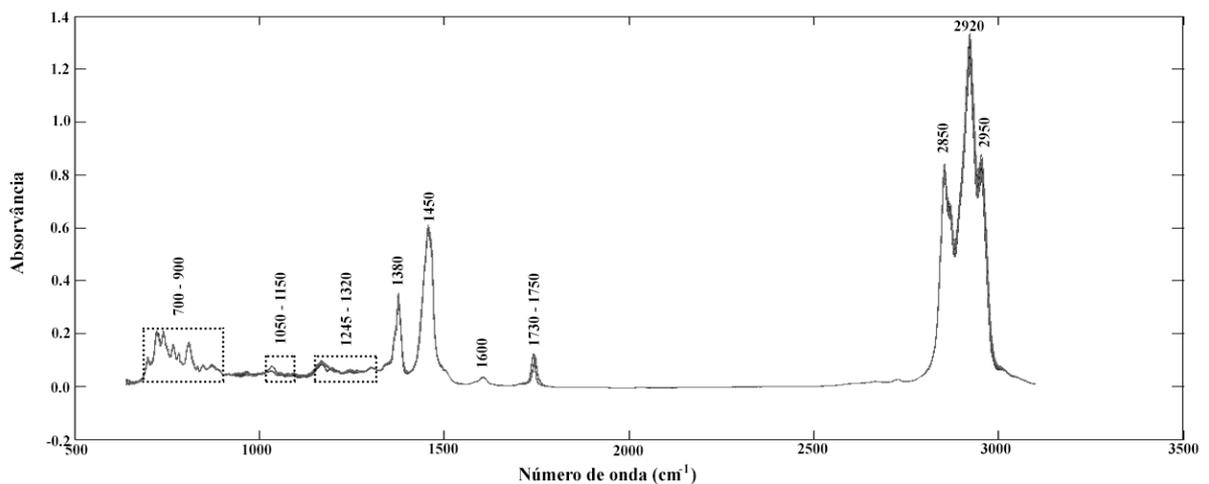
### 5.4 Modelos PLS1-DA para classificar diferentes tipos de B5: BES, BMPM, BMFU, BMS, BEFU e BEPM

Os espectros MIR de 6 tipos de B5 são mostrados na Figura 33. A proeminente banda em  $2920\text{ cm}^{-1}$  é atribuída a estiramento  $\text{CH}_3$ . A banda em  $2850\text{ cm}^{-1}$  atribuída ao estiramento  $\text{CH}_2$  e a banda em  $2950\text{ cm}^{-1}$ , corresponde a estiramento assimétrico  $\text{CH}_2$ . A banda em  $1730\text{ cm}^{-1}$

## CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL

1750  $\text{cm}^{-1}$  corresponde a C=O e o pico perto de 1600  $\text{cm}^{-1}$  é atribuído ao estiramento C=C. A banda de 1450  $\text{cm}^{-1}$ , pode ser atribuída a deformação assimétrica de  $\text{CH}_3$ , o sinal em 1400  $\text{cm}^{-1}$  é atribuído a vibração da deformação simétrica angular no plano e em 1380  $\text{cm}^{-1}$  é atribuída a deformação assimétrica  $\text{CH}_3$ . A região de absorção observada de 1245-1320  $\text{cm}^{-1}$  é atribuída a deformação simétrica angular no plano de =C-H. A banda em 900-700  $\text{cm}^{-1}$  é devido ao estiramento  $\text{C}_1\text{-C}_2$ , estiramento C-O e estiramento C-C perto de 1050-1150  $\text{cm}^{-1}$  (MIRANDA et al., 2014).

Figura 33 - Espectros MIR de 132 amostras de B5: BES, BMPM, BMFU, BMS, BEFU e BEPM.



Fonte: o Autor.

Espectros MIR das seis classes deferentes de cada tipo de B5 são muito similares (Figura 33). Só a inspeção visual dos espectros MIR, não permite identificar o tipo de biodiesel presente na mistura B5, de acordo com a matéria-prima usada na produção de B100. Devido a grande similaridade nos perfis espectrais, é necessário o uso de PLS-DA para discriminar estes tipos de amostras.

A Tabela 5 ilustra resultados dos parâmetros de classificação obtidos por cada modelo PLS1-DA.

## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

**Tabela 5** - Parâmetros de classificação obtidos por PLS1-DA

Tipos de B5 /Parâmetros	BES	BMPM	BMFU	BMS	BEFU	BEPM
VL	6	8	9	6	7	4
RMSEC	0,0928	0,1006	0,1293	0,1383	0,0857	0,0946
RMSECV	0,1103	0,1296	0,1685	0,1645	0,1065	0,1020
RMSEP	0,1542	0,1427	0,1895	0,1674	0,1363	0,1021
<b>Sensibilidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b>Especificidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<i>Threshold</i>	0,4090	0,6467	0,4665	0,4873	0,5291	0,3717

Fonte: o Autor.

A Figura 34 demonstra uma separação clara dos valores estimados, em detrimento disso é observada a discriminação da classe das outras cinco classes, mostrando que os espectros MIR contêm informações químicas que torna possível para PLS1-DA discriminar os seis tipos de misturas B5.

É interessante notar que *threshold* de discriminação não está no centro da diferença entre os valores estimados, isto é, 0 e 1 (ROGGO et al., 2007) na qual pode ser observada na Figura 34b e 32f. Esta assimetria pode ser explicada pela diferença de dispersão dos valores estimados para cada classe discriminada das outras classes.

Analisando a Figura 34, é observada que a estimativa dos valores da classe para treinamento (símbolos cheios) e conjunto de teste (símbolos vazios) para discriminação de classe pelo modelo PLS1-DA: (a) BES dos outros biodieseis, (b) BMPM dos outros biodieseis, (c) BMFU dos outros biodieseis, (d) BMS dos outros biodieseis, (e) BEFU dos outros biodieseis e (f) BEPM dos outros biodieseis (■) BES, (▲) BMPM, (▼) BMFU, (●) BMS, (◆) BEFU, (◄) BEPM e (---) *threshold* de discriminação.

### CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 34a – Estimativa de (■) BES.

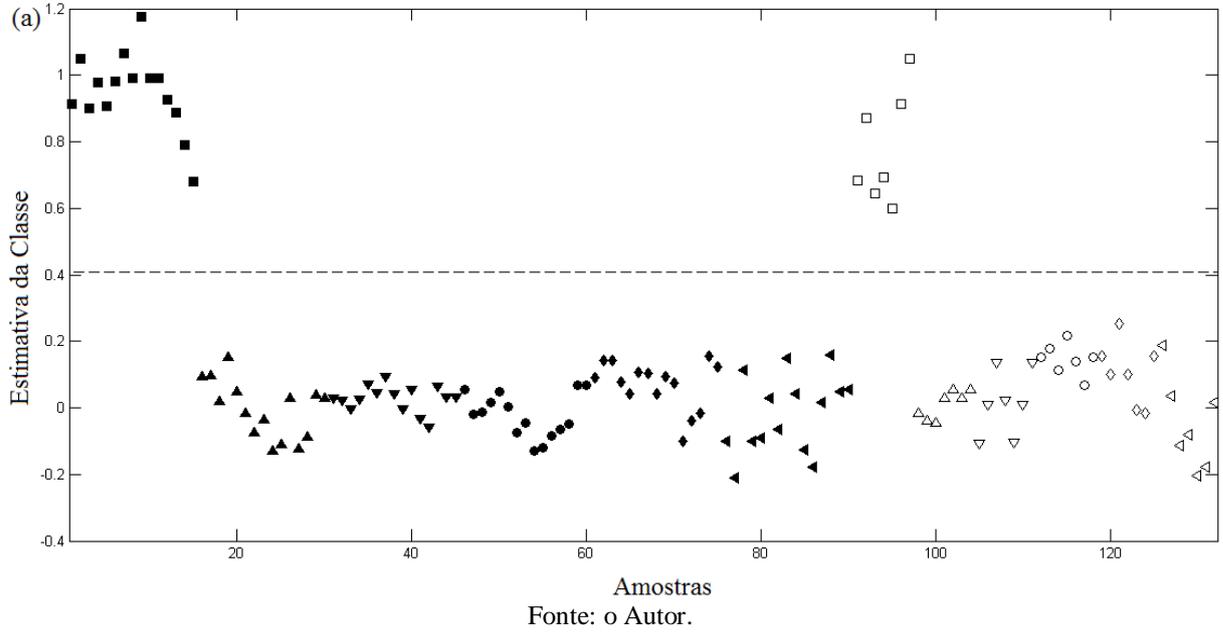
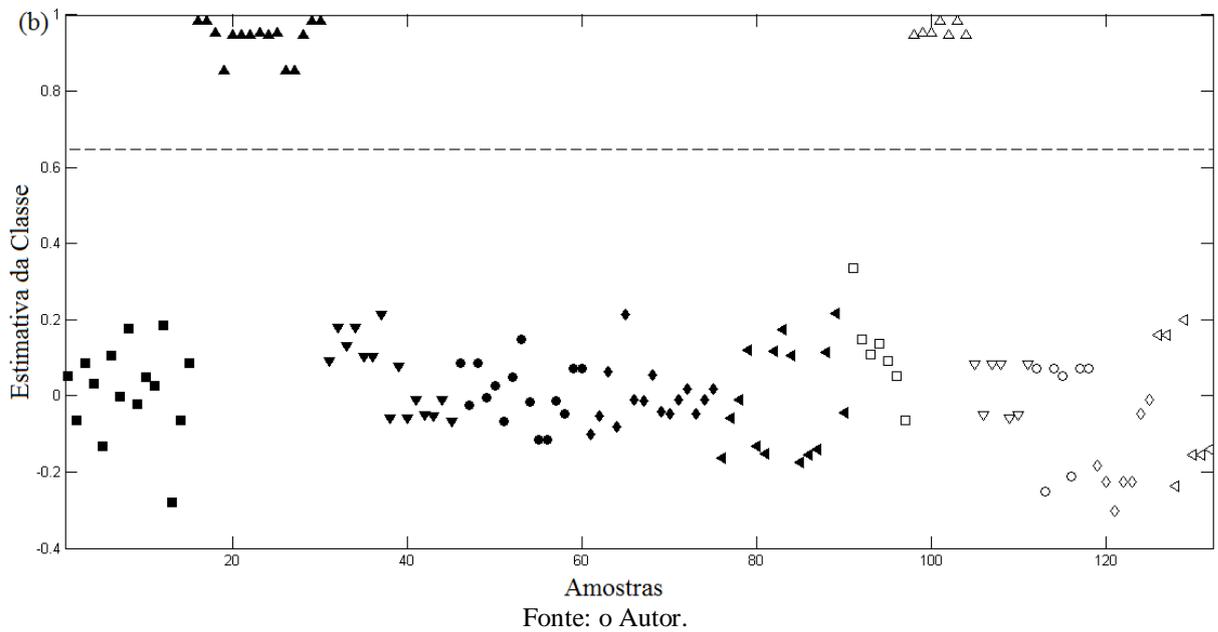


Figura 34b – Estimativa de (▲) BMPM.



## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 34c – Estimativa de (▼) BMFU.

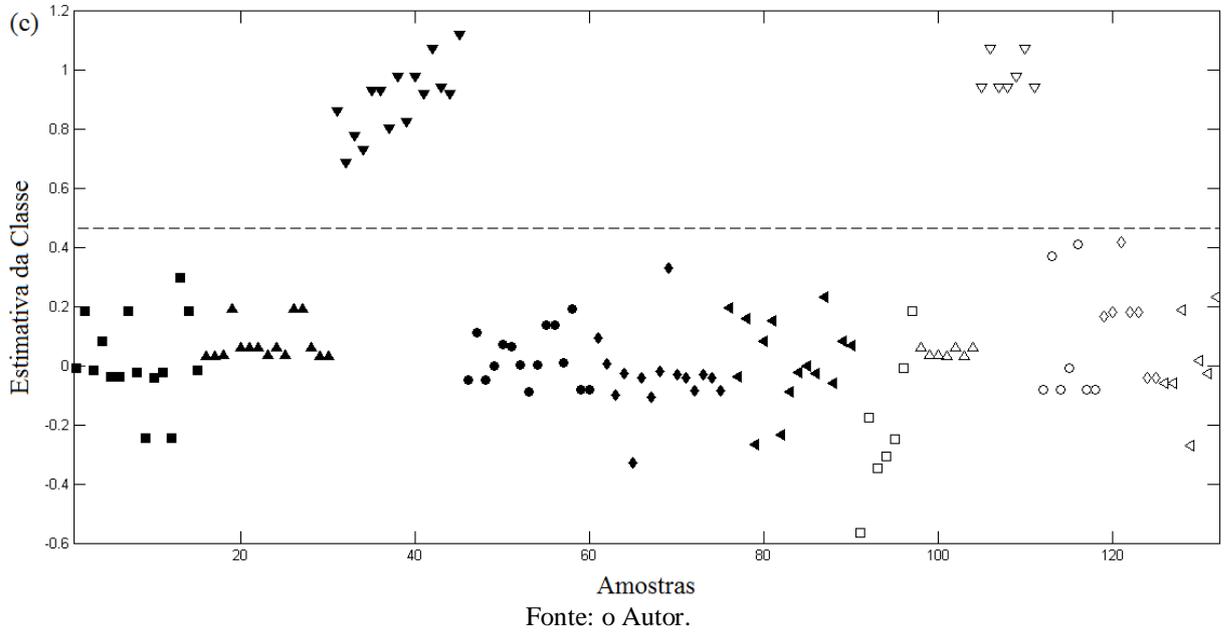
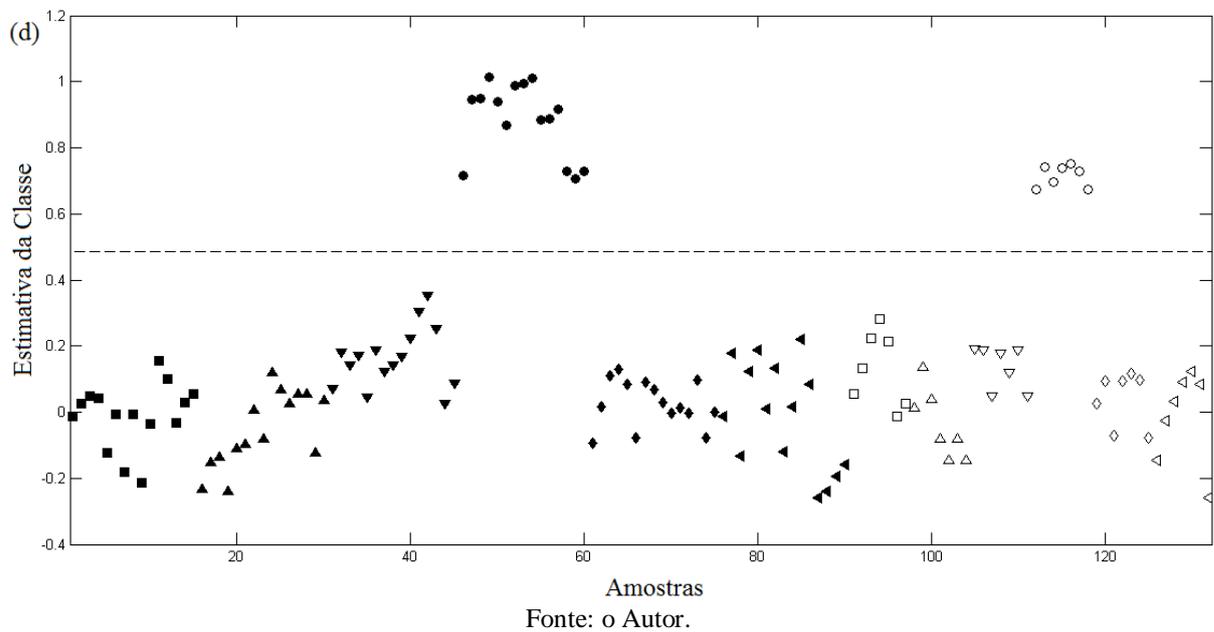


Figura 34d – Estimativa de (●) BMS.



## CAPÍTULO 5: CLASSIFICAÇÃO DE BIODIESEIS E SUAS MISTURAS COM DIESEL

Figura 34e – Estimativa de (♦) BEFU.

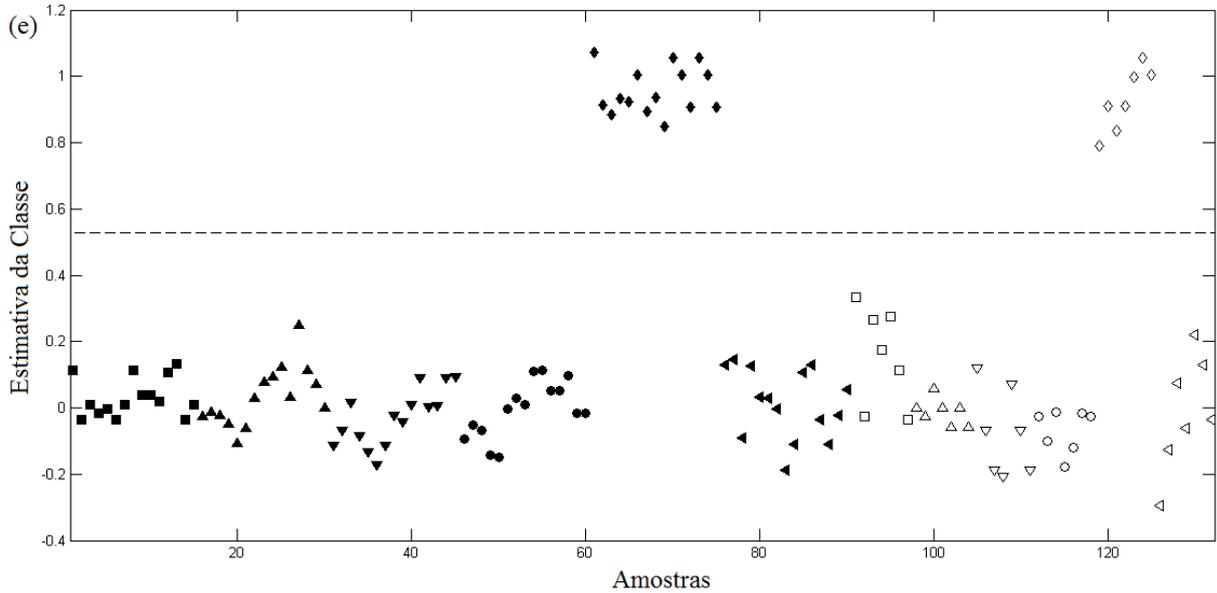
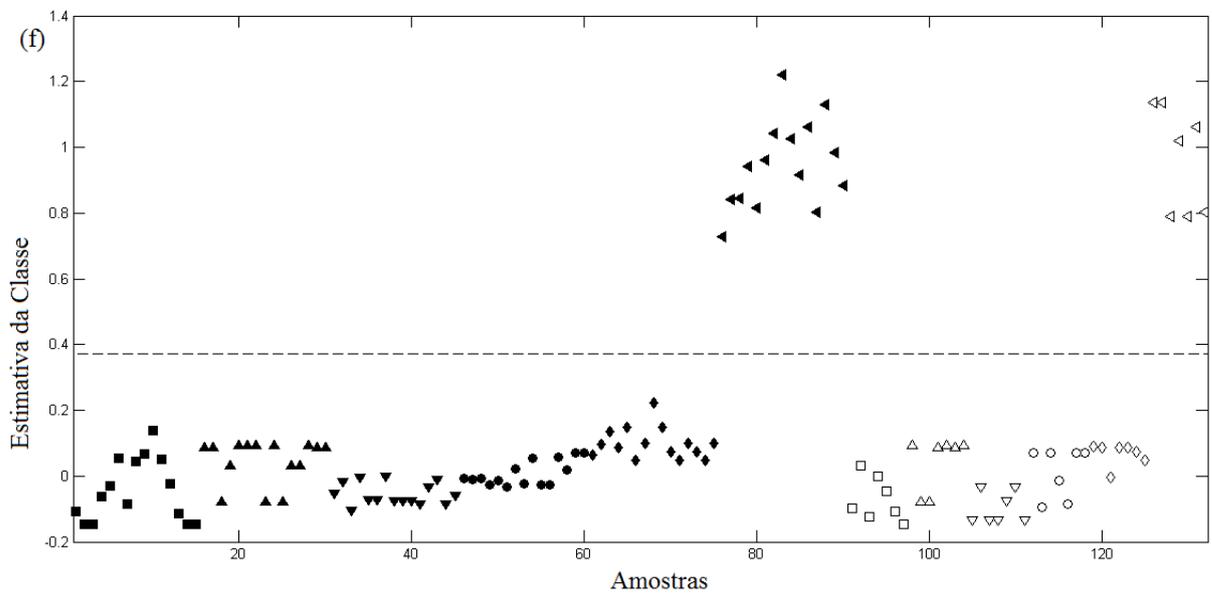


Figura 34f – Estimativa de (◀) BEPM.

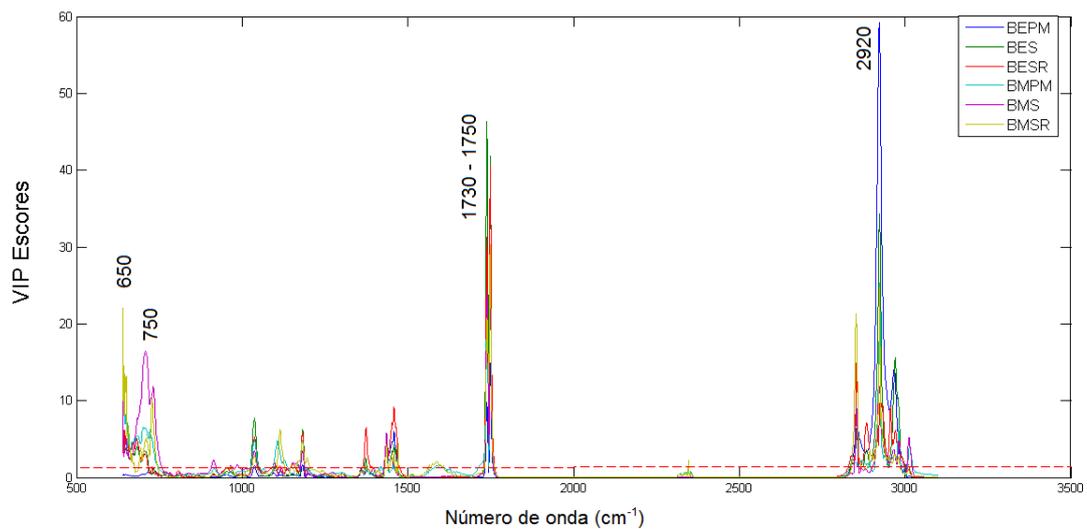


Os escores da VIP, apresentados na Figura 35, mostram as bandas MIR que contribuem na diferenciação entre os tipos de misturas B5. Valores da VIP maior que 1, indicam que a

## CAPÍTULO 5: CLASSIFICAÇÃO DE BODIESEIS E SUAS MISTURAS COM DIESEL

região espectral contribui na separação da classe das outras. Estas regiões podem ser distinguidas como importantes para ótimo desempenho do modelo PLS1-DA (ALMEIDA et al., 2013). Neste caso, as regiões que mais influenciaram na discriminação foram 1730-1750  $\text{cm}^{-1}$  correspondente a C=O e região de 2920  $\text{cm}^{-1}$  atribuída ao estiramento simétrico  $\text{CH}_3$ . No entanto, as regiões com número de onda 650  $\text{cm}^{-1}$  e 750  $\text{cm}^{-1}$  foram também responsáveis principalmente para discriminar biodieseis de BMFU e BMS.

Figura 35 - Variável Importante na Projeção (VIP) escores para modelo PLS1-DA



Fonte: o Autor.

## 6 MODELOS DE ADULTERAÇÕES EM MISTURAS B5

### 6.1 Objetivos Específicos

Este capítulo tem como objetivo:

- ✓ adulterar artificialmente 15 lotes diferentes de biodiesel (B5) na faixa de concentração de 1-20% para cada tipo (BES, BMPM, BMFU, BMS, BEFU e BEPM) usando gasolina, OLAR, soja e OFU;
- ✓ construir modelos PLS-DA para detectar cada tipo de adulterante (gasolina, OLAR, soja, OFU) em cada um dos tipos de B5 (BES, BMPM, BMFU, BMS, BEFU e BEPM) usando dados de espectroscopia MIR.
- ✓ validar os modelos PLS-DA desenvolvidos, através dos valores da sensibilidade e especificidade, para cada parâmetro nas amostras de conjunto teste na classe de interesse.

### 6.2 Metodologia

Foram utilizados 22 lotes diferentes de amostras puras de cada mistura B5 de BES, BMPM, BMFU, BMS, BEFU e BEPM. Destas amostras separaram-se alíquotas de 15 amostras para conjunto de treinamento e 7 para conjunto teste, de forma independente.

Em relação às adulterações artificiais, estas foram realizadas adicionando os adulterantes em cada uma das misturas B5, com concentrações entre 1,0 e 20,0% (m/m), em 15 lotes de amostras.

Para cada tipo de adulterante (gasolina, OLAR, óleo de soja e UFO), foram preparadas a partir da solução estoque, em que as 15 amostras foram separadas em dois grupos de forma independente, sendo 10 para o conjunto de treinamento e 5 para o conjunto teste.

Assim, em cada modelo PLS1-DA, para detectar adulterantes foram usadas no total 37 amostras, das quais 22 amostras puras (não adulteradas) e 15 amostras adulteradas.

Os espectros foram obtidos usando um espectrômetro FT-IR SpectrumTwo (Perkin Elmer) utilizando um acessório horizontal de reflectância total atenuada (HATR) de Seleneto de Zinco (ZnSe) com as seguintes configurações: região de 4000 a 600  $\text{cm}^{-1}$ , resolução de 4  $\text{cm}^{-1}$  e 16 varreduras por espectro. Os espectros MIR, foram obtidos em quintuplicata para cada amostra, para a construção dos modelos PLS-DA, para detectar adulterantes.

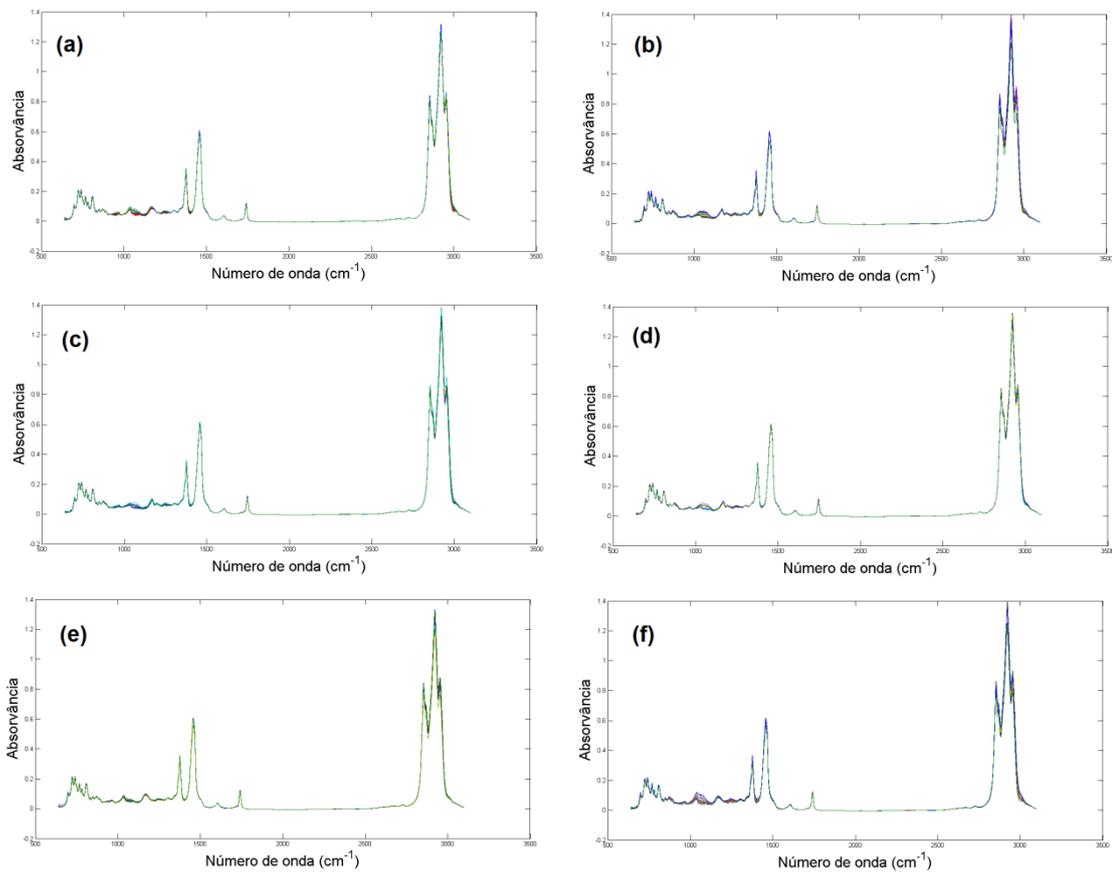
Os espectros obtidos tiveram suas linhas de base corrigidas utilizando o método de baseline nas regiões espectrais de 1850 a 2500 e de 3150 a 4000  $\text{cm}^{-1}$  e foram analisados com PLS\_Toolbox 7.5 (Eigenvector Research) e em ambiente MATLAB 7.5 (MathWorks Inc.). Nos modelos construídos, o conjunto de dados foi dividido em um grupo para treinamento e outro para teste. Os grupos de amostras foram centrados na média e empregou-se validação cruzada por *leave one out*.

### 6.3 Modelos PLS1-DA para detectar adulteração por gasolina em misturas B5

Nos espectros das misturas de B5 de BES, BMPM, BMFU, BMS, BEFU e BEPM, adulteradas com gasolina, verifica-se que há uma proeminente banda em 2920  $\text{cm}^{-1}$  que é atribuída ao estiramento simétrico  $\text{CH}_3$ ; 2850  $\text{cm}^{-1}$  e 2950  $\text{cm}^{-1}$  associados ao estiramento simétrico e assimétrico de  $\text{CH}_2$ , respectivamente (MIRANDA et al., 2014 ); 1750-1730  $\text{cm}^{-1}$  corresponde às vibrações de  $\text{C}=\text{O}$  (GAYDOU; DUPUY, 2011); 1600  $\text{cm}^{-1}$  é atribuída a  $\text{C}=\text{C}$ ; 1450 e 1380  $\text{cm}^{-1}$  deformação assimétrica e simétrica de  $\text{CH}_3$  respectivamente (MIRANDA et al., 2014 ) (Figura 36).

Finalmente a vibração observada próxima de 1100 até 1350  $\text{cm}^{-1}$  provavelmente seja do estiramento  $\text{C}=\text{O}(\text{OR})$  e na região de 700-900  $\text{cm}^{-1}$  é atribuída a vibração do estiramento  $\text{CH}$  (GAYDOU; DUPUY, 2011). Os espectros MIR adulterados com gasolina apresentam uma sobreposição na região de 850-1300  $\text{cm}^{-1}$  (Figura 36).

Figura 36 – Espectros de amostras das misturas de B5 puras e adulteradas com gasolina: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM.



Fonte: o Autor.

Os seis modelos PLS1-DA desenvolvidos para classificar amostras puras (não adulteradas) das adulteradas com gasolina, os parâmetros obtidos estão apresentados na Tabela 6.

**Tabela 6** - Parâmetros de classificação obtidos por PLS1-DA para detectar amostras puras das adulteradas por gasolina.

Parâmetros	Tipos de B5 adulteradas por gasolina					
	BES	BMPM	BMFU	BMS	BEFU	BEPM
Variáveis Latentes	3	5	4	4	8	5
RMSEC	0,1115	0,0349	0,0723	0,0664	0,0443	0,0283
RMSECV	0,1415	0,0668	0,0990	0,0944	0,1099	0,0368
RMSEP	0,1716	0,0736	0,1201	0,1391	0,1231	0,0544
<b>Sensibilidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b>Especificidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
Valor de <i>Threshold</i>	0,6300	0,4107	0,4260	0,5900	0,5774	0,3933

Fonte: o Autor.

Verifica-se que os modelos PLS1-DA desenvolvidos são eficientes na detecção de adulterantes empregando dados de espectroscopia MIR. A Figura 37 apresenta os valores de classe estimada para o conjunto de treinamento e de teste. Amostras que se encontram acima do valor de *threshold* são classificadas como pertencente à mistura de B5 pura e abaixo do valor de *threshold* como amostras da mistura de B5 adulterada com gasolina.

Os modelos PLS1-DA desenvolvidos mostram que os perfis dos espectros MIR possuem variabilidade de uma classe para outra, neste caso, dos diferentes tipos de biodieseis de acordo com o tipo de óleo e álcool usado uma vez que todas as misturas B5 foram adulteradas com gasolina. Portanto, esta variabilidade entre os perfis dos espectros MIR faz com que o valor estimado de *threshold* pelo modelo PLS-DA seja diferente para cada tipo de biodiesel (PASTORE et al., 2011).

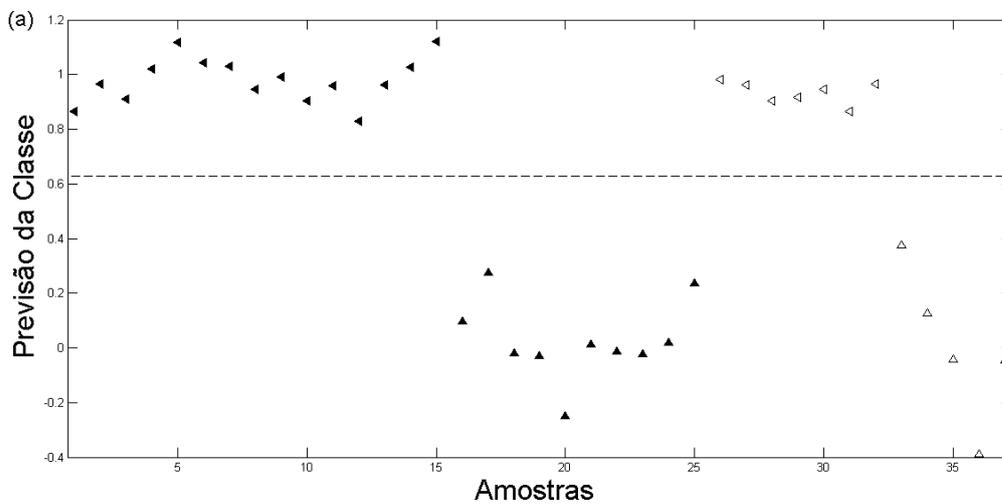
Nos seis modelos construídos, observa-se que todas as amostras do conjunto teste foram classificadas como pertencentes as respectivas classes (Figura 37).

Analisando a Figura 37, é observada que a estimativa dos valores da classe para treinamento (símbolos cheios) e conjunto de teste (símbolos vazios) para discriminação entre amostras adulteradas e puras de B5 pelo modelo PLS1-DA: (a) amostras de BES não adulteradas das adulteradas com gasolina, (b) amostras de BMPM não adulteradas das adulteradas com gasolina, (c) amostras de BMFU não adulteradas das adulteradas com

## CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

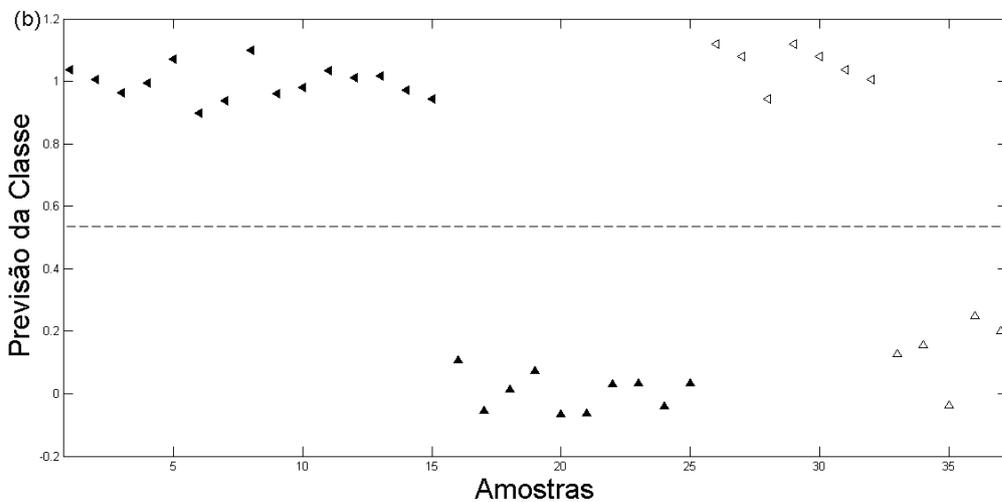
gasolina, (d) amostras de BMS não adulteradas das adulteradas com gasolina, (e) amostras de BEFU não adulteradas das adulteradas com gasolina (f) amostras de BMS não adulteradas das adulteradas com gasolina, onde ( $\blacktriangleleft$ ) amostras não adulteradas de BES, BEFU, BEPM, ( $\blacksquare$ ) amostras não adulteradas de BMPM, BMFU, BMS, ( $\blacktriangle$ ) Misturas de B5 adulteradas por gasolina (- - -) *threshold* de discriminação.

Figura 37a – Estimativa de ( $\blacktriangleleft$ ) amostras não adulteradas de BES das ( $\blacktriangle$ ) misturas de B5 adulteradas por gasolina.



Fonte: o Autor.

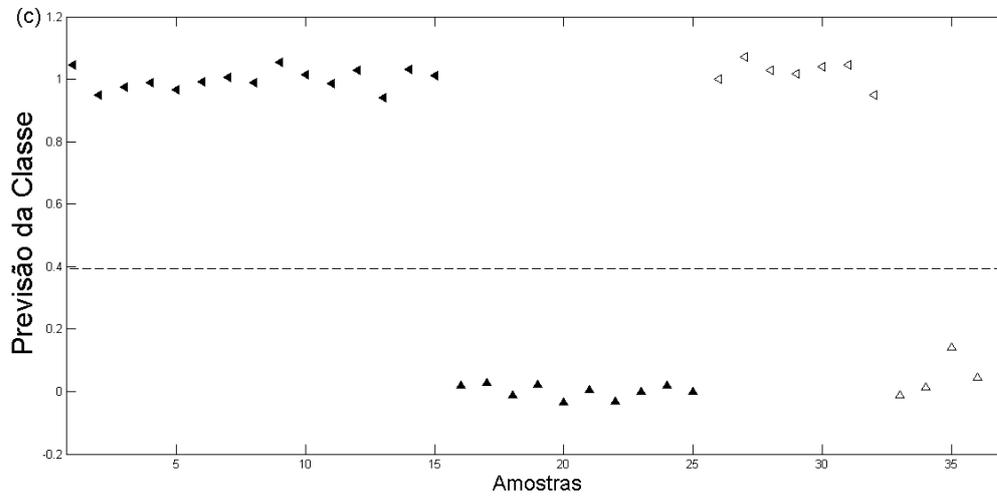
Figura 37b – Estimativa de ( $\blacktriangleleft$ ) amostras não adulteradas de BEFU das ( $\blacktriangle$ ) misturas de B5 adulteradas por gasolina.



Fonte: o Autor.

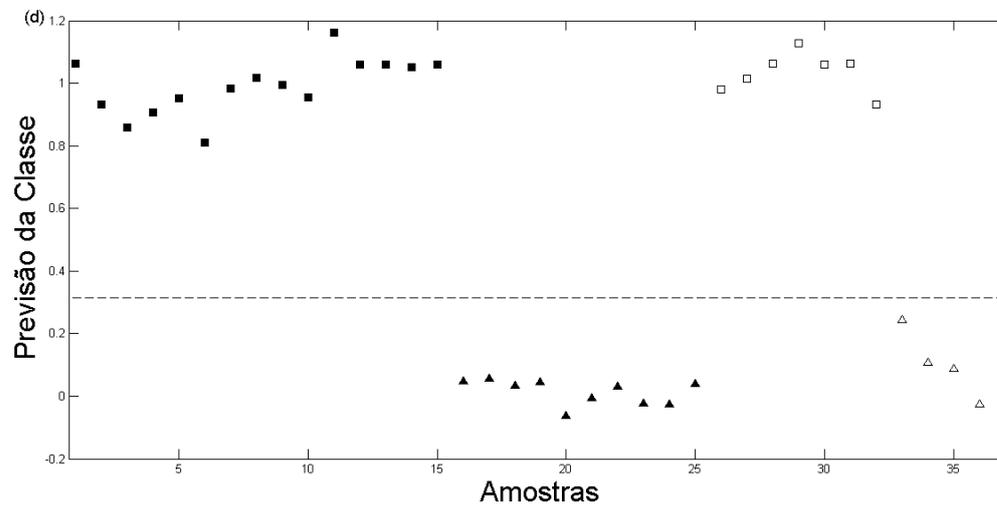
## CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

Figura 37c – Estimativa de (◄) amostras não adulteradas de BEPM das (▲) misturas de B5 adulteradas por gasolina.



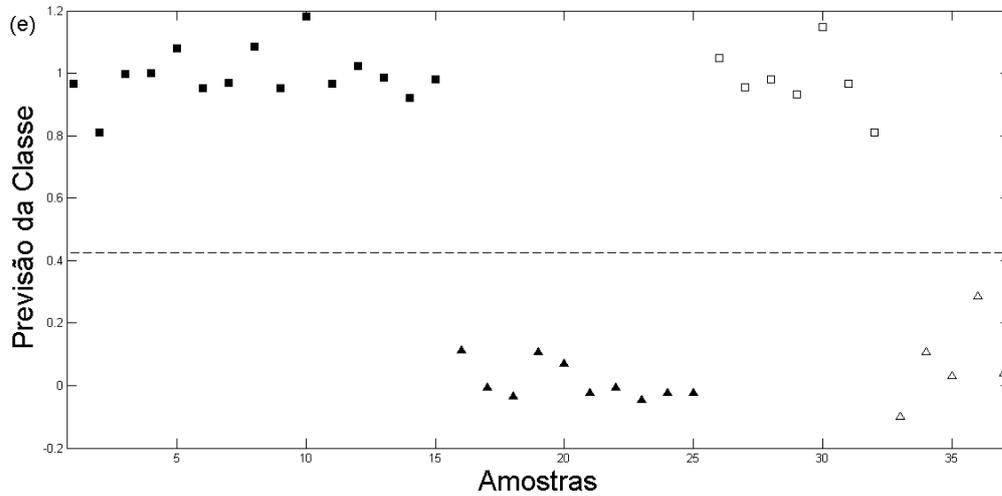
Fonte: o Autor.

Figura 37d – Estimativa de (■) amostras não adulteradas de BMPM das (▲) misturas de B5 adulteradas por gasolina.



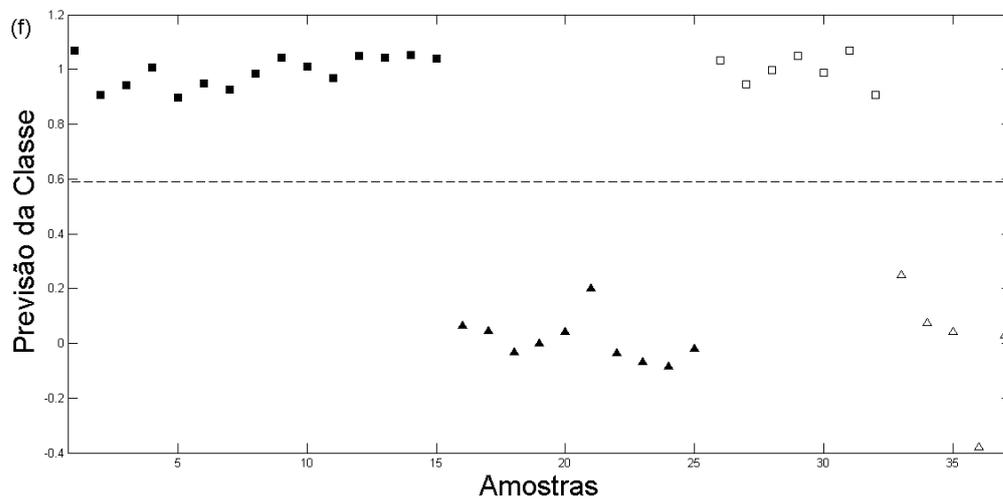
Fonte: o Autor.

Figura 37e – Estimativa de (■) amostras não adulteradas de BMFU das (▲) misturas de B5 adulteradas por gasolina.



Fonte: o Autor.

Figura 37f – Estimativa de (■) amostras não adulteradas de BMS das (▲) misturas de B5 adulteradas por gasolina.

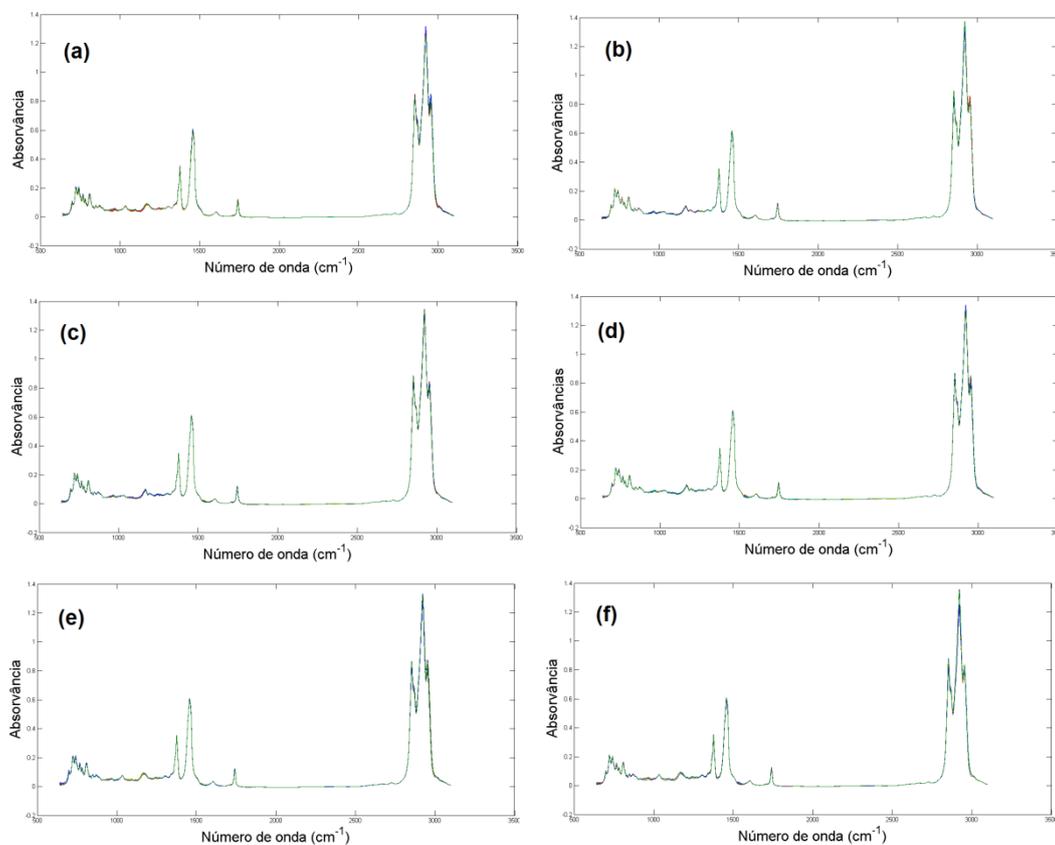


Fonte: o Autor.

#### 6.4 Modelos PLS1-DA para detectar adulteração por OLAR em misturas B5

Os espectros das misturas B5 de BES, BMPM, BMFU, BMS, BEFU e BEPM de amostras puras e adulteradas por OLAR, são muito semelhantes, uma vez que não se nota diferenças nos espectros MIR, através uma simples inspeção pela visão humana (Figura 38) (SOUZA et al., 2014).

Figura 38 – Espectros de amostras das misturas de B5 puras e adulteradas por OLAR: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM.



Fonte: o Autor.

Os modelos PLS1-DA foram obtidos a partir do conjunto amostral constituído por misturas B5 puras e adulteradas com OLAR.

Os resultados obtidos para cada uma das duas classes (amostras adulteradas e amostras não adulteradas) de misturas B5 de BES, BMPM, BMFU, BMS, BEFU e BEPM, estão representados na Figura 39.

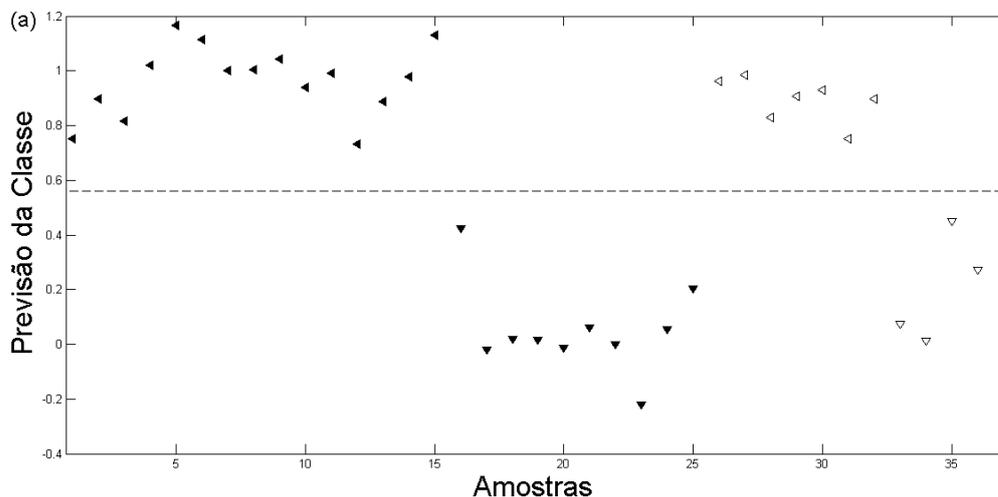
---

 CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

A linha tracejada na Figura 39 indica o valor de *threshold* entre os valores previstos. Valores acima do valor de *threshold* indica que a amostra pertence à classe modelada e valores previstos abaixo de *threshold* indica que a amostra não pertence à classe modelada. Esta linha é calculada pelo modelo PLS-DA e é estimada a partir das amostras do conjunto de treinamento e os valores do conjunto teste pelo algoritmo PLS.

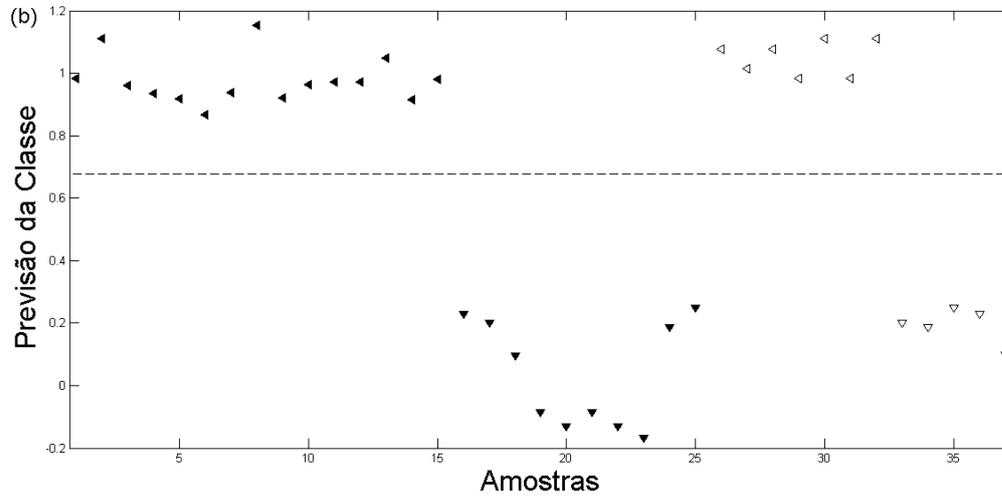
Analisando a Figura 39, é observada que a estimativa dos valores da classe para treinamento (símbolos cheios) e conjunto de teste (símbolos vazios) para discriminação entre amostras adulteradas e puras de misturas B5 pelo modelo PLS1-DA: (a) amostras de BES não adulteradas das adulteradas com OLAR, (b) amostras de BMPM não adulteradas das adulteradas com OLAR, (c) amostras de BMFU não adulteradas das adulteradas com OLAR, (d) amostras de BMS não adulteradas das adulteradas com OLAR, (e) amostras de BEFU não adulteradas das adulteradas com OLAR (f) amostras de BMS não adulteradas das adulteradas com OLAR, onde (◀) amostras não adulteradas de BES, BEFU, BEPM, (■) amostras não adulteradas de BMPM, BMFU, BMS, (▼) misturas de B5 adulteradas por OLAR ( - - - ) *threshold* de discriminação.

Figura 39a – Estimativa de (◀) amostras não adulteradas de BES das (▼) misturas de B5 adulteradas por OLAR.



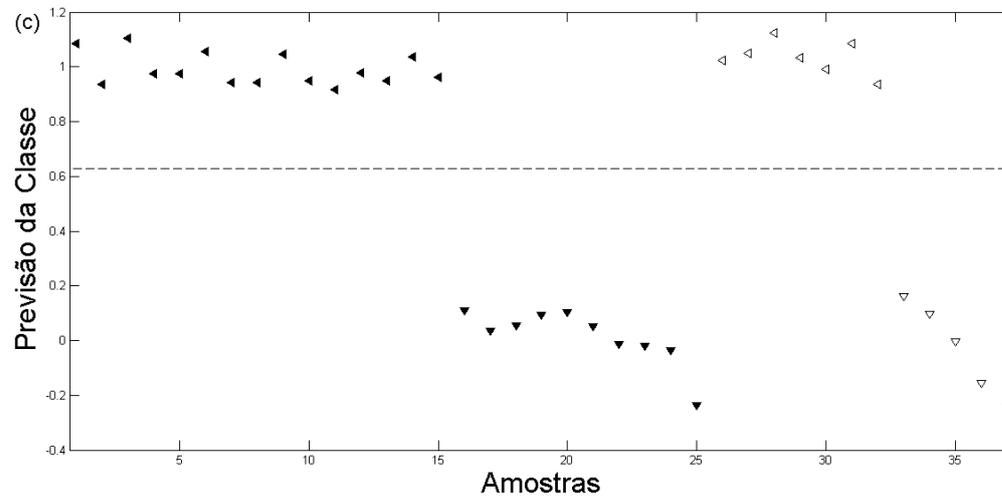
Fonte: o Autor.

Figura 39b – Estimativa de (◀) amostras não adulteradas de BEFU das (▼) misturas de B5 adulteradas por OLAR.



Fonte: o Autor.

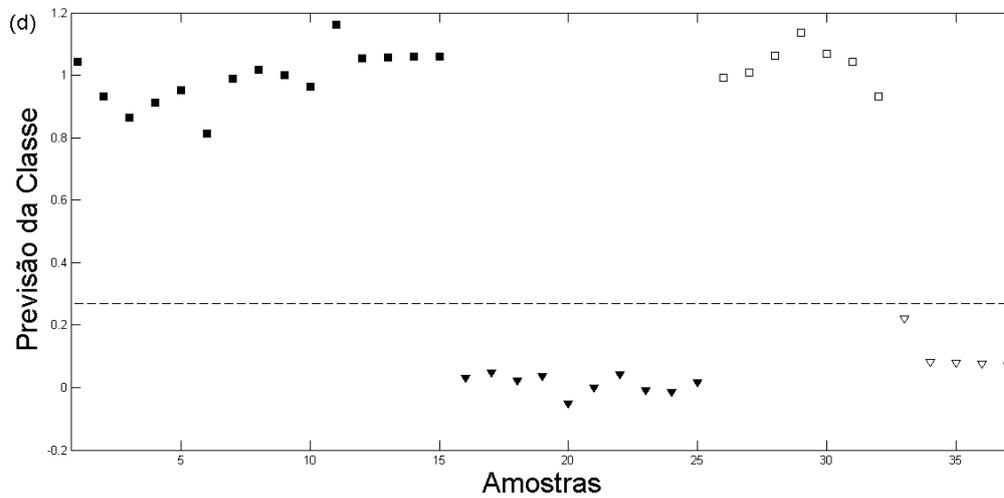
Figura 39c – Estimativa de (◀) amostras não adulteradas de BEPM das (▼) misturas de B5 adulteradas por OLAR.



Fonte: o Autor.

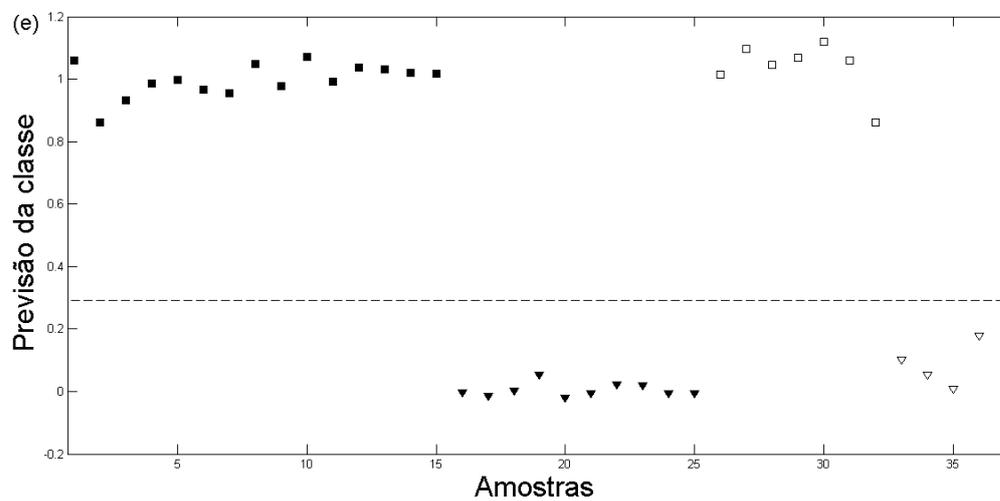
## CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

Figura 39d – Estimativa de (■) amostras não adulteradas de BMPM das (▼) misturas de B5 adulteradas por OLAR.



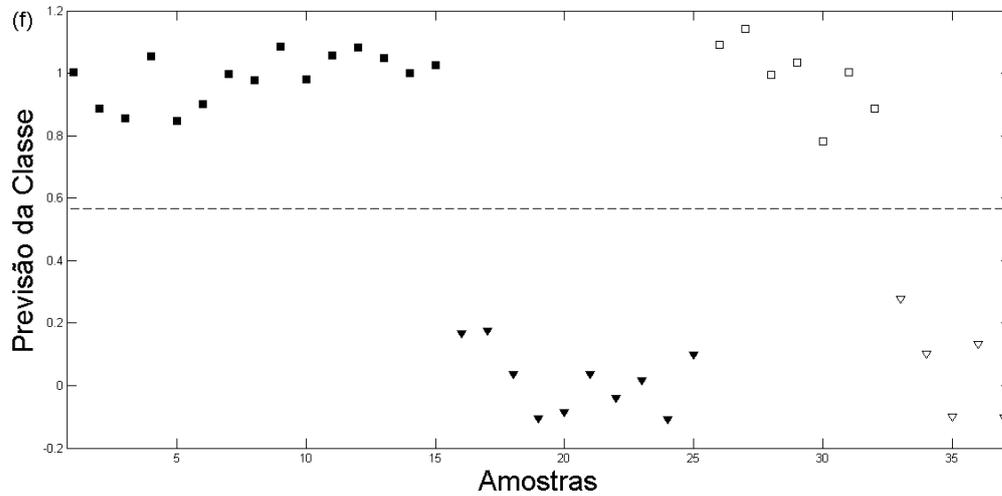
Fonte: o Autor.

Figura 39e – Estimativa de (■) amostras não adulteradas de BMFU das (▼) misturas de B5 adulteradas por OLAR.



Fonte: o Autor.

Figura 39f – Estimativa de (■) Amostras não adulteradas de BMS das (▼) Misturas de B5 adulteradas por OLAR.



Fonte: o Autor.

A Figura 39 indica que ambas as classes (amostras adulteradas e amostras não adulteradas) foram classificadas corretamente pelos seis modelos PLS1-DA, tanto as amostras do conjunto de treinamento quanto do conjunto de teste. Os parâmetros de sensibilidade e especificidade obtidas pelo modelo PLS-DA estão apresentados na Tabela 7. Os resultados indicam 100% de classificação correta nas amostras de teste de amostras adulteradas e não adulteradas.

Estes resultados confirmam que o uso da Espectroscopia MIR associado ao PLS-DA foram capazes de discriminar amostras de B5 de BES, BMPM, BMFU, BMS, BEFU e BEPM, adulteradas e não adulteradas com 100% de classificação correta tanto no treinamento quanto no teste.

## CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

**Tabela 7** - Parâmetros de classificação obtidos por PLS1-DA para detectar amostras puras das adulteradas por OLAR

Parâmetros	Tipos de B5 adulteradas por OLAR					
	BES	BMPM	BMFU	BMS	BEFU	BEPM
Variáveis Latentes	2	4	6	3	2	2
RMSEC	0,1458	0,0697	0,0431	0,0881	0,1205	0,0767
RMSECV	0,1715	0,0931	0,0828	0,1253	0,1276	0,0939
RMSEP	0,1917	0,0939	0,0918	0,1341	0,1400	0,1101
<b>Sensibilidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b>Especificidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
Valor de <i>Threshold</i>	0,5600	0,2684	0,2909	0,5680	0,6764	0,6272

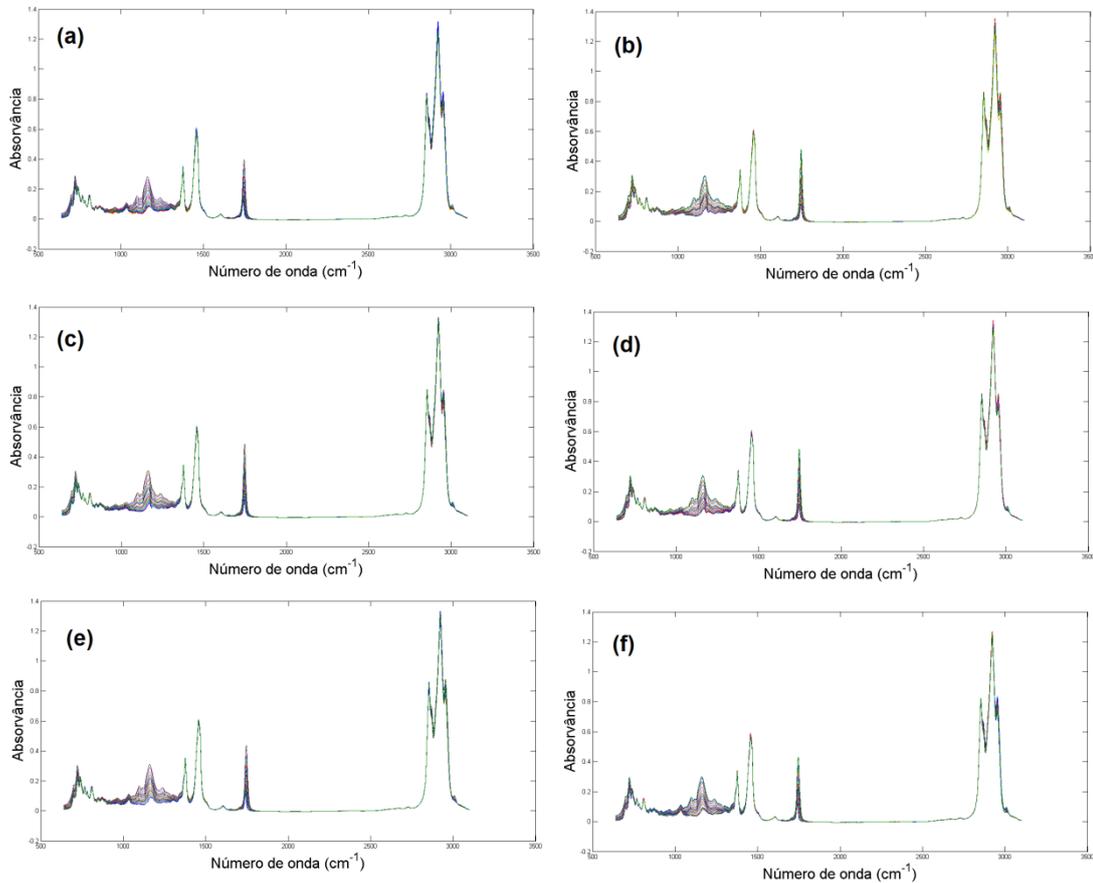
Fonte: o Autor.

A partir dos resultados obtidos pelos modelos, foi possível determinar as amostras puras das adulteradas, após a classificação de uma amostra desconhecida utilizando cada um dos seis modelos PLS1-DA. Ao classificar amostras desconhecidas, o modelo estará a reconhecer padrões dado que a amostra não faz parte do conjunto treinamento, porém, importa referir que o modelo reconhece as amostras desconhecidas uma vez que o perfil delas está definido no modelo através de amostras do conjunto de treinamento.

### 6.5 Modelos PLS1-DA para detectar adulteração por óleo de soja em misturas B5

Os espectros das misturas B5 puras e adulteradas com óleo de soja apresentam uma sobreposição que diferenciam dos adulterantes, apresentados anteriormente. A referida sobreposição é observada na região entre 1720-1780  $\text{cm}^{-1}$ , correspondente às vibrações de C=O, presente na estrutura química do óleo de soja (Figura 40) (DA COSTA FILHO, 2014).

Figura 40 – Espectros de amostras das misturas de B5 puras e adulteradas por óleo de soja: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM.



Fonte: o Autor.

A análise por PLS-DA foi realizada com intuito de distinguir os espectros MIR entre amostras não adulteradas e adulteradas com óleo de soja, uma vez que a análise qualitativa dos espectros não permite classificar amostras adulteradas das não adulteradas.

No entanto, os seis modelos desenvolvidos não apresentaram erro de classificação, em detrimento de ter apresentados 100% de classificação correta (Tabela 8).

## CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

**Tabela 8** - Parâmetros de classificação obtidos por PLS1-DA para detectar amostras puras das adulteradas por óleo de soja.

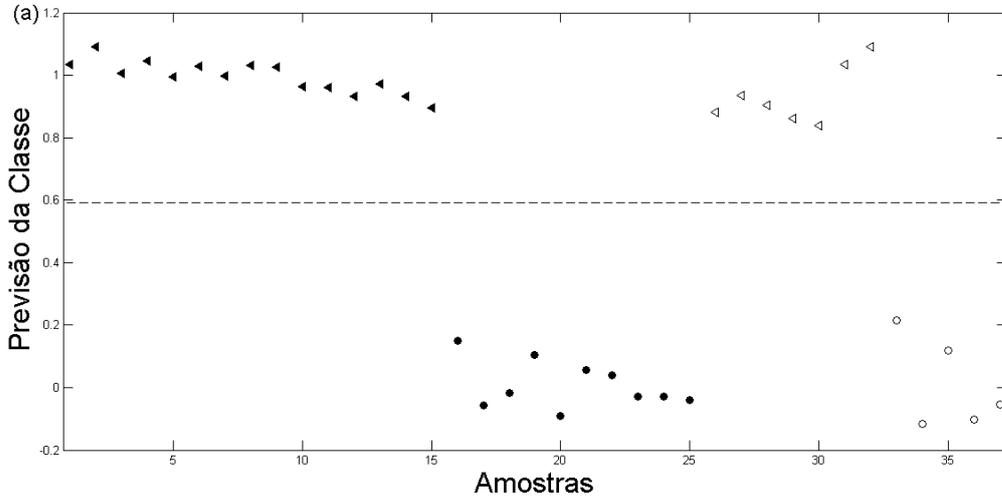
Parâmetros	Tipos de B5 adulteradas por óleo de soja					
	BES	BMPM	BMFU	BMS	BEFU	BPEM
Variáveis Latentes	6	6	3	4	3	4
RMSEC	0,0602	0,0293	0,0898	0,0542	0,1028	0,0570
RMSECV	0,1181	0,0685	0,1234	0,0719	0,1277	0,0918
RMSEP	0,1189	0,1026	0,1297	0,0882	0,1614	0,0944
<b>Sensibilidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b>Especificidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
Valor de <i>Threshold</i>	0,5918	0,5563	0,6687	0,4060	0,4747	0,5127

Fonte: o Autor.

As informações fornecidas pelos espectros MIR juntamente com o PLS-DA, auxilia na identificação de misturas de B5 adulteradas (amostras não conformes) (Figura 41).

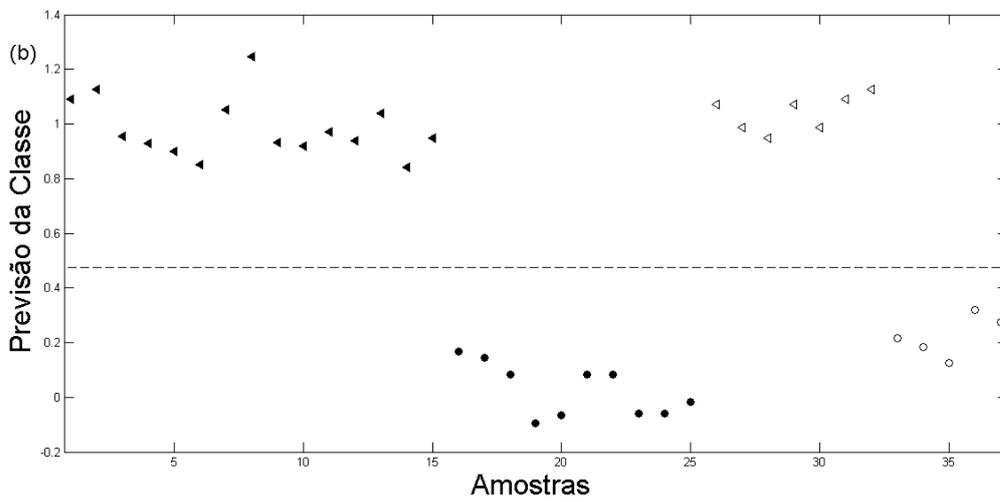
Analisando a Figura 41, é observada que a estimativa dos valores da classe para treinamento (símbolos cheios) e conjunto de teste (símbolos vazios) para discriminação entre amostras adulteradas e puras de misturas B5 pelo modelo PLS1-DA: (a) amostras de BES não adulteradas das adulteradas com óleo de soja, (b) amostras de BMPM não adulteradas das adulteradas com óleo de soja, (c) amostras de BMFU não adulteradas das adulteradas com óleo de soja, (d) amostras de BMS não adulteradas das adulteradas com óleo de soja, (e) amostras de BEFU não adulteradas das adulteradas com óleo de soja (f) amostras de BMS não adulteradas das adulteradas com óleo de soja, onde (◀) amostras não adulteradas de BES, BEFU, BEPM, (■) amostras não adulteradas de BMPM, BMFU, BMS, (●) misturas de B5 adulteradas por óleo de soja (- - -) *threshold* de discriminação.

Figura 41a – Estimativa de (◄) amostras não adulteradas de BES das (●) misturas de B5 adulteradas por óleo de soja.



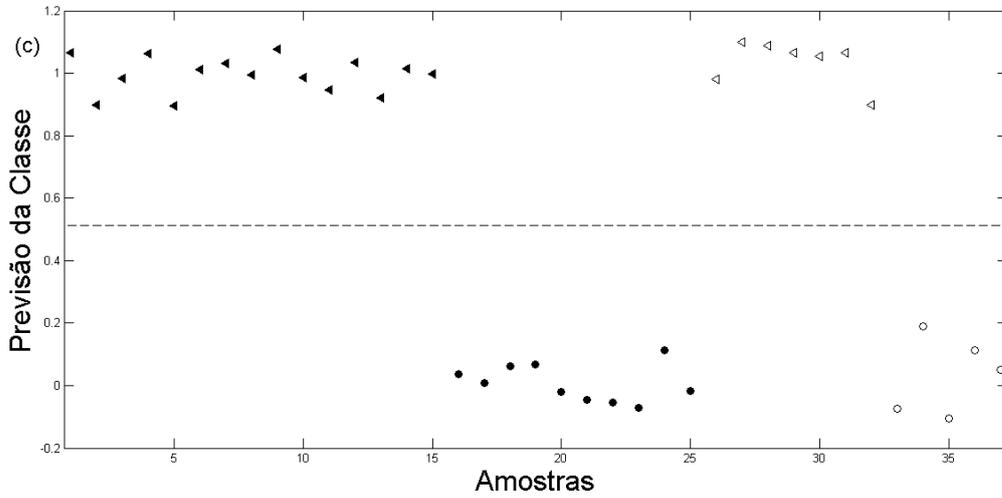
Fonte: o Autor.

Figura 41b – Estimativa de (◄) amostras não adulteradas de BEFU das (●) misturas de B5 adulteradas por óleo de soja.



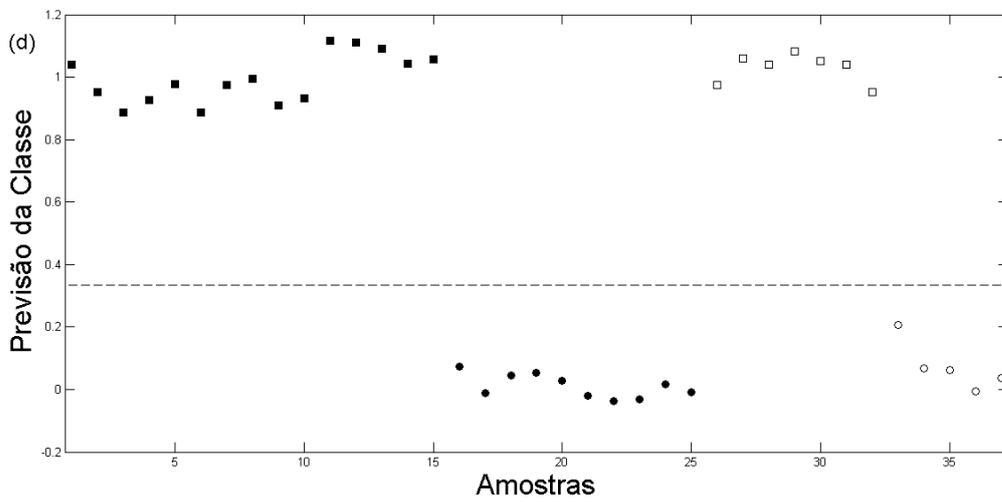
Fonte: o Autor.

Figura 41c – Estimativa de (◄) amostras não adulteradas de BEPM das (●) misturas de B5 adulteradas por óleo de soja.



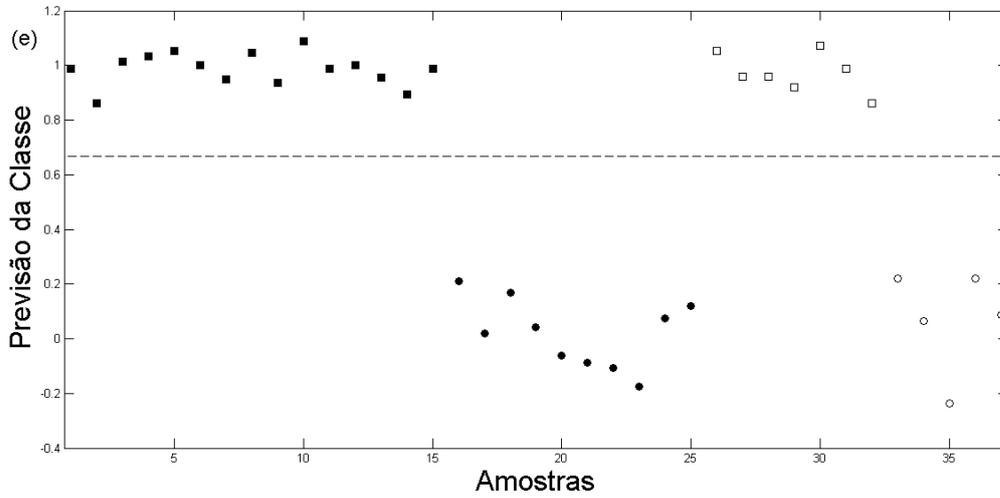
Fonte: o Autor.

Figura 41d – Estimativa de (■) amostras não adulteradas de BMPM das (●) misturas de B5 adulteradas por óleo de soja.



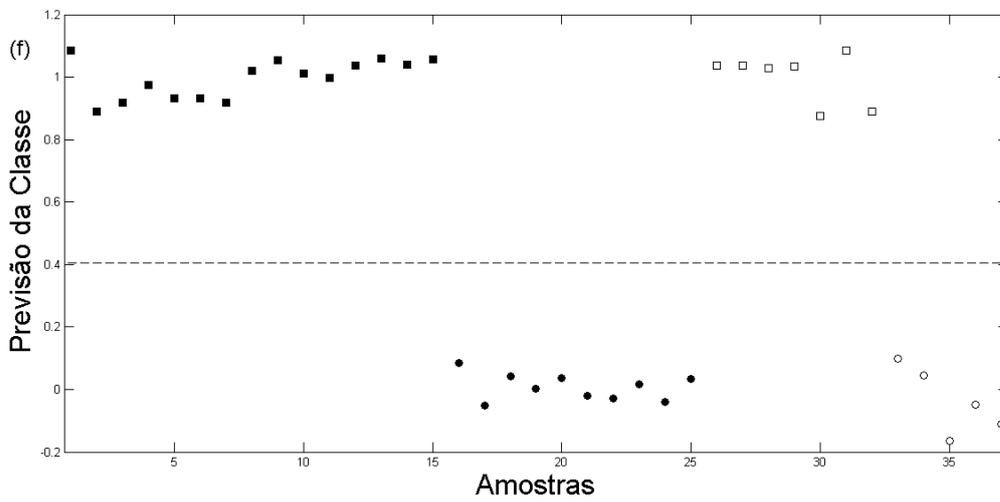
Fonte: o Autor.

Figura 41e – Estimativa de (■) amostras não adulteradas de BMFU das (●) misturas de B5 adulteradas por óleo de soja.



Fonte: o Autor.

Figura 41f – Estimativa de (■) amostras não adulteradas de BMS das (●) misturas de B5 adulteradas por óleo de soja.

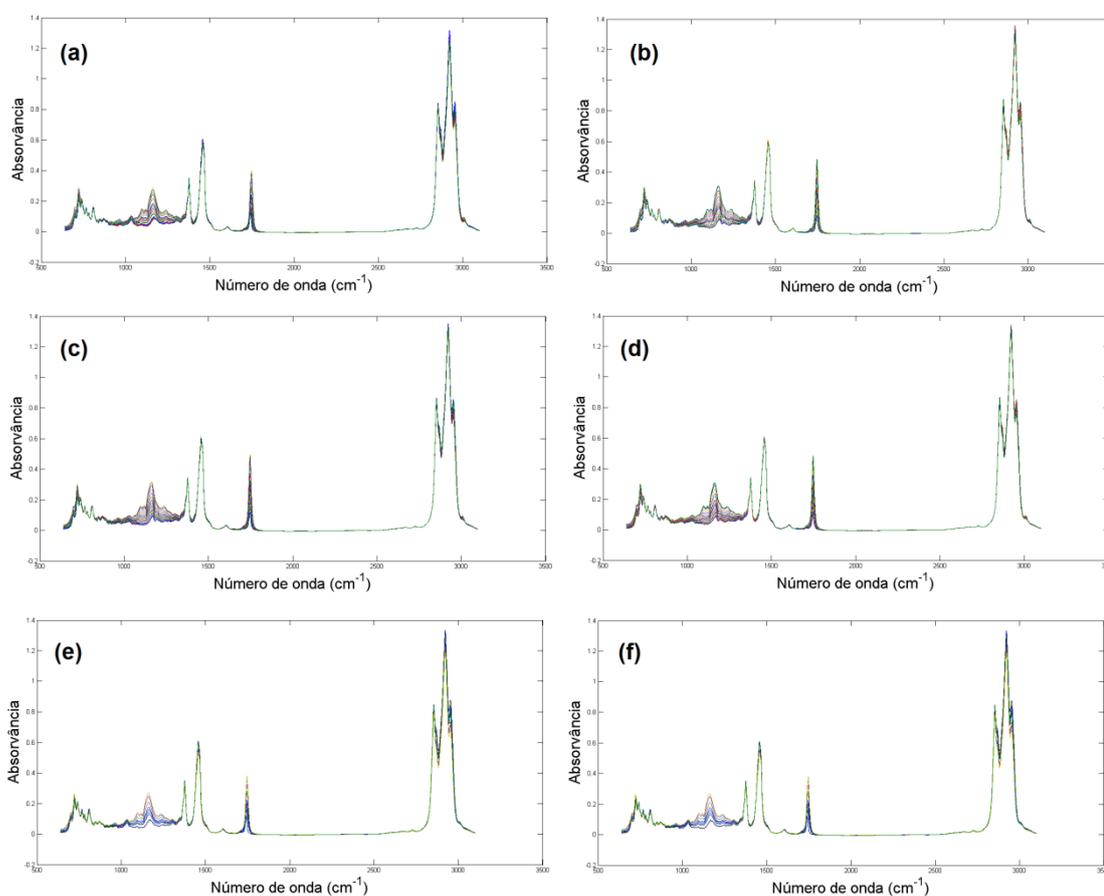


Fonte: o Autor.

### 6.6 Modelos PLS1-DA para detectar adulteração por óleo de fritura usado em misturas B5

Os espectros das misturas B5 puras e adulteradas com OFU estão apresentados na Figura 42. Porém, o perfil espectral é muito idêntico do comportamento apresentado pelos espectros de B5 adulterados com óleo de soja. Em detrimento disso, pode se afirmar que o OFU é de fato óleo de soja usado, isto é, degradado (BRANDAO; BRAGA; SUAREZ, 2012).

Figura 42 - Misturas de B5 não adulteradas e adulteradas por OFU: (a) BES, (b) BMPM, (c) BMFU, (d) BMS, (e) BEFU e (f) BEPM.



Fonte: o Autor.

Portanto, os espectros MIR, apresentam semelhanças e conseqüentemente não é possível através de simples inspeção visual a diferenciação de amostras adulteradas e não adulteradas, tornando-se necessário a utilização de métodos quimiométricos.

Os resultados dos seis modelos de análise multivariada construído, baseado no método PLS-DA e nos espectros MIR das amostras não adulteradas e adulteradas com OFU, não

## CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

apresentaram erro de classificação, tendo apresentado valores de sensibilidade e especificidade igual 1, corresponde a 100% de classificação correta nas amostras do conjunto teste (Tabela 9).

**Tabela 9** - Parâmetros de classificação obtidos por PLS1-DA para detectar amostras não adulteradas das adulteradas por OFU.

Parâmetros	Tipos de B5 adulteradas por OFU					
	BES	BMPM	BMFU	BMS	BEFU	BEPM
Variáveis Latentes	3	5	3	3	4	5
RMSEC	0,1153	0,0264	0,0603	0,0626	0,1089	0,0417
RMSECV	0,1430	0,0494	0,0852	0,0744	0,2101	0,0517
RMSEP	0,1531	0,0521	0,1106	0,1107	0,2190	0,0914
<b>Sensibilidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b>Especificidade</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
Valor de <i>Threshold</i>	0,6048	0,4354	0,4257	0,4229	0,6431	0,5000

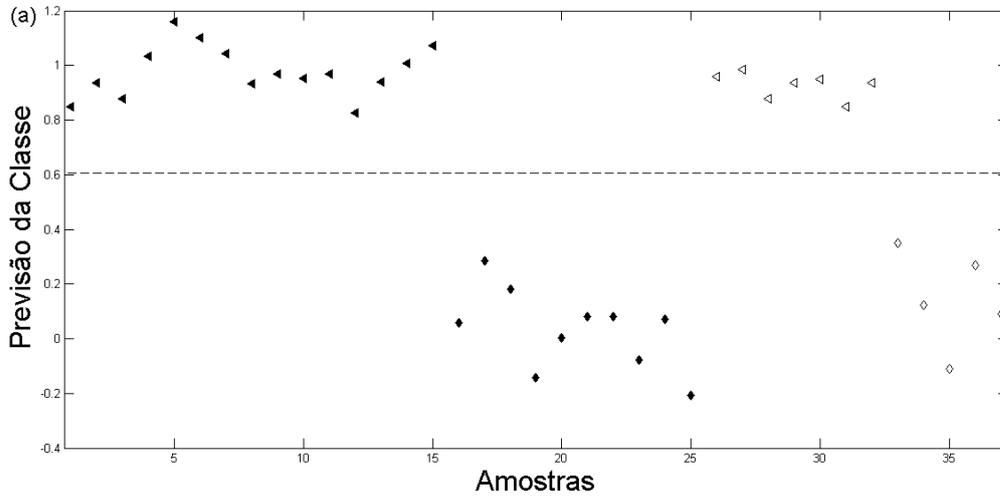
Fonte: o Autor.

A análise dos modelos PLS1-DA utilizados para discriminação por classes das amostras de B5 não adulteradas (amostras conformes) e adulteradas com OFU (amostras não conformes), mostram que as amostras foram classificadas nas respectivas classes.

Na Figura 43 são mostradas as classes previstas para as amostras não adulteradas das adulteradas, alto nível de desempenho, isto é, com sensibilidade e especificidade igual a 1.

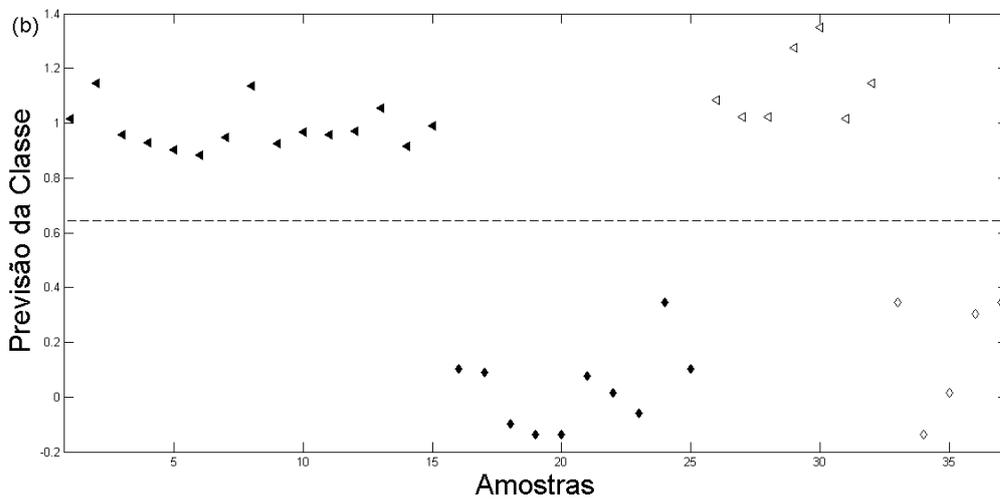
Analisando a Figura 43, é observada que a estimativa dos valores da classe para treinamento (símbolos cheios) e conjunto de teste (símbolos vazios) para discriminação entre amostras adulteradas e puras de misturas B5 pelo modelo PLS1-DA: (a) amostras de BES não adulteradas das adulteradas com OFU, (b) amostras de BMPM não adulteradas das adulteradas com OFU, (c) amostras de BMFU não adulteradas das adulteradas com OFU, (d) amostras de BMS não adulteradas das adulteradas com OFU, (e) amostras de BEFU não adulteradas das adulteradas com OFU (f) amostras de BMS não adulteradas das adulteradas com OFU, onde (◀) amostras não adulteradas de BES, BEFU, BEPM, (■) amostras não adulteradas de BMPM, BMFU, BMS, (♦) Misturas de B5 adulteradas por OFU (- - -) *threshold* de discriminação.

Figura 43a – Estimativa de (◄) amostras não adulteradas de BES das (◆) misturas de B5 adulteradas por OFU.



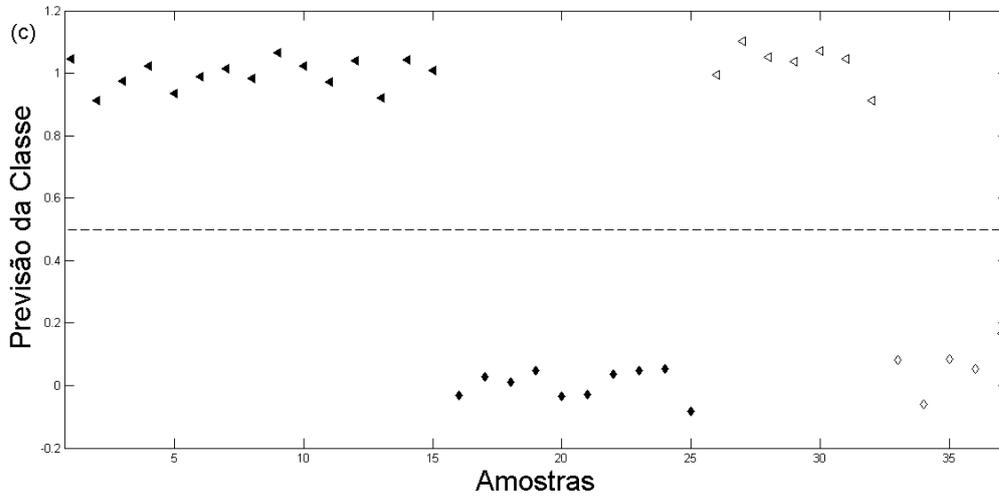
Fonte: o Autor.

Figura 43b – Estimativa de (◄) amostras não adulteradas de BEFU das (◆) misturas de B5 adulteradas por OFU.



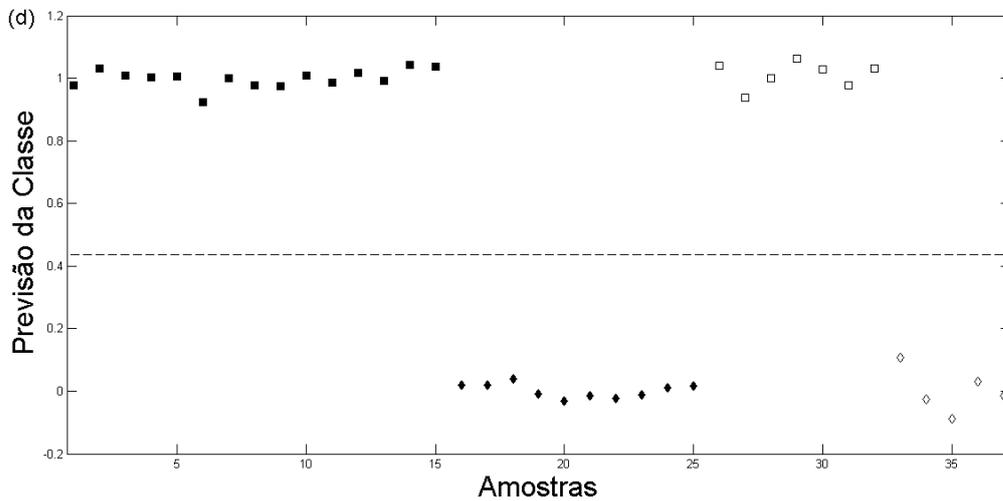
Fonte: o Autor.

Figura 43c – Estimativa de (◀) amostras não adulteradas de BEPM das (◆) misturas de B5 adulteradas por OFU.



Fonte: o Autor.

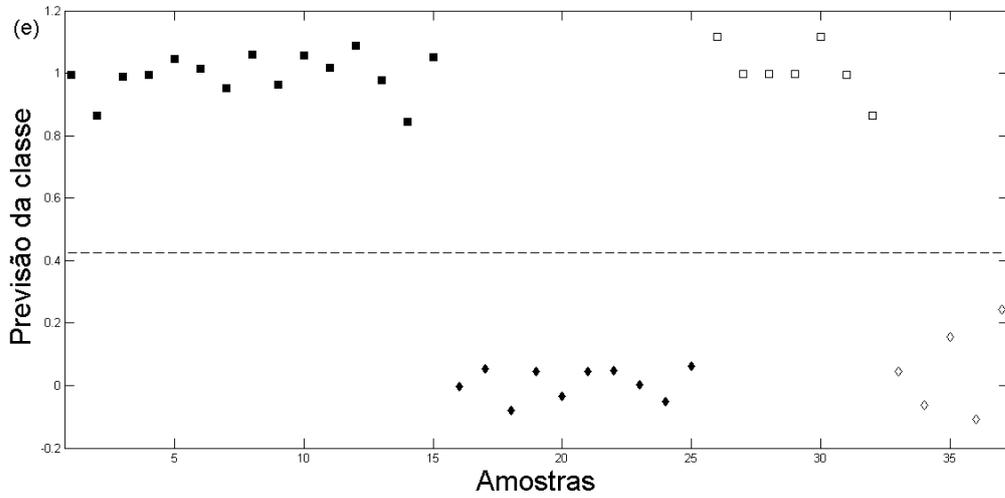
Figura 43d – Estimativa de (■) amostras não adulteradas de BMPM das (◆) misturas de B5 adulteradas por OFU.



Fonte: o Autor.

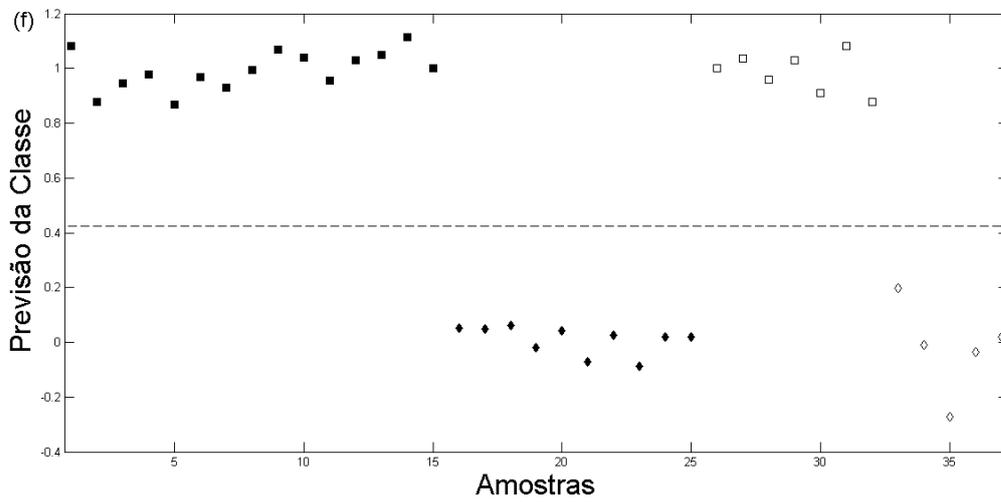
CAPÍTULO 6: MODELOS DE ADULTERAÇÕES EM MISTURAS B5

Figura 43e – Estimativa de (■) amostras não adulteradas de BMFU das (◆) misturas de B5 adulteradas por OFU.



Fonte: o Autor.

Figura 43f – Estimativa de (■) amostras não adulteradas de BMS das (◆) misturas de B5 adulteradas por OFU.



Fonte: o Autor.

## 7 CONCLUSÕES

Neste trabalho, foi possível classificar biodieseis metílicos e etílicos puros (B100) e em misturas com óleo diesel (B5), através de metodologias baseadas na utilização da técnica de espectroscopia MIR aliada ao PLS-DA.

PLS1-DA combinada com a espectroscopia MIR permitiu classificar os biodieseis de acordo com a origem (tipo de óleo usado) e rota usada no processo de produção (metílica ou etílica).

O emprego do PLS2-DA aliada à espectroscopia na região do infravermelho médio demonstrou ser um método analítico para classificação simultânea de diferentes tipos de biodieseis provenientes de óleos (óleo de soja e óleo de fritura usado) e rotas (metílica e etílica) no mesmo modelo, portanto, a classificação simultânea pode ser usada na indústria com intuito de realizar as análises em menor tempo.

Este estudo apresentou 100% de classificação correta, o que mostra que a espectroscopia MIR combinada com PLS1-DA, podem ser aplicadas com segurança na classificação entre amostras puras de biodiesel/diesel B5 e adulteradas. A correta classificação de amostras adulteradas em relação às amostras puras de B5 no conjunto teste, indica que os modelos desenvolvidos podem ser usados no controle de qualidade de biodiesel. Portanto, os resultados sugerem que a metodologia proposta é uma alternativa promissora para ser usada na detecção de adulteração em mistura biodiesel/diesel.

As metodologias desenvolvidas apresentam-se como uma interessante alternativa de triagem, sendo uma ferramenta eficiente para o controle de qualidade de biodieseis que visam classificar amostras com objetivo de identificar a origem, rotas e possíveis adulterações, usando uma análise rápida, amostragem direta, sem necessidade de pré-tratamento das amostras, não destrutivas, garantindo a integridade da amostra que poderá ser novamente analisada, um reduzido volume de amostra gasto, de aproximadamente 0,5 ml por amostra. Portanto, podem ser utilizadas ao longo da linha de produção e nos laboratórios.

## 8 REFERÊNCIAS

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS.

**Boletim Mensal do Biodiesel - Julho de 2014.** SUPERINTENDÊNCIA DE REFINO, P. D. G. N. E. P. D. B. Brasil, 2014. 13 p. Disponível em: <<http://www.anp.gov.br/?dw=71842>>. Acesso em: 10 set. 2014.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS.

**Regulamento Técnico ANP N° 4/2012.** Brasil, 2012. 10 p. Disponível em: <[http://nxt.anp.gov.br/nxt/gateway.dll/leg/resolucoes\\_anp/2012/maio/ranp%2014%20-%202012.xml?fn=document-frameset.htm\\$f=templates\\$3.0](http://nxt.anp.gov.br/nxt/gateway.dll/leg/resolucoes_anp/2012/maio/ranp%2014%20-%202012.xml?fn=document-frameset.htm$f=templates$3.0)>. Acesso em: 24 set. 2014.

ALBERICI, R. M.; SIMAS, R. C.; DE SOUZA, V.; DE SÁ, G. F.; DARODA, R. J.; EBERLIN, M. N. Analysis of fuels via easy ambient sonic-spray ionization mass spectrometry. **Analytica Chimica Acta**, Amsterdã, v. 659, p. 15-22, 2010.

ALMEIDA, E. W. C. et al. Estudo teórico experimental de espectros infravermelho de ésteres de ácido graxo presentes na composição do biodiesel de soja. **Química Nova**, São Paulo, v. 35, n.9, p. 1752–57, 2012.

ALMEIDA, M. R. et al. Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation. **Talanta**, Londres, v. 117, p. 305–11, 2013.

ALVES, J. C.; HENRIQUES, C. B.; POPPI, R. J. Classification of diesel pool refinery streams through near infrared spectroscopy and support vector machines using C-SVC and nu-SVC. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, Amsterdã, v. 117, p. 389-96, 2014.

ANDRADE, J. M. et al. A review of the main factors influencing the FT-IR-PLS abilities exemplified with petrochemical qualimetric applications. **Talanta**, Londres, v. 44, p. 2167–84, 1997.

AOUIDI, F. et al. Discrimination of five Tunisian cultivars by Mid InfraRed spectroscopy combined with chemometric analyses of olive *Olea europaea* leaves. **Food Chemistry**, Londres, v. 131, n. 1, p. 360-66, 2012.

- ARAUJO, V. K. W. S.; HAMACHER, S.; SCAVARDA, L. F. Economic assessment of biodiesel production from waste frying oils. **Bioresource Technology**, v. 101, p. 4415-4422, 2010.
- BARKER, M.; RAYENS, W. Partial least squares for discrimination. **Journal of Chemometric**, Chichester, v. 17, p. 166-173, 2003.
- BALABIN, R. M.; SAFIEVA, R. Z. Biodiesel classification by base stock type (vegetable oil) using near infrared spectroscopy data. **Analytica Chimica Acta**, Amsterdã, v. 689, n. 2, p. 190-97, 2011.
- BALABIN, R. M.; SYUNYAEV, R. Z. Petroleum resins adsorption onto quartz sand: near infrared (NIR) spectroscopy study. **Journal of Colloid and Interface Science**, New York, v. 318, n. 2, p. 167-74, 2008.
- BAPTISTA, P. et al. Multivariate near infrared spectroscopy models for predicting the iodine value, CFPP, kinematic viscosity at 40 degrees C and density at 15 degrees C of biodiesel. **Talanta**, Londres, v. 77, n. 1, p. 144-51, 2008.
- BARTH, A. Infrared spectroscopy of proteins. **Biochimica et Biophysica Acta**, Amsterdã, v. 1767, n. 9, p. 1073-101, 2007.
- BEEBE, K.R.; PELL, R.J.; SEASHOLTZ, M.B. **Chemometrics: a practical guide**. New York: Wiley-Interscience, 1998. 348p.
- BORRAS, E. et al. Fast and robust discrimination of almonds (*Prunus amygdalus*) with respect to their bitterness by using near infrared and partial least squares-discriminant analysis. **Food Chemistry**, Londres, v. 153, p. 15-9, 2014.
- BRANDAO, L. F.; BRAGA, J. W.; SUAREZ, P. A. Determination of vegetable oils and fats adulterants in diesel oil by high performance liquid chromatography and multivariate methods. **Journal of Chromatography A**, Amsterdã, v. 1225, p. 150-57, 2012.
- BRERETON, R. G. **Chemometrics: Data Analysis for the Laboratory and Chemical Plant**. Chichester: John Wiley & Sons, 2003a.
- \_\_\_\_\_. Introduction to multivariate calibration in analytical chemistry. **Analyst**, Londres, v. 125, p. 2125-54, 2000b.

- BRUNS, R.; FAIGLE, J. F. G. Quimiometria. **Química Nova**, São Paulo, v. 8, n. 2, p. 84–99, 1985.
- CHAKRABORTY, S. et al. Analysis of petroleum contaminated soils by spectral modeling and pure response profile recovery of n-hexane. **Environmental Pollution**, Barking, v. 190, p. 10-8, Jul 2014.
- CHEN, H.; BAKSHI, B. R.; GOEL, P.K. Toward Bayesian chemometrics- A tutorial on some recent advances. **Analytica Chimica Acta**, Amsterdã, v. 602, p. I-I6, 2007.
- CHEN, L.-Y. et al. Fuel properties and combustion characteristics of jatropha oil biodiesel–diesel blends. **Journal of the Taiwan Institute of Chemical Engineers**, [S.l.], v. 44, n. 2, p. 214-20, 2013.
- CHONG, I.-G.; JUN, C.-H. Performance of some variable selection methods when multicollinearity is present. **Chemometrics and Intelligent Laboratory Systems**, Amsterdã, v. 78, p. 103–12, 2005.
- CHUI, Q. S. H.; ZUCCHINI, R. R.; LICHTIG, J. Qualidade de medições em química analítica. Estudo de caso: determinação de cádmio por espectrofotometria de absorção atômica com chama. **Química Nova**, São Paulo, v. 24, p. 374-80, 2001.
- CLAYDEN, J. et al. **Organic Chemistry**. 1<sup>st</sup> Ed. New York: Oxford University Press, 2011.
- CORGOZINHO, C. N. C.; PASA, V. M. D.; BARBEIRA, P. J. S. Determination of residual oil in diesel oil by spectrofluorimetric and chemometric analysis. **Talanta**, Londres, v. 76, p. 479-484, 2008.
- CORREIA, P. R. M.; FERREIRA, M. M. C. Reconhecimento de Padrões por Métodos não Supervisionados: explorando procedimentos quimiométricos para tratamento de dados analíticos. **Química Nova**, São Paulo, v. 30, n. 2, p. 481-487, 2007.
- COZZOLINO, D. An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals. **Food Research International**, Barking, v. 60, p. 262-65, 2014.
- COZZOLINO, D.; ROUMELIOTIS, S.; EGLINTON, J. An attenuated total reflectance mid infrared (ATR-MIR) spectroscopy study of gelatinization in barley. **Carbohydrate Polymers**, Barking, v. 108, p. 266-71, 2014a.

- \_\_\_\_\_. Evaluation of the use of attenuated total reflectance mid infrared spectroscopy to determine fatty acids in intact seeds of barley (*Hordeum vulgare*). **LWT - Food Science and Technology**, Londres, v. 56, n. 2, p. 478-83, 2014b.
- CVENGROŠ, J.; CVENGROŠOVÁ, Z. Used frying oils and fats and their utilization in the production of methyl esters of higher fatty acids. **Biomass and Bioenergy**, Oxford, v. 27, n. 2, p. 173-81, 2004.
- DA COSTA FILHO, P. A. Developing a rapid and sensitive method for determination of *trans*-fatty acids in edible oils using middle-infrared spectroscopy. **Food Chemistry**, Londres, v. 158, p. 1-7, 2014.
- DA SILVA, M. P. F. et al. Classification of gasoline as with or without dispersant and detergent additives using infrared spectroscopy and multivariate classification. **Fuel**, Londres, v. 116, p. 151-57, 2014.
- DA SILVA, V. A. G. et al. Non-destructive identification of different types and brands of blue pen inks in cursive handwriting by visible spectroscopy and PLS-DA for forensic analysis. **Microchemical Journal**, New York, v. 116, p. 235-43, 2014.
- DE SOUSA, L. S. et al. Use of natural antioxidants in soybean biodiesel. **Fuel**, Londres, v. 134, p. 420-28, 2014.
- DEMTRÖDER, W. **Laser Spectroscopy: basic concepts and instrumentaion**. 3<sup>rd</sup> Ed. New York: Springer, 2003.
- DIAS, J. M. et al. Study of an ethylic biodiesel integrated process: Raw-materials, reaction optimization and purification methods. **Fuel Processing Technology**, Amsterdã, v. 124, p. 198-205, 2014.
- DIVAKARA, B. N. et al. Biology and genetic improvement of *Jatropha curcas* L.: A review. **Applied Energy**, Londres, v. 87, n. 3, p. 732-42, 2010.
- ENCINAR, J. M.; GONZÁLEZ, J. F.; RODRÍGUEZ-REINARES, A. Ethanolysis of used frying oil. Biodiesel preparation and characterization. **Fuel Processing Technology**, Amsterdã, v. 88, n. 5, p. 513-22, 2007.
- EUROPEAN COMMISSION- EC. **Directive of the European Parliament and of the Council**. Brussels: EN 0288/2012: 23 p. 2012. Disponível em:

<[http://ec.europa.eu/energy/renewables/biofuels/doc/biofuels/com\\_2012\\_0595\\_en.pdf](http://ec.europa.eu/energy/renewables/biofuels/doc/biofuels/com_2012_0595_en.pdf)>.

Acesso em: 15 jun. 2014.

FERREIRA, M.M.C. Multivariate QSAR. **Journal of the Brazilian Chemical Society**, São Paulo, v. 13, n. 6, p. 742-753, 2002.

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S. Quimiometria I: calibração multivariada, um tutorial. . **Química Nova**, São Paulo, v. 22, p. 724-31, 1999.

GALTIER, O. et al. Comparison of PLS1-DA, PLS2-DA and SIMCA for classification by origin of crude petroleum oils by MIR and virgin olive oils by NIR for different spectral regions. **Vibrational Spectroscopy**, Amsterdã, v. 55, p. 132-40, 2011.

GAUGLITZ, G.; VO-DINH, T. **Handbook of Spectroscopy**. Weinheim: Wiley- VCH, 2003.

GAYDOU, V.; DUPUY, J. K. N. Evaluation of multiblock NIR/MIR PLS predictive models to detect adulteration of diesel/biodiesel blends by vegetal oil. **Chemometrics and Intelligent Laboratory Systems**, Amsterdã, v. 106, p. 190–197, 2011.

HELLAND, I. S. On the structure of partial least square regression. **Communications in statistics – simulation and computation**, New York, v. 17, p. 581-607, 1998.

JIN, B. et al. Comprehensive utilization of the mixture of oil sediments and soapstocks for producing FAME and phosphatides. **Fuel Processing Technology**, Amsterdã, v. 89, n. 1, p. 77-82, 2008.

KOWALSKI, B. R. Chemometrics: Views and Propositions. **Journal of Chemical Information and Modeling**, Washington, v. 15, n. 4, p. 201-03, 1975.

KRÄMER, K.; EBEL, S. Application of NIR reflectance spectroscopy for the identification of pharmaceutical excipients. **Analytica Chimica Acta**, Amsterdã, v. 420, n. 2, p. 155-61, 2000.

KUMAR TIWARI, A.; KUMAR, A.; RAHEMAN, H. Biodiesel production from jatropha oil (*Jatropha curcas*) with high free fatty acids: An optimized process. **Biomass and Bioenergy**, Oxford, v. 31, n. 8, p. 569-75, 2007.

- LAM, M. K.; LEE, K. T.; MOHAMED, A. R. Homogeneous, heterogeneous and enzymatic catalysis for transesterification of high free fatty acid oil (waste cooking oil) to biodiesel: a review. **Biotechnology Advances**, New York, v. 28, n. 4, p. 500-18, 2010.
- LEUNG, D. Y. C.; WU, X.; LEUNG, M. K. H. A review on biodiesel production using catalyzed transesterification. **Applied Energy**, Londres, v. 87, n. 4, p. 1083-95, 2010.
- LOPES, A. F. G. et al. Speed of sound in pure fatty acid methyl esters and biodiesel fuels. **Fuel**, Londres, v. 116, p. 242-54, 2014.
- MARTENS, H.; NAES, T. **Multivariate calibration**. New York: Wiley, 1996.
- MASSART, D. L.; VANDEGINSTE, B. G. M.; BUYDENS, S. J.; LEWI, P. J.; SMEYERS-VERBEKE. **Handbook of Chemometrics and Qualimetrics: Parte B**. Amsterdam: Elsevier, 1997.
- MAZIVILA, S. J. et al. Discrimination of the type of biodiesel/diesel blend (B5) using mid-infrared spectroscopy and PLS-DA. **Fuel**, Londres, v. 142, p. 222-226, 2015a.
- \_\_\_\_\_. Fast Detection of Adulterants/Contaminants in Biodiesel/Diesel Blend (B5) Employing Mid-Infrared Spectroscopy and PLS-DA. **Energy & Fuels**, Washington, v. 29, p.227-232, 2015b.
- MEHMOOD, T. et al. A review of variable selection methods in Partial Least Squares Regression. **Chemometrics and Intelligent Laboratory Systems**, Amsterdã, v. 118, p 62–69, 2012.
- MENDONÇA, S.; LAVIOLA, B. G. **Comunicado Técnico: Uso Potencial e Toxidez da Torta de Pinhão-manso**. AGROENERGIA, E. Brasília, DF: Embrapa: 8 p. 2009.
- MIRANDA, A. M. et al. Line shape analysis of the Raman spectra from pure and mixed biofuels esters compounds. **Fuel**, Londres, v.115, p. 118–25, 2014.
- MITTELBAACH, M. Diesel fuel derived from vegetable oils, VI: Specifications and quality control of biodiesel. **Bioresource Technology**, Essex, v. 56, n. 1, p. 7-11, 1996.

- MUELLER, D. et al. Fourier transform infrared spectroscopy (FTIR) and multivariate analysis for identification of different vegetable oils used in biodiesel production. **Sensors**, Basel, v. 13, n. 4, p. 4258-71, 2013.
- OM TAPANES, N. C. et al. Transesterification of Jatropha curcas oil glycerides: Theoretical and experimental studies of biodiesel reaction. **Fuel**, Londres, v. 87, n. 10-11, p. 2286-95, 2008.
- OTTO, M. **Chemometrics: Statistics and Computer application in Analytical Chemistry**. Weinheim: Wiley- VCH, 1999.
- PASTORE, T. C. et al. Near infrared spectroscopy (NIRS) as a potencial tool for monitoring trade of similar Woods: Discrimination of true mahogany, cedar, andiroba, and curupixá. **Holzforshung**, Berlim, v.66, p. 73-80, 2011.
- PAVIA, D. L. et al. **Introdução à espectroscopia**. Tradução de Pedro Barros. São Paulo: Cengage Learning, 2012. Título original: Introduction to spectroscopy 4. ed. norte-americana.
- PEREZ, A. et al. Winterization of peanut biodiesel to improve the cold flow properties. **Bioresour Technology**, Essex, v. 101, n. 19, p. 7375-81, 2010.
- PETROBRAS. **Petrobras Biocombustível: Suprimento e Tecnologia Agrícola para Biodiesel**. 16 p. Disponível em:  
<<http://www.redebiodiesel.com.br/arquivos/download/64.pdf>>. Acesso em: 15 jun. 2014.
- PONTES, M. J. et al. Screening analysis to detect adulteration in diesel/biodiesel blends using near infrared spectrometry and multivariate classification. **Talanta**, Londres, v. 85, n. 4, p. 2159-65, 2011.
- RINCÓN, L. E.; JARAMILLO, J. J.; CARDONA, C. A. Comparison of feedstocks and technologies for biodiesel production: An environmental and techno-economic evaluation. **Renewable Energy**, Oxford, v. 69, p. 479-87, 2014.
- ROGGO, Y. et al. A review of near infrared spectroscopy and chemometrics in pharmaceutical Technologies. **Journal of Pharmaceutical and Biomedical Analysis**, Amsterdã, v. 44, p. 683–700, 2007.

- SÁNCHEZ-ALONSO, I., CARMONA, P., CARECHE, M. Vibrational spectroscopic analysis of hake (*Merluccius merluccius* L.) lipids during frozen storage. **Food Chemistry**, Londres, v. 132, p. 160–67, 2012.
- SAVORANI, F. et al. A primer to nutritional metabolomics by NMR spectroscopy and chemometrics. **Food Research International**, Barking, v. 54, n. 1, p. 1131-45, 2013.
- SCHALE, S. P.; LE, T. M.; PIERCE, K. M. Predicting feedstock and percent composition for blends of biodiesel with conventional diesel using chemometrics and gas chromatography-mass spectrometry. **Talanta**, Londres, v. 94, p. 320-27, 2012.
- SCHUT, M. et al. Multi-actor governance of sustainable biofuels in developing countries: The case of Mozambique. **Energy Policy**, Surrey, v. 65, p. 631–43, 2014.
- SEKULIC, S. et al. Nonlinear multivariate calibration methods in analytical chemistry, **Analytical Chemistry**, Washington, v. 65, n. 19, A835-A845, 1993.
- SERRANO-LOURIDO, D. et al. Classification and characterisation of Spanish red wines according to their appellation of origin based on chromatographic profiles and hemometric data analysis. **Food Chemistry**, Londres, v. 135, p. 1425–31, 2012.
- SHAH, S.; GUPTA, M. N. Lipase catalyzed preparation of biodiesel from *Jatropha* oil in a solvent free system. **Process Biochemistry**, Londres, v. 42, n. 3, p. 409-14, 2007.
- SHARMA, Y. C.; SINGH, B.; UPADHYAY, S. N. Advancements in development and characterization of biodiesel: A review. **Fuel**, Londres, v. 87, n. 12, p. 2355-73, 2008.
- SMITH, B. C. **Fundamentals of Fourier Transform Infrared Spectroscopy**. 2<sup>a</sup> Ed. New York: Taylor & Francis Group, 2011.
- SOARES, I. P. et al. Multivariate Calibration by Variable Selection for Blends of Raw Soybean Oil/Biodiesel from Different Sources Using Fourier Transform Infrared Spectroscopy (FTIR) Spectra Data. **Energy & Fuels**, Washington, v. 22, p. 2079–83, 2008.
- SOUZA, L. M. et al. Quantification of residual automotive lubricant oil as an adulterant in Brazilian S-10 diesel using MIR spectroscopy and PLS. **Fuel**, Londres, v.130, p. 257–262, 2014.

- WESTERHUIS, J. A. et al. Assessment of PLS-DA cross validation. **Metabolomics**, [S.l.], v. 4, p. 81–89, 2008.
- WOLD, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, Amsterdã, v. 58, p. 109-30, 2001.
- WOLD, S.; ESBENSEN, K.; GELADI, P. **Chemometrics and Intelligent Laboratory Systems**, Amsterdã, v. 2, p. 37-52, 1987.
- XU, L. et al. Rapid discrimination of pork in Halal and non-Halal Chinese ham sausages by Fourier transform infrared (FTIR) spectroscopy and chemometrics. **Meat Science**, Barking, v. 92, n. 4, p. 506-10, 2012.
- YINGYING, L. et al. Biodiesel Production from Crude *Jatropha curcas* L. Oil with Trace Acid Catalyst. **Chinese Journal of Chemical Engineering**, [S.l.], v. 20, n. 4, p. 740-46, 2012.
- ZHANG, C.; SU, J. Application of near infrared spectroscopy to the analysis and fast quality assessment of traditional Chinese medicinal products. **Acta Pharmaceutica Sinica B**, [S.l.], v. 4, n. 3, p. 182-92, 2014.
- ZHANG, W.-B. Review on analysis of biodiesel with infrared spectroscopy. **Renewable and Sustainable Energy Reviews**, [S.l.], v. 16, n. 8, p. 6048-58, 2012.
- ZÍLIO, J. et al. Avaliação de Biocombustíveis em Moçambique. 515 p. Disponível em: <[https://energypedia.info/images/0/0b/PT\\_Avalia%C3%A7%C3%A3o\\_dos\\_Biocombust%C3%ADveis\\_em\\_Mo%C3%A7ambique\\_Econergy\\_International\\_Corporation%3Bet.al.pdf](https://energypedia.info/images/0/0b/PT_Avalia%C3%A7%C3%A3o_dos_Biocombust%C3%ADveis_em_Mo%C3%A7ambique_Econergy_International_Corporation%3Bet.al.pdf)>. Acesso em: 25 jun. 2014.

**Anexos: Publicações decorrentes do trabalho**

[1] MAZIVILA, S. J.; SANTANA, F. B; MITSUTAKE, H.; GONTIJO, L. C.; SANTOS, D. Q.; BORGES NETO, W. Discrimination of the type of biodiesel/diesel blend (B5) using mid-infrared spectroscopy and PLS-DA. **Fuel** 142 (2015) 222-226.

[2] MAZIVILA, S. J.; GONTIJO, L. C.; SANTANA, F. B; MITSUTAKE, H.; SANTOS, D. Q.; BORGES NETO, W. Fast Detection of Adulterants/Contaminants in Biodiesel/Diesel Blend (B5) Employing Mid-Infrared Spectroscopy and PLS-DA. **Energy & Fuels** 29 (2015) 227-232.

[3] MAZIVILA, S. J.; MITSUTAKE, H.; SANTANA, F. B; GONTIJO, L. C.; SANTOS, D. Q.; BORGES NETO, W. Fast classification of different oils and routes used in biodiesel production using mid infrared spectroscopy and PLS2-DA. Journal of the Brazilian Chemical Society. **Artigo aceito para publicação.**

[4] SOUZA, L.M.; SANTANA, F.B.; GONTIJO, L.C.; MAZIVILA, S.J.; BORGES NETO, W. QUANTIFICATION OF ADULTERATIONS IN EXTRA VIRGIN FLAXSEED OIL USING MIR AND PLS. Food Chemistry. **Artigo aceito com correções.**