



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**  
**UNIVERSIDADE FEDERAL DOS VALES DO**  
**JEQUITINHONHA E MUCURI**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM**  
**BIOCOMBUSTÍVEIS**



**JOSÉ EDUARDO BUIATTE**

**USO DE ESPECTROSCOPIA NO INFRAVERMELHO MÉDIO E**  
**ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS**  
**PARCIAIS NO CONTROLE DE QUALIDADE DE BIODIESEL DE**  
**PINHÃO MANSO E CRAMBE**

**UBERLÂNDIA**  
**2019**



**JOSÉ EDUARDO BUIATTE**

**USO DE ESPECTROSCOPIA NO INFRAVERMELHO MÉDIO E  
ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS  
PARCIAIS NO CONTROLE DE QUALIDADE DE BIODIESEL DE  
PINHÃO MANSO E CRAMBE**

Tese apresentada ao Programa de Pós-Graduação em Biocombustíveis da Universidade Federal dos Vales do Jequitinhonha e Mucuri e Universidade Federal de Uberlândia como exigência parcial para obtenção de título de Doutor em Biocombustíveis.

Orientador: Prof. Dr. Douglas Queiroz Santos

**UBERLÂNDIA**

**2019**

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU  
com dados informados pelo(a) próprio(a) autor(a).

B932 2019	<p>Buiatte, José Eduardo, 1960- Uso de Espectroscopia no Infravermelho Médio e Análise Discriminante por Quadrados Mínimos Parciais no Controle de Qualidade de Biodiesel de Pinhão Manso e Crambe [recurso eletrônico] / José Eduardo Buiatte. - 2019.</p> <p>Orientador: Douglas Queiroz Santos. Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Biocombustíveis. Modo de acesso: Internet. Disponível em: <a href="http://doi.org/10.14393/ufu.te.2019.2541">http://doi.org/10.14393/ufu.te.2019.2541</a> Inclui bibliografia.</p> <p>1. Biocombustível. I. Santos, Douglas Queiroz, 1980-, (Orient.). II. Universidade Federal de Uberlândia. Pós-graduação em Biocombustíveis. III. Título.</p> <p style="text-align: right;">CDU: 662.756</p>
--------------	---

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:  
Gizele Cristine Nunes do Couto - CRB6/2091  
Nelson Marcos Ferreira - CRB6/3074



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**  
Coordenação do Programa de Pós-Graduação em  
Biocombustíveis

Av. João Naves de Ávila, 2121 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902

Telefone: (34) 3239-4385 / 4208 - www.iq.ufu.br - ppbic@iqufu.ufu.br



**ATA DE DEFESA - PÓS-GRADUAÇÃO**

Programa de Pós-Graduação em:	Biocombustíveis				
Defesa de:	Doutorado, 11, PPBIC				
Data:	13 de dezembro de 2019	Hora de início:	09:00	Hora de encerramento:	<b>12:30</b>
Matrícula do Discente:	11623PGB002				
Nome do Discente:	José Eduardo Buiatte				
Título do Trabalho:	USO DE ESPECTROSCOPIA NO INFRAVERMELHO MÉDIO E ANÁLISE DISCRIMINANTE POR QUADRADOS MÍNIMOS PARCIAIS NO CONTROLE DE QUALIDADE DE BIODIESEL DE PINHÃO MANSO E CRAMBE				
Área de concentração:	Biocombustíveis				
Linha de pesquisa:	Produtos e Coprodutos				
Projeto de Pesquisa de vinculação:					

Reuniu-se no Anfiteatro/Sala Auditório Prof. Dr. Manuel Gonzalo Hernández-Terrones, piso superior do Bloco 5I no Campus Santa Mônica, da Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Biocombustíveis, assim composta: Professores Doutores: Carlos Alberto de Oliveira e Jefferson Luis Ferrari, da Universidade Federal de Uberlândia; João Paulo Victorino Santos, do Instituto Federal de Goiás; Frederico Garcia Pinto, Universidade Federal de Viçosa; e Douglas Queiroz Santos, orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Douglas Queiroz Santos, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

**Aprovado(a).**

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Douglas Queiroz Santos, Professor(a) do Ensino Básico, Técnico e Tecnológico**, em 13/12/2019, às 12:30, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Carlos Alberto de Oliveira, Professor(a) do Magistério Superior**, em 13/12/2019, às 12:35, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Jefferson Luis Ferrari, Professor(a) do Magistério Superior**, em 13/12/2019, às 12:35, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **JOÃO PAULO VICTORINO SANTOS, Usuário Externo**, em 13/12/2019, às 12:37, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Frederico Garcia Pinto, Usuário Externo**, em 13/12/2019, às 12:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://www.sei.ufu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **1701795** e o código CRC **0B41E127**.

*Dedico esta Tese principalmente a Deus e aos meus pais, Laerte e Abigail, pelo dom da vida. À minha esposa Raquel, que se manteve ao meu lado desde o início. Às minhas filhas Isabella e Izadora, por sempre me apoiar e auxiliar nos trabalhos. Minha avó Astina Campos por ter me dado o incentivo inicial. À todos os meus irmãos e família que estiveram ao meu lado e souberam entender e me apoiar incondicionalmente na busca por mais esse objetivo de vida. Amo vocês.*

*"Bom mesmo, é ir à luta com muita alegria, determinação e perseverança, abraçando a vida, ajudando nosso semelhante e vivendo com paixão, sabendo perder com classe e vencer com ousadia, pois, o triunfo pertence a quem vai à luta, se atreve e não desiste de seus sonhos. A vida é algo muito importante para ser insignificante".*

*Charles Chaplin*

## AGRADECIMENTOS

Agradeço principalmente a DEUS por ter me permitido a vida e a capacidade de realizar esse trabalho, dando-me saúde, paz e persistência em todos esses anos; ‘Obrigado meu PAI pelo dom da vida’. A todos os meus santos protetores que intercederam junto a DEUS por mim.

Ao meu orientador e amigo Prof. Dr. Douglas Queiroz Santos pelos ensinamentos, paciência e persistência na realização desse trabalho. Em especial, ao Prof. Dr. Waldomiro Borges Neto pela parceria e por disponibilizar o Laboratório de Quimiometria do Triângulo (LQT) para as análises dando-me a oportunidade de desenvolver esse projeto e acreditar na minha capacidade. Certamente carregarei muito das suas experiências e ensinamentos durante toda a minha vida. Vocês são um espelho para todos nós como profissionais e como pessoas Prof. Dr. Douglas Queiroz Santos e Prof. Dr. Waldomiro Borges Neto.

Aos meus colegas do LQT: Eloiza Guimarães, Lucas Caixeta Gontijo, Hery Mitsutake, Letícia Maria Sousa, Felipe Bachion Santana, Edvando Souza Telles, Baltazar Vasco Siteo, Ademar Domingos Viagem Máquina, Lucas Gustavo da Costa, Tathiana Dienifer Godinho, Raquel Boaventura de Moraes e Maria Teresa Costa Carvalho, pelo incentivo, companheirismo, amizade, paciência, discussões, viagens... Foram muitos os momentos de tensões, alegrias, gargalhadas e festas.

A todas as pessoas que um dia fizeram parte da minha formação profissional, pessoal e religiosa e que, direta ou indiretamente contribuíram para a realização deste trabalho.

A Universidade Federal de Uberlândia e à Universidade Federal dos Vales do Jequitinhonha e Mucuri, seus coordenadores: Prof. Dr. Alexandre Soares dos Santos e Prof. Dr. Ricardo Reis Soares, ao corpo docente, direção e administração pelo apoio.

Ao Instituto de Química pelo fornecimento de toda a estrutura física e tecnológica para a realização do trabalho.

Aos companheiros de profissão pelo apoio e estímulo durante minha jornada: André, Aníbal, Edmilson, Edvando, Magayver, Ivan, Moacir, Otávio e Roni.

Ao secretário Gabriel Fonseca, pelo acompanhamento dos processos técnico-administrativos.

A TRANSPETRO pelo fornecimento do diesel utilizado no preparo das amostras.

## RESUMO

O biodiesel é um combustível renovável derivado de óleos vegetais ou gorduras animais que tem se destacado na matriz energética Brasileira e de vários países, como opção energética e contraponto a dependência de combustíveis fósseis. Introduzido na matriz energética brasileira no ano de 2005, ele é comercializado em mistura com o diesel na proporção estabelecida pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP). Embora existam órgãos reguladores e fiscalizadores que garantam a qualidade dos combustíveis comercializados, podem ocorrer determinadas formas de adulterações e contaminações que comprometem a qualidade deste combustível. Neste trabalho, a técnica da Espectroscopia no Infravermelho Médio (MIR) aliada ao método quimiométrico da Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) foi aplicada para desenvolver um método eficiente na classificação de biodiesel de Pinhão Manso e de Crambe em misturas com o diesel, tanto em relação ao teor de biodiesel presente na mistura com o diesel quanto a presença de adulterantes como: óleo de Soja, óleo residual, óleo automotivo usado e gasolina. Foram utilizadas entre 58 a 60 amostras para os conjuntos de treinamento dos modelos, com os teores variando de 1,00 a 30,00 (% m/m) de biodiesel e adulterante. Os conjuntos testes, isto é, para a validação do modelo, foram construídos com aproximadamente 30 amostras cada. A eficiência destes modelos foi avaliada com base nas figuras de mérito, sensibilidade e especificidade. Os valores de limites a 95% de confiança (*threshold*) foram escolhidos com base no teorema de Bayes, segundo o qual os valores de  $y_{\text{previstos}}$  seguem uma distribuição semelhante aos que serão observados para as amostras futuras, de forma a minimizar a ocorrência de falsos positivos e falsos negativos nas amostras classificadas. A relação entre a especificidade e a sensibilidade também foi avaliada graficamente usando as curvas Característica de Operação do Receptor (ROC). Estes parâmetros apresentaram valores variando entre 0,93 a 1,00, o que representa classificação correta das amostras de calibração e previsão acima de 93,30%. Os bons resultados dos modelos comprovam que de fato, o método analítico é viável e eficiente para o controle qualitativo destes combustíveis.

**Palavras-chave:** Biocombustível. FT-MIR. PLS-DA. *Crambe abyssynica* H. *Jatropha curcas* L. Quimiometria.

## ABSTRACT

Biodiesel is a renewable fuel derived from vegetable oils or animal fats that has stood out in the Brazilian energy matrix and in several countries as an energy option and counter points the dependence on fossil fuels. The biodiesel was included into the Brazilian energy matrix in 2005. It is usually marketed in mixture with the mineral diesel in proportions legally established by the National Agency of Petroleum, Natural Gas and Biofuels (ANP). Although there are regulatory and oversight organizations that assure the quality of marketed fuels, certain forms of adulterations and contaminations that affect the quality of this fuel do sometimes occur. In this work the use of medium infrared spectroscopy (MIR) data analyzed with the chemometric method of Discriminating Analysis by Partial Least Squares discriminating analysis, according to the partial least squares (PLS-DA) criterion, was used in an attempt to develop an efficient method to classify biodiesels derived from oils from *mansonia pinus* and *Crambe* in mixtures with the diesel, both in relation to the biodiesel in the mixture with the diesel along with common adulterants as like: soybean oil, waste oil, used automotive oil and gasoline. Between 58 and 60 samples were used to check the models, with contents varying between 1.00 and 30.00 mass% biodiesel and adulterants. The test sets, that is, for the validation of the model, were constructed with approximately 30 samples each. The efficiency of these models was assessed on the basis of figures of merit, sensitivity and specificity. The threshold values were chosen based on the Bayes', whereby the  $y_{\text{predicted}}$  values result in a similar distribution to the values that will be observed for future samples, theorem, in order to minimize the occurrence of false positives and false negatives for the classified samples. The relation between specificity and sensitivity were also evaluated graphically using the receiver operating characteristic (ROC) curves. These parameters presented values ranging from 0.93 to 1.00, which represents correct classification of the calibration and forecast samples above 93.30%. From the good results of these models, it was prove it that, in fact, the analytical method is indeed reliable and that the qualitative control of these fuels is efficient enough.

**Keywords:** Biofuel. FT-MIR. PLS-DA. *Crambe abissynica* H. *Jatropha curcas* L. Chemometrics

## LISTA DE FIGURAS

<b>Figura 1</b> – Matriz energética de Transportes no Brasil e no mundo.....	19
<b>Figura 2</b> – Esquema da reação de transesterificação de óleos vegetais .....	21
<b>Figura 3</b> – Esquema da reação de esterificação de óleos vegetais .....	21
<b>Figura 4</b> – Evolução do percentual de teor de biodiesel presente no diesel fóssil no Brasil..	23
<b>Figura 5</b> – Matérias primas utilizadas para a produção de biodiesel no Brasil .....	24
<b>Figura 6</b> – Partes da planta de <i>Crambe</i> : (a) árvore; (b) fruto e (c) sementes .....	26
<b>Figura 7</b> – Partes da planta de Pinhão Manso: (a) árvore; (b) corte longitudinal do fruto e (c) fruto aberto .....	27
<b>Figura 8</b> – Distribuição percentual das não conformidades do óleo diesel .....	29
<b>Figura 9</b> – (a) Reflectância de radiação no interior do cristal de ZnSe; (b) acessório de ATR .....	32
<b>Figura 10</b> – Organização dos dados para a construção da matriz <b>X</b> .....	33
<b>Figura 11</b> – Representação da matriz de dados <b>X</b> e o vetor <b>y</b> usados no PLS-DA .....	34
<b>Figura 12</b> – Resultado gráfico do modelo PLS-DA .....	37
<b>Figura 13</b> – Fluxograma da reação de esterificação do biodiesel .....	42
<b>Figura 14</b> – Esquema da reação de transesterificação do biodiesel .....	43
<b>Figura 15</b> – Espectros MIR de diesel puro (azul), B9 de Pinhão Manso (vermelho) e B10 de <i>Crambe</i> (verde) .....	46
<b>Figura 16</b> – Espectros MIR das amostras usadas para a construção dos modelos de (a) teor de biodiesel metílico de Pinhão Manso e (b) teor de biodiesel metílico de <i>Crambe</i> .....	47
<b>Figura 17</b> – Estimativas de B7 de biodiesel metílico de Pinhão Manso.....	48
<b>Figura 18</b> – Estimativas de B10 de biodiesel metílico de <i>Crambe</i> .....	49
<b>Figura 19</b> – ROC dos modelos PLS-DA para teor de biodiesel metílico de Pinhão Manso.....	54
<b>Figura 20</b> – ROC dos modelos PLS-DA para teor de biodiesel metílico de <i>Crambe</i> .....	55
<b>Figura 21</b> – Pesos dos modelos PLS-DA para teor de biodiesel metílico de (a) Pinhão Manso e (b) <i>Crambe</i> .....	56

<b>Figura 22</b> – VIP dos modelos PLS-DA para teor de biodiesel metílico de (a) Pinhão Manso e (b) Crambe.....	57
<b>Figura 23</b> – Espectros MIR das amostras de biodiesel/diesel de Pinhão Manso adulterados com (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina .....	59
<b>Figura 24</b> – Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com óleo de Soja .....	60
<b>Figura 25</b> – Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com óleo residual .....	61
<b>Figura 26</b> – Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com óleo automotivo .....	62
<b>Figura 27</b> – Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com gasolina .....	63
<b>Figura 28</b> - ROC das amostras de biodiesel/diesel de Pinhão Manso para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com óleo de Soja .....	66
<b>Figura 29</b> – ROC das amostras de biodiesel/diesel de pinhão manso para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com óleo residual .....	67
<b>Figura 30</b> – ROC das amostras de biodiesel/diesel de Pinhão Manso para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com óleo automotivo usado .....	68
<b>Figura 31</b> –ROC das amostras de biodiesel/diesel de Pinhão Manso para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com gasolina... ..	69
<b>Figura 32</b> –VIP para PLS-DA das amostras de Pinhão Manso para as classes adulteradas para os conjuntos de adulteração com (a) óleo de Soja; (b) óleo residual; c) óleo automotivo usado e (d) gasolina .....	70
<b>Figura 33.</b> Espectros MIR de amostras biodiesel/diesel de Crambe adulteradas com (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina .....	72
<b>Figura 34</b> – Estimativas das amostras autênticas de biodiesel/diesel de Crambe para o conjunto de adulteração com óleo de Soja .....	73
<b>Figura 35</b> – Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com óleo residual .....	74
<b>Figura 36</b> - Estimativas das amostras autêntica de biodiesel/diesel de Crambe para o conjunto de adulteração com óleo automotivo .....	75

<b>Figura 37</b> – Estimativas das amostras autênticas de biodiesel/diesel de Crambe para o conjunto de adulteração com gasolina .....	76
<b>Figura 38</b> – ROC das amostras de biodiesel/diesel de Crambe para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com óleo de Soja .....	79
<b>Figura 39</b> – ROC das amostras de biodiesel/diesel de Crambe para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com óleo residual .....	80
<b>Figura 40</b> – ROC das amostras de biodiesel/diesel de Crambe para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com óleo automotivo usado .....	81
<b>Figura 41</b> – ROC das amostras de biodiesel/diesel de Crambe para as classes Adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com Gasolina .....	82
<b>Figura 42</b> – VIP para PLS-DA das amostras de Crambe para as classes adulteradas (classe 1) e autênticas (classe 2) para os conjuntos de adulteração com (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina .....	83

## LISTA DE TABELAS

<b>Tabela 1</b> – Regiões espectrais do infravermelho .....	31
<b>Tabela 2</b> – (a) Tabela de confusão para os resultados do modelo PLS-DA para teor de biodiesel metílico de Pinhão Manso.....	50
(b) Tabela de confusão para os resultados do modelo PLS-DA para teor de biodiesel metílico de Crambe.....	51
<b>Tabela 3</b> – Parâmetros de classificação obtidos pelos modelos PLS-DA para o teor de biodiesel de Pinhão Manso e Crambe na mistura BX.....	53
<b>Tabela 4</b> – Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Pinhão Manso com óleo de Soja e óleo residual.....	64
<b>Tabela 5</b> – Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Pinhão Manso com óleo automotivo usado e gasolina.....	64
<b>Tabela 6</b> – Parâmetros de classificação obtidos pelo modelo PLS-DA para adulteração de biodiesel metílico de Pinhão Manso adulterado com (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo e (d) gasolina.....	65
<b>Tabela 7</b> – Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Crambe com (a) óleo de Soja e (b) óleo residual .....	77
<b>Tabela 8</b> – Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Crambe com (a) óleo automotivo usado e (b) gasolina .	77
<b>Tabela 9</b> – Parâmetros de classificação obtidos pelo modelo PLS-DA para adulteração de biodiesel metílico de Crambe adulterado com (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo e (d) gasolina.....	78

## LISTA DE ABREVIATURAS E SIGLAS

<b>ABNT</b>	Associação Brasileira de Normas Técnicas
<b>AGL</b>	Ácidos Graxos Livres
<b>ANP</b>	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
<b>ATR</b>	Reflectância Total Atenuada, do inglês <i>Attenuated Total Reflectance</i>
<b>ASTM</b>	American Society for Testing and Materials
<b>AUC</b>	Área sob a curva (do inglês <i>Area Underthe</i> )
<b>ESI-MS</b>	Espectrometria de massas com Ionização Elestrospray
<b>EN</b>	Norma Europeia
<b>FN</b>	Falso Negativo
<b>FP</b>	Falso Positivo
<b>FT-MIR</b>	Infravermelho Médio com a Transformada de Fourier (do inglês, <i>Fourier Transform Mid Infrared</i> )
<b>HPLC</b>	Cromatografia Líquida de Alta Eficiência
<b>KNN</b>	K-ésimo Vizinhos mais Próximos (do inglês <i>K-Nearest Neighbors</i> )
<b>MME</b>	Ministério de Minas e Energia
<b>NAS</b>	Sinal Analítico Líquido (do inglês <i>Net Analyte Signal</i> )
<b>PCA</b>	Análise de Componentes Principais (do inglês <i>Principal Components Analysis</i> )
<b>PLS</b>	Quadrados Mínimos Parciais (do inglês <i>Partial Least Squares</i> )
<b>PLS-DA</b>	Análise Discriminante por Quadrados Mínimos Parciais (do inglês <i>Partial Least Squares – Discriminant Analysis</i> )
<b>PMQC</b>	Programa de Monitoramento da Qualidade dos Combustíveis
<b>PNPB</b>	Programa de Nacional de Produção e Uso do Biodiesel
<b>RMSEC</b>	Erro Quadrático Médio de Calibração (do inglês <i>Root Mean Square Error of Calibration</i> )
<b>RMSEP</b>	Erro Quadrático Médio de Previsão (do inglês <i>Root Mean Square Error of Prediction</i> )

<b>RMSEV</b>	Erro Quadrático Médio de Validação Cruzada (do inglês <i>Root Mean Square Error of Cross-Validation</i> )
<b>ROC</b>	Característica de Operação do Receptor (do inglês <i>Receiver Operating Characteristics</i> )
<b>SABESP</b>	Companhia de Saneamento Básico do Estado de São Paulo
<b>SIMCA</b>	Modelagem Independente Flexível por Analogia de Classe (do inglês <i>Soft Independent Modelling of Class Analogy</i> )
<b>SVM</b>	Máquinas de Vetores de Suporte (do inglês <i>Support Vector Machines</i> )
<b>VIP</b>	Importância da Variável na Projeção (do inglês <i>Variable Importance in Projection</i> )
<b>VL</b>	Variável Latente
<b>VN</b>	Verdadeiro Negativo
<b>VP</b>	Verdadeiro Positivo

# SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	<b>17</b>
<b>1.1 Revisão Bibliográfica</b> .....	<b>19</b>
1.1.1 Diesel e biodiesel .....	19
1.1.2 Teor, adulteração e não conformidades do diesel .....	28
1.1.3 Espectrometria no Infravermelho Médio .....	30
1.1.4 Quimiometria .....	33
1.1.5 Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) .....	34
<i>1.1.5.1 Parâmetros de avaliação do modelo PLS-DA</i> .....	36
<i>1.1.5.2 Representação gráfica do modelo PLS-DA</i> .....	37
1.1.6 Aplicações de Espectroscopia no Infravermelho aliada a métodos quimiométricos para a análise de biodiesel/diesel .....	38
<b>2. OBJETIVOS</b> .....	<b>41</b>
<b>2.1 Objetivo Geral</b> .....	<b>41</b>
<b>2.2 Objetivos Específicos</b> .....	<b>41</b>
<b>3. PROCEDIMENTO EXPERIMENTAL</b> .....	<b>42</b>
3.1 Amostras de dieseis e biodieseis .....	42
3.2 Preparo de amostras de misturas biodiesel/diesel .....	44
3.3 Obtenção dos espectros MIR .....	44
3.4 Construção dos modelos PLS-DA .....	44
<b>4. RESULTADOS E DISCUSSÃO</b> .....	<b>46</b>
<b>4.1 Caracterizado dos espectros MIR</b> .....	<b>46</b>
<b>4.2 Construção do modelo PLS-DA</b> .....	<b>47</b>
4.2.1 Teor de biodiesel de Pinhão Manso e Crambe .....	47
4.2.2 Adulteração da mistura biodiesel de Pinhão Manso/diesel usando (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina .....	58
4.2.3 Adulteração da mistura biodiesel de Crambe/diesel usando (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina .....	71
<b>5. CONCLUSÕES</b> .....	<b>84</b>
<b>REFERÊNCIAS</b> .....	<b>85-91</b>

## 1. INTRODUÇÃO

O biodiesel é um combustível de origem renovável, produzido a partir de diferentes tipos de oleaginosas e gorduras e inserido na matriz energética brasileira desde 2005. É utilizado em misturas com o diesel e, ao longo dos anos, foi assumindo diferentes proporções entre biodiesel e diesel na mistura comercializada.

Dentre as variadas oleaginosas utilizadas na produção de biodiesel, a Soja aparece em primeiro lugar, seguido do sebo bovino e algodão. No entanto, outras oleaginosas apresentam vantagens em relação a estas matérias primas, tais como: alto teor de óleos, facilidade no plantio e colheita, adaptabilidade em diferentes tipos de solos e climas, não pertencerem a cadeia alimentícia, dentre outras. Diante das novas alternativas em relação às oleaginosas utilizadas na produção de biodiesel, pode-se destacar o Crambe (*Crambe abyssinica Hochst*) e Pinhão Manso (*Jatropha curcas L*).

Para a comercialização deste combustível, deve-se atender uma série de especificações estabelecidas pela Agência Nacional de Petróleo, Gás Natural e Biocombustível (ANP) com o intuito de garantir sua qualidade. Com relação ao teor de biodiesel no diesel é obrigatória a adição de 10% de biodiesel em todo diesel comercializado no país desde março/2018 (CNPE, 2017). No entanto, essa mistura pode sofrer adulterações por outros combustíveis, óleos vegetais, solventes e até mesmo óleos residuais, visando obter lucros indevidos no caso das adulterações uma vez que, os óleos adicionados são mais baratos e possuem propriedades físico-químicas semelhantes tanto ao diesel quanto ao biodiesel. Além disso, durante o processo de produção do biodiesel pode ocorrer a presença de resíduos indesejáveis e também algum tipo de contaminação durante o transporte e armazenamento. Essas adulterações e contaminações causam mudanças nas propriedades do combustível podendo danificar os motores dos automóveis causando prejuízos aos consumidores.

Ao longo de toda a cadeia produtiva do biodiesel até chegar ao consumidor, uma série de análises físico-químicas são realizadas visando comprovar a conformidade das amostras analisadas aos parâmetros de qualidade estabelecidos pela ANP. A realização dos ensaios pode levar tempo e gerar resíduos, utilizando técnicas diferentes.

Portanto, a busca por métodos analíticos alternativos capazes de suprirem essa demanda, principalmente com capacidade de amostragem direta e que minimizem a geração de resíduos são alvos de diversos grupos de pesquisas.

Diante disto, este trabalho teve como objetivo desenvolver um método usando a técnica de Espectroscopia no Infravermelho Médio com a Transformada de Fourier (FT-MIR) aliada ao método quimiométrico de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) capaz de classificar amostras de misturas de biodiesel de Pinhão Manso e Crambe em misturas com diesel visando o controle de qualidade em relação ao teor de biodiesel e possíveis adulterações que possa ser implementado em análises de rotina e controle de qualidade desses combustíveis.

O presente trabalho foi dividido em quatro partes. A primeira, apresentará uma revisão bibliográfica sobre o diesel e o biodiesel enfatizando o processo de produção do biodiesel, as matérias primas utilizadas bem como os casos de adulteração e não conformidades do diesel comercializado. Além disso, será descrita a aplicação da Espectroscopia no Infravermelho Médio e da PLS-DA no monitoramento da qualidade de combustíveis. Na segunda parte serão enfatizados os objetivos do trabalho seguidos do procedimento experimental (terceira parte) com a descrição sobre a produção de biodiesel, preparo de amostras, aquisição dos espectros e a construção dos modelos PLS-DA. Finalmente, nos resultados e discussão serão analisados os dados obtidos em cada um dos modelos de controle de qualidade de misturas de biodiesel/diesel em que o biodiesel foi produzido a partir do óleo de Pinhão Manso e Crambe. Serão detalhadas também as potencialidades da Espectroscopia MIR aliada ao PLS-DA.

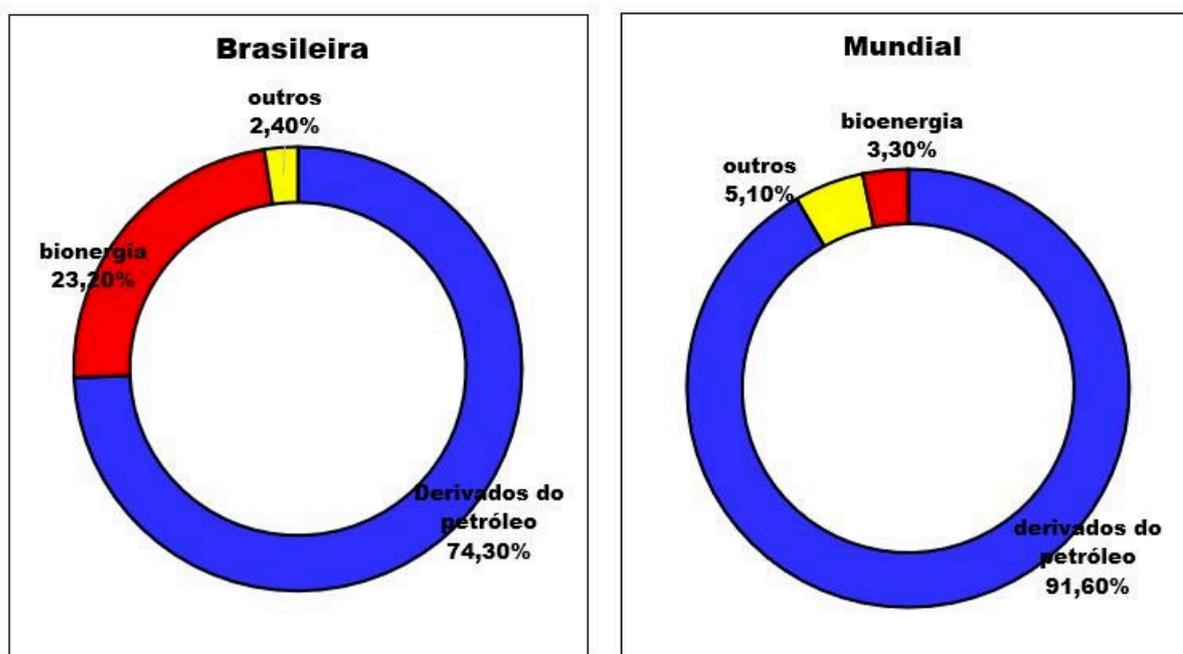
## 1.1 Revisão Bibliográfica

### 1.1.1 Diesel e biodiesel

Desde a Revolução Industrial, a competitividade econômica dos países e a qualidade de vida de seus cidadãos são intensamente influenciadas pela energia. Em um mercado global e em face das crescentes preocupações com o meio ambiente, essa influência se mostra cada vez mais decisiva (TOLMASQUIM; GUERREIRO; GORINI, 2007).

Grande parte da energia consumida no mundo provém do petróleo, do carvão e do gás natural. No entanto, essas fontes não são renováveis. Embora maior parte de energia consumida nos meios de transporte tanto no Brasil como no mundo ainda é de fontes de origem fósseis, observa-se que o Brasil consome energia renovável em número muito maior em relação ao restante do mundo (Figura 1). De acordo com dados do Ministério de Minas e Energia (MME) em 2018, a participação de etanol e biodiesel na matriz energética brasileira ficou em torno de 23,2%, enquanto que a nível mundial, a participação foi apenas de 3,4% de combustíveis de origem renovável (MME, 2018).

**Figura1.** Matriz energética de transportes no Brasil e no mundo



Fonte: Ministério de Minas e Energia (MME), 2018.

Dentre os combustíveis derivados do petróleo, o mais utilizado é o diesel com mais de 44,4% do processo de consumo ficando a gasolina em segundo lugar com aproximadamente 26,6% (Anuário Estatístico- ANP- 2019).

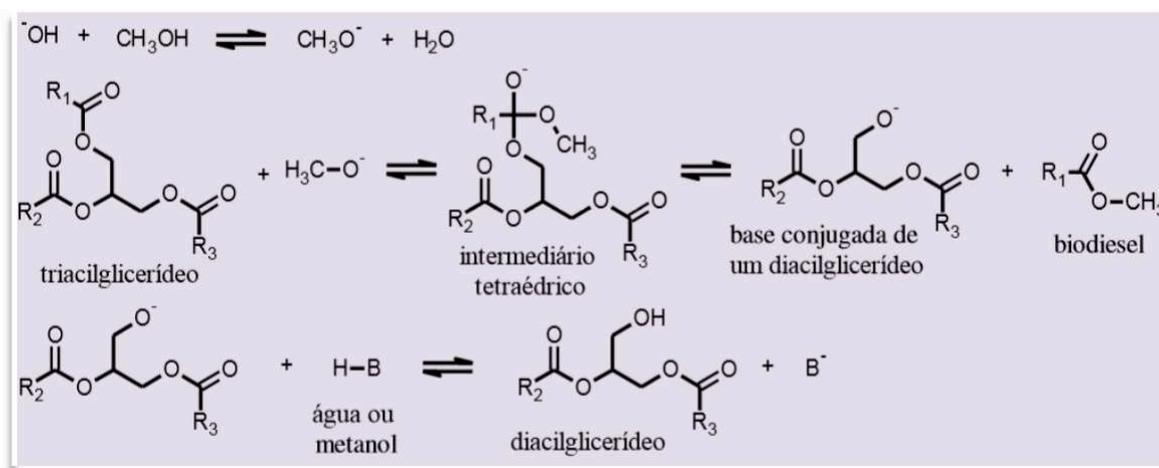
De acordo com a ANP, o óleo diesel é definido como “um combustível líquido derivado do petróleo, composto por hidrocarbonetos com cadeias de 8 a 16 carbonos, e, em menor proporção: nitrogênio, enxofre e oxigênio”. O diesel pode ser usado em transportes aéreos, marítimos e rodoviários, sendo que o último pode ser classificado como tipo A (isento de biodiesel) e tipo B (com adição de biodiesel) (Agência Nacional de Petróleo, Gás Natural e – ANP, 2017). Uma grande parte dos países, inclusive o Brasil, utiliza o diesel tipo B como proposta de diminuição do uso de fontes energéticas provenientes de combustíveis fósseis (LÔBO, FERREIRA, DA CRUZ, 2009).

A ANP define o biodiesel como um “combustível composto por alquil éster de ácidos carboxílicos de cadeia longa, produzido a partir da reação de transesterificação e/ou esterificação de ácidos graxos, de gorduras de origem vegetal ou animal” (ANP, 2014). Este biocombustível apresenta propriedades semelhantes ao diesel mineral podendo substituí-lo em qualquer proporção. Além disso, possui algumas características que representam vantagens sobre os derivados do petróleo como: pequena quantidade de teor de enxofre e de compostos aromáticos, alto número de cetano (qualidade de combustível medida entre a injeção de combustível e o início da combustão), teor médio de oxigênio, maior ponto de fulgor (características de baixa volatilidade dando maior segurança no transporte e armazenamento) (FERRARI, OLIVEIRA, SCABIO, 2005). O aumento no uso do biodiesel leva a uma redução nos gases componentes do efeito estufa, como o CO<sub>2</sub>, levando em consideração o ciclo do carbono durante o crescimento da oleaginosa, matéria particulada e emissões de enxofre (MONTEIRO et al., 2009). Devido ao desenvolvimento tecnológico cada vez mais exigente, torna-se possível uma análise em relação à matéria-prima mais abundante, barata e de propriedades que garantam a qualidade do biodiesel (RAMOS et al., 2017).

O biodiesel pode ser produzido de diferentes maneiras, mas principalmente pela reação de transesterificação e/ou esterificação. A transesterificação em meio básico é a mais utilizada sendo que um triacilglicerídeo com um álcool de cadeia curta produz monoésteres graxos e glicerina como coproduto. O alcóxido (álcool + hidróxido) reage com um carbono do triacilglicerídeo gerando um intermediário que produz um monoéster. Este, por sua vez, gera um diacilglicerídeo que é atacado formando um monoacilglicerídeo que, por fim, produz uma terceira molécula de monoéster e glicerol (Figura 2). O processo de produção de biodiesel com metanol é mais comum uma vez que seu custo no mercado internacional menor em relação ao etanol e possui maior rendimento na reação. Por outro lado, o Brasil possui

tecnologia consolidada de produção do etanol, podendo optar pela produção de biodiesel através da rota etílica (RAMOS et al. 2017).

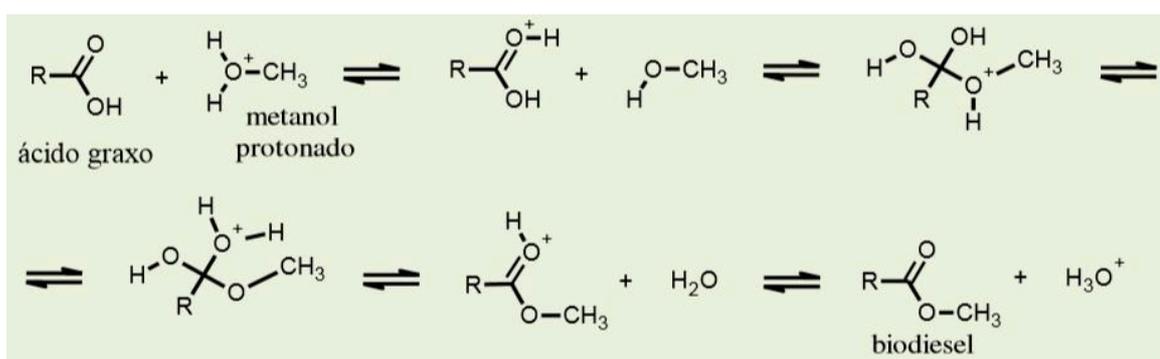
**Figura 2.** Esquema da reação de transesterificação de óleos vegetais.



Fonte: Ramos et al.,2017

No caso da utilização de óleos com alto teor de ácidos graxos livres (AGL), superior a 1%, a transesterificação em meio básico não é o ideal. Para este tipo de matéria prima é necessário o uso de catalisador ácido para evitar a formação de sabão sendo necessária a esterificação. Neste tipo de reação são adicionados ácidos como, por exemplo, H<sub>2</sub>SO<sub>4</sub> para esterificar os ácidos graxos livres e os ésteres metílicos, como demonstra o esquema na Figura 3.

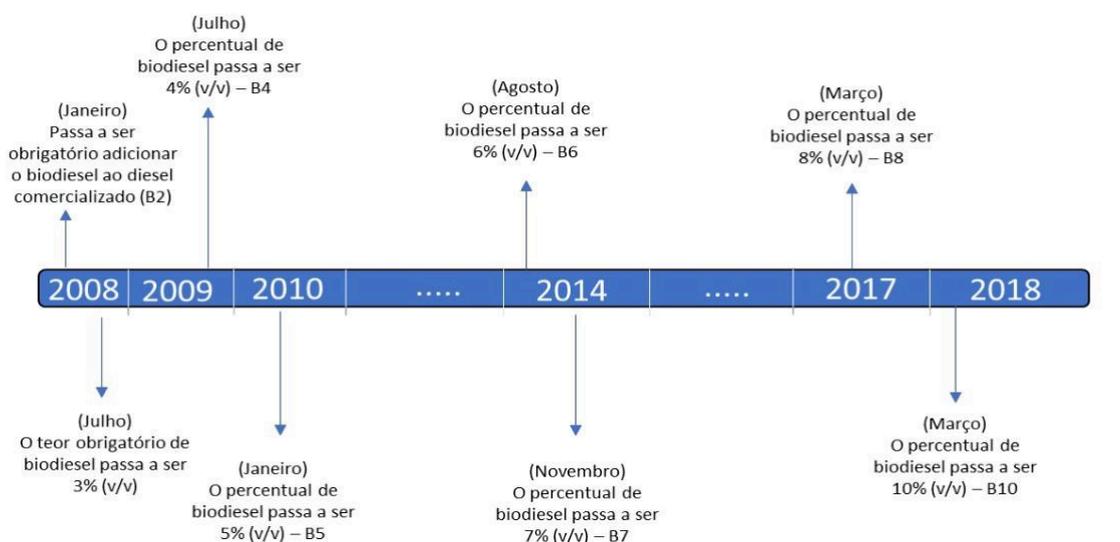
**Figura 3.** Esquema da reação de esterificação de óleos vegetais.



Fonte: Ramos et al., 2017.

O glicerol é um coproduto da reação de produção do biodiesel e devido às características relacionadas à não toxicidade e ausência de cor e odor, tem uma ampla aplicação sendo utilizado em vários setores da indústria como: cosméticos, saboaria, farmacêutico, alimentício, entre outros. É chamado de glicerina os produtos comerciais com aproximadamente 95% de glicerol, mas atualmente, com o aumento da produção de biodiesel em todo o mundo, uma fração com cerca de 80% de glicerol é normalmente denominada de glicerina loira ou bruta (MOTA; da SILVA; VASCONCELOS, 2012). No entanto, de acordo com a ANP, a quantidade de glicerina produzida é superior ao consumo. Uma das alternativas usada para o consumo desse excesso de glicerina é a exportação. Em 2014, o Brasil exportou 211 mil toneladas de glicerina (PEITER et al., 2016). Além disso, a glicerina purificada possui várias outras utilidades como: na produção de medicamentos e maquiagem (PAGLIARO et al., 2007; SINGHABHANDHU, 2010), na indústria alimentícia (SANTOS, 2009), têxtil (KNOTHE et al., 2006; SINGHABHANDHU, 2010), fabricação de explosivos (PAGLIARO et al., 2007), na produção de ração animal (MENTEN et al., 2008), dentre outras utilidades.

De acordo com a Lei 11.097 de Janeiro de 2005, o biodiesel foi inserido na matriz energética brasileira na forma de B2 (2% de biodiesel no diesel) e passou a ser obrigatório a partir de 2008 (ANP, 2008). Devido ao aumento do consumo energético, consequência do desenvolvimento tecnológico, demográfico e dependência de derivados do petróleo intensificou-se a busca por fontes alternativas de energia, principalmente renováveis (KNOTHE; GERPEN; KRAHL, 2006, MITTELBACH; REMSCHMIDT, 2005). Assim, a porcentagem de biodiesel no diesel foi aumentando gradativamente como mostra a Figura 4. Desde março/2018, o percentual obrigatório do biodiesel é de 10%, mistura essa denominada de B10 (ANP, 2018a).

**Figura 4.** Evolução do percentual de teor de biodiesel presente no diesel fóssil no Brasil

**Fonte:** ANP, 2018a

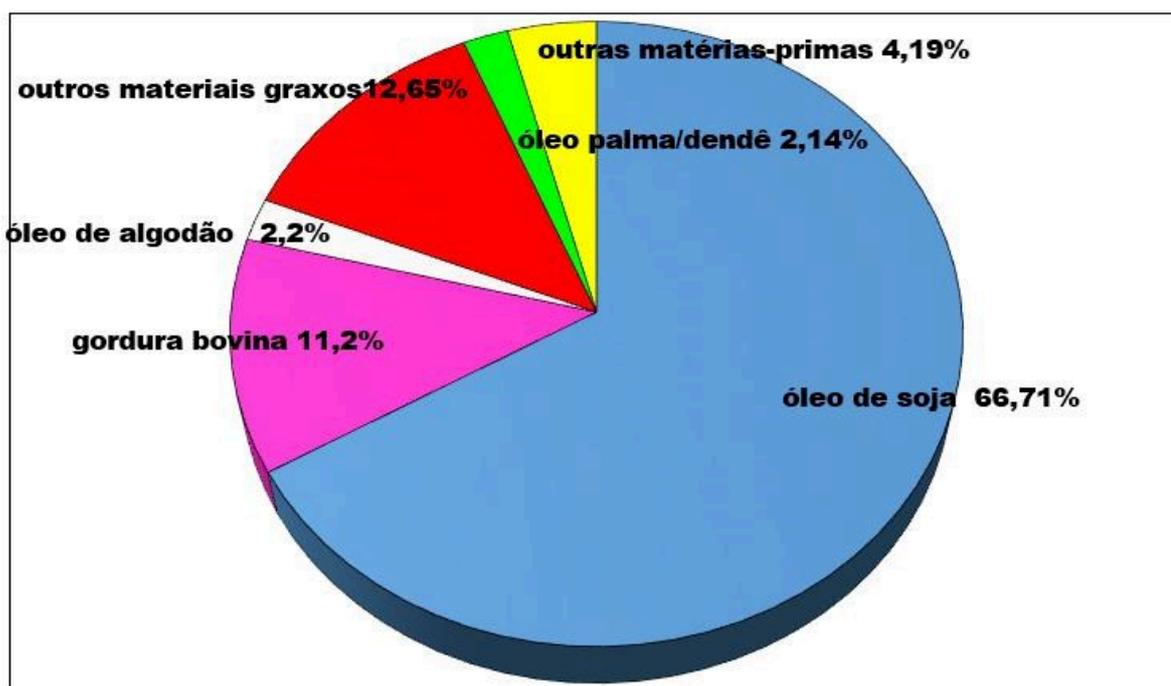
Atualmente, o Brasil tem um consumo de diesel superior à sua produção, levando à importação deste combustível e gerando um déficit na balança comercial no país para este produto. Em 2013, esse déficit era de 9,919 mil m<sup>3</sup>, mas vem aumentando a cada ano; em 2017, era de 12,454 mil m<sup>3</sup>, gerando uma dependência externa de 24,7%, isto é, a representatividade das importações sobre o volume total transacionado do produto no território nacional (ANP, 2018c). Além deste problema comercial, o óleo diesel é um combustível de origem não renovável derivado de petróleo, cujas reservas são limitadas. Apesar desta limitação, não há um consenso em relação ao tamanho das reservas, uma vez que esta estimativa depende da descoberta de novas reservas, produtividade nos poços e consumo mundial (SORRELL et al., 2012). Em 2017, as reservas comprovadas em todo o mundo eram de 1,7 trilhões de barris e o crescente consumo mundial totalizou 98,2 milhões de barris/dia (ANP, 2018a).

Portanto, o uso do biodiesel é uma alternativa promissora para a redução do consumo do diesel. Além de reduzir a dependência em relação ao petróleo, a produção deste biocombustível fortalece o agronegócio e cria um mercado para óleos vegetais e gorduras animais (POUSA; SANTOS; SOARES, 2007). Também se verifica uma redução no êxodo rural e geração de empregos no campo, uma vez que empresas com o Selo de Combustível Social tem preferência em leilões da ANP. Este selo é dado às companhias cuja parte da matéria prima vem da agricultura familiar, além das mesmas prestarem assistência técnica a estes agricultores (PADULA et al., 2012). Além das vantagens já citadas, por ser derivado da biomassa, o biodiesel é renovável considerado menos poluente do que o óleo diesel. Também

apresenta vantagens ecológicas pela diminuição dos gases de combustão emitidos dos motores devido uma pequena quantidade de óxidos de enxofre, vantagens financeiras devido ao mercado de “bônus de carbono”, diversificação da matriz energética e consequente diminuição do petróleo (BALAT; BALAT, 2009).

No levantamento do perfil nacional das matérias primas feito pela ANP em setembro de 2019 (Figura 5), o óleo de Soja continua como a principal matéria prima para a produção de biodiesel com 66,7% da produção nacional, seguida pela gordura bovina com 11,12% e óleo de algodão com 2,20% , outros materiais graxos com 12,65% , óleo de palma 2,14% e outras matérias primas como: óleo de fritura, óleo de canola, óleo de milho, gordura de porco e gordura de frango com 4,19%. Porém essas porcentagens podem se diferenciar em relação a cada região de acordo com as características e o processo de produção das oleaginosas consolidada. Na região SUDESTE, por exemplo, a Soja variou sua porcentagem entre 10 e 41% entre os meses de janeiro a setembro de 2019 e o óleo de fritura entre 10 e 21% e a gordura bovina entre 20 e 37% sendo o estado que mais reaproveita resíduos para a produção de biodiesel (ANP, 2019).

**Figura 5.** Matérias primas utilizadas para a produção de biodiesel no Brasil



Fonte: ANP, 2019.

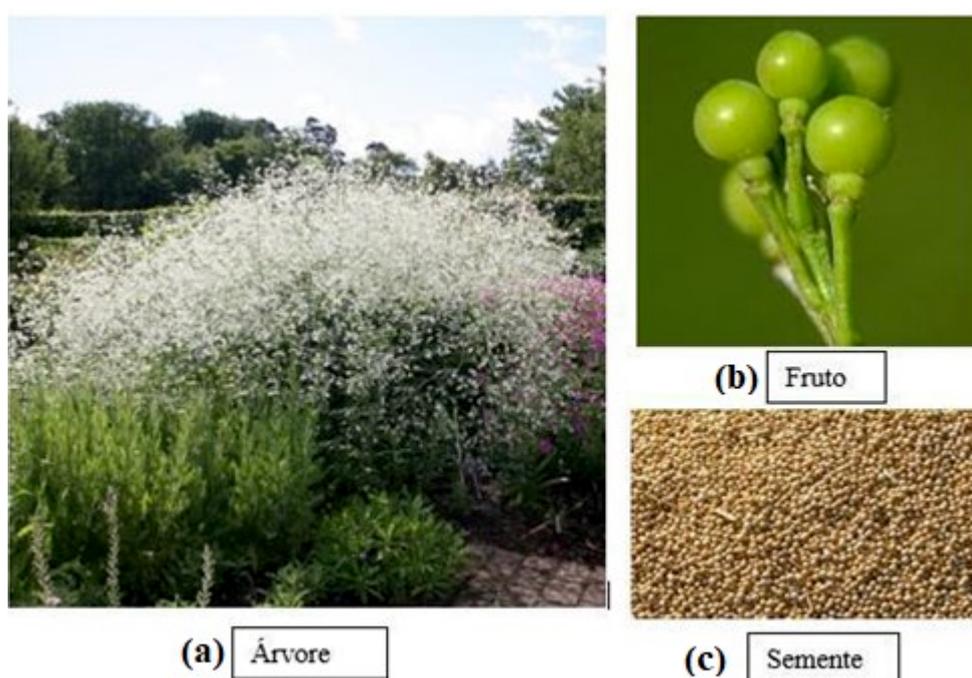
Em relação ao mercado mundial, em países como Brasil, Argentina, Indonésia e outros países em desenvolvimento, espera-se a expansão devido às políticas favoráveis nesse setor. No entanto, na União Europeia com a aprovação de um plano de ação climática e energética aumentou a participação de fontes renováveis na sua matriz energética e os EUA tornou o maior exportador mundial do biocombustível. No Brasil, em virtude do Programa de Nacional de Produção e Uso do Biodiesel (PNPB), iniciado em 2005, já foram produzidos, até dezembro de 2018, mais de 34,7 bilhões de litros deste biocombustível. Comparativamente, o Brasil vem mantendo sua posição de segundo maior produtor e consumidor de biodiesel no ranking internacional, antecedido pelos EUA, e sucedido pela Alemanha e Argentina (MME, 2018). O uso de óleos comestíveis e refinados pode elevar o custo de produção do biodiesel ou a uma competitividade aumentando os preços desses alimentos (KARPAGAM, et al., 2015).

Assim, vários estudos e pesquisas buscam a viabilidade do uso de oleaginosas alternativas e de preferência que não sejam usadas como alimento humano. O Brasil é um país com características de solo e clima favoráveis para a maioria das oleaginosas que vem sendo adaptadas em diferentes regiões do país. Dentre elas podemos destacar a macaúba, milho, algodão, Crambe, Pinhão Manso, dendê, baru, babaçu, pequi, moringa oleífera, dentre outras (QUINTELLA et al., 2009). Outras matérias primas como as microalgas estão apresentando resultados promissores para a produção de biodiesel devido à maior produtividade em óleos, ciclo de vida em poucos dias, além de não competir com a indústria alimentícia. Por outro lado, as microalgas possuem um alto custo no cultivo, na secagem no material utilizado, alta acidez e necessidade de grande quantidade na colheita (ATABANI, et al., 2013, AHMAD et al., 2011). Também deve-se destacar a importância do uso de óleos residuais oriundos de processos de frituras de alimentos, cuja destinação remedia uma importante fonte de contaminação ambiental (GUI; LEE; BHATIA, 2008). Estudos indicam que cada litro jogado no esgoto contamina 25 mil litros de água podendo aumentar 40% de custo de tratamento de efluentes devido à presença de óleo residual (COSTA NETO et al., 2000; SABESP, 2010).

Dentre as oleaginosas não comestíveis destaca-se o Crambe e o Pinhão Manso. O Crambe (*Crambe abyssinica* Hochst) (Figura 6) é uma oleaginosa que apresenta potencial como matéria prima para a produção de óleo com fins energéticos, devido sua tolerância à seca, produtividade e alto teor de óleo, apresentando uma alternativa para a safrinha (plantio que ocorre após a colheita de Soja) (ONOREVOLI et al., 2014). Trata-se de uma espécie nativa da região Mediterrânea, da Etiópia e Tanzânia cujas sementes contêm de 35,6 – 42,8% de óleo, e seu plantio também pode ocorrer em áreas marginais não-aptas para as tradicionais culturas de alimentos, evitando a competição com este setor (FALASCA et al., 2010). Esta

planta foi introduzida no Brasil há pouco tempo para produção de sementes. O processo de secagem das sementes é um dos fatores que afetam sua qualidade. A diminuição do teor de água das sementes em níveis adequados permite preservá-las por um período prolongado, possibilitando menor perda de vigor e germinação ao longo do armazenamento. Com relação ao biodiesel produzido, Rosa e colaboradores (2014) verificaram que o consumo específico e a eficiência de conversão energética eram estatisticamente iguais ao do diesel, além de ter uma redução significativa da emissão dos gases CO, NO, NO<sub>x</sub> e SO<sub>2</sub>. WAZILEWSKI et al., 2013, analisaram este biodiesel e concluíram que este apresenta uma maior estabilidade ao estresse térmico e contaminação como bronze ou ferro do que o biodiesel produzido a partir da Soja.

**Figura 6.** Partes da planta de Crambe: (a) árvore; (b) fruto e (c) semente

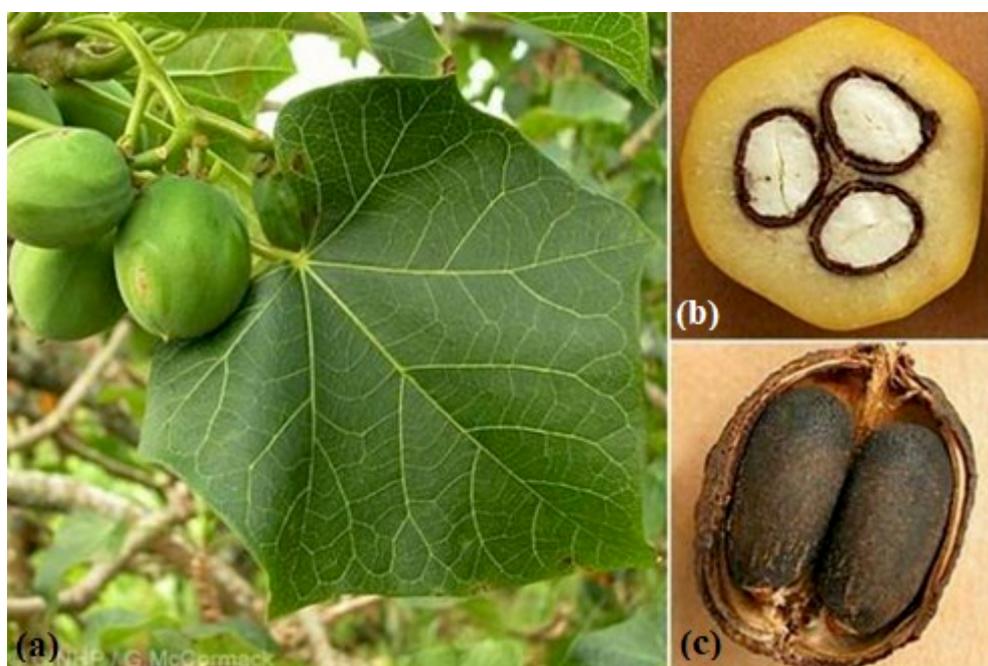


**Fonte:** O autor, 2018.

O Pinhão Manso (*Jatropha curcas* L.) (Figura 7) foi muito estudado devido ao alto rendimento de óleo (43-59%) e baixo custo de produção. Essa última característica é devido a sua rusticidade, o que permite seu cultivo em regiões com diferentes regimes de chuva (944 a 3121 mm/ano) sem necessidade de irrigação, além de crescer em solos de baixa fertilidade, reduzindo sua competição por terra agrícola e por se prestar de forma completa para a agricultura familiar. Estes fatores fazem com que seu custo seja reduzido e uma alternativa para a produção de biodiesel (LAVIOLA et al., 2017). Trata-se de uma árvore nativa do México, América Central, América do Sul, Índia e África, cuja altura pode chegar até 7 m.

Ela produz sementes após um ano de cultivo e pode viver de 30-50 anos (ATIBANI et al., 2013). Devido ao elevado teor de ácidos graxos saturados e um baixo teor de ácidos graxos insaturados, o biodiesel produzido a partir do Pinhão Manso apresenta boa estabilidade à oxidação e excelente qualidade (FATEMI; HAMMOND, 1980; KNOTHE et al., 2006), pois, possui boa estabilidade oxidativa facilitada pela presença de ácidos graxos monoinsaturados na mistura de biodiesel, além de melhoria do desempenho a baixas temperaturas (LAVIOLA et al., 2017). No entanto, esses processos de descasque e colheita são feitos, na maioria dos casos, manualmente, o que pode tornar o Pinhão Manso uma oleaginosa não competitiva. Uma das alternativas seria o processo de mecanização do plantio, à colheita, um aumento no rendimento do óleo e melhoria na sustentabilidade da indústria, podendo suprir à crescente demanda mundial (LIM et al., 2015). Além disso, o Pinhão Manso possui uma maturação dos frutos de forma irregular dificultando a colheita. Novas estratégias estão sendo pesquisadas para o melhoramento genético e biotecnológico associados à seleção de plantas para o desenvolvimento de uma variedade comercial (LAVIOLA et al., 2017).

**Figura 7.** Partes da planta de Pinhão Manso: (a) árvore; (b) corte longitudinal do fruto e (c) fruto aberto



**Fonte:** O autor, 2018.

Biodieseis produzidos de diferentes fontes de ácidos graxos e álcoois apresentam características diferentes. Neste sentido, Pinho e colaboradores (2014) demonstraram que biodiesel produzido de diferentes oleaginosas podem afetar no rendimento do teor de biodiesel e dependendo da oleaginosa podem ser adotadas políticas diferentes relacionadas aos biodieseis produzidos de óleos comestíveis e não comestíveis devido à competição com a indústria de alimentos (PINHO et al., 2014).

Chatarino e colaboradores (2007) distinguiram biodieseis tanto em relação ao tipo de oleaginosa quanto ao tipo de álcool usando Espectrometria de massas com Ionização Elestrospray (ESI-MS) (CHATARINO et al., 2007).

César e colaboradores (2015) estudaram vários aspectos como extração, economia, propriedades físico-químicas do óleo de macaúba para a produção de biodiesel. As perspectivas são boas não somente devido ao alto teor de óleo, mas pela baixa viscosidade. O óleo do mesocarpo possui uma grande quantidade de ácido láurico facilitando a reação de esterificação, pois o mesmo possui cadeia curta (constituída por doze átomos de carbono) sem insaturações e baixo ponto de fusão resultando num biodiesel de alta qualidade (CÉSAR et al., 2015).

### **1.1.2. Teor, adulteração e não conformidades do diesel**

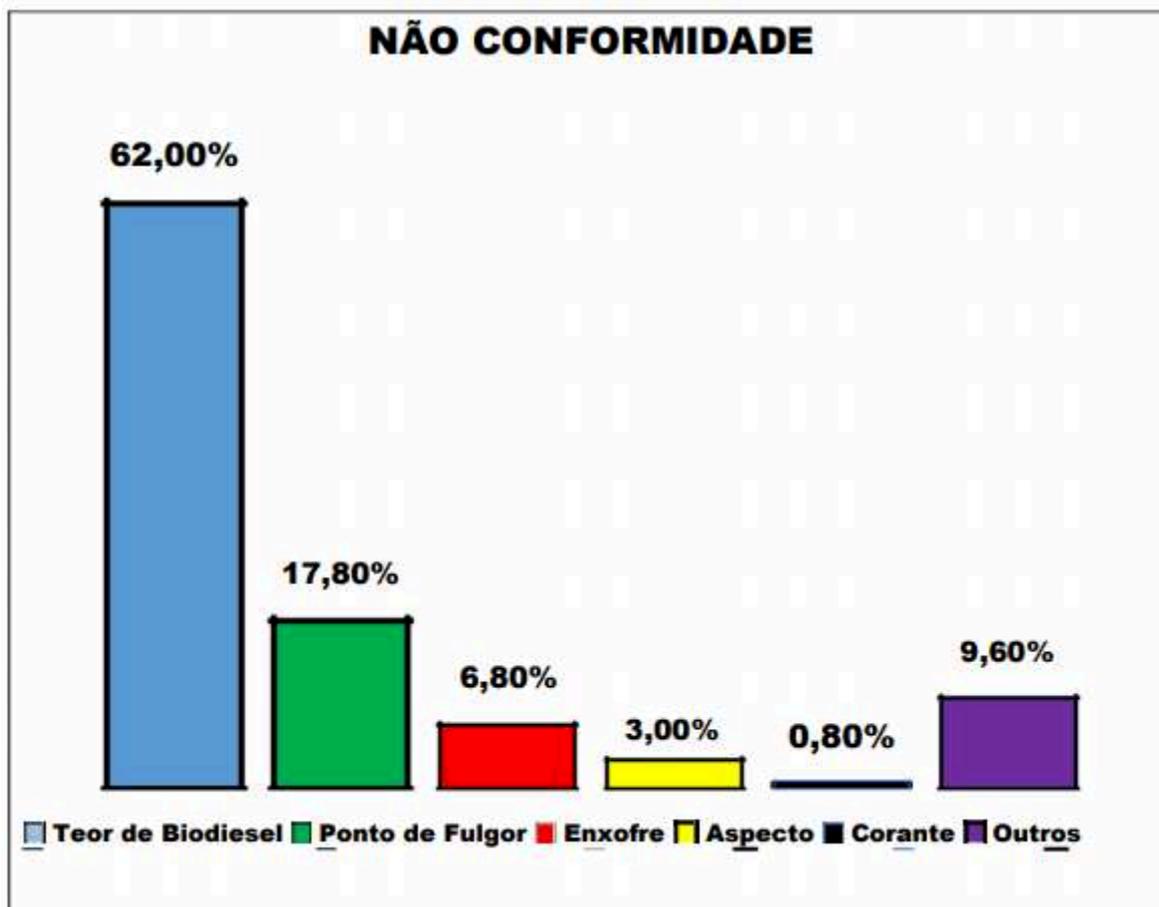
O óleo diesel deve atender a uma série de especificações de qualidade determinados pela ANP, para ser comercializado, além da obrigatoriedade de adição de 10% (v/v) de biodiesel no diesel. Os padrões de qualidade estão relacionados ao processo de produção do biodiesel e a escolha da matéria-prima (ANP, 2014). Uma das vantagens do uso das misturas biodiesel/diesel é o aproveitamento da rede de distribuição de combustível já existente, e esta mistura pode ser usada nos motores atuais sem qualquer modificação em uma porcentagem de até 20,0% (v/v) (MOFIJUR et al., 2013).

O Programa de Monitoramento de Qualidade de Biocombustíveis (PMQC) é o instrumento utilizado pela ANP para verificar a qualidade dos principais combustíveis líquidos comercializados no Brasil identificando a existência de produtos que não atendem às especificações e planejam ações de fiscalização do abastecimento.

Em 2018, partir de uma das análises de monitoramento feita pela PMQC, foram avaliadas 33.513 amostras de diesel comercial e dessas 1.861 estavam não conformes. De acordo com a Figura 8, 62,1% das não conformidades eram relativas ao teor de biodiesel (verificação do cumprimento ao dispositivo legal que determina a adição de biodiesel ao óleo diesel); 9,6% a cor ASTM (cor ASTM fora de especificação pode ser indicativo de degradação ou contaminação) e massa específica a 20 °C; 17,8% a ponto de fulgor; 6,8%

concentração de enxofre no combustível; 0,8% corante e 3% aspecto (indicação visual de qualidade e de possíveis contaminações (ANUÁRIO ESTATÍSTICO – ANP- 2019). ades de óleo diesel.

**Figura 8.** Distribuição percentual das não conformidades do óleo diesel.



Fonte: ANP, 2019

As não conformidades com relação ao teor de biodiesel não são a única forma de adulteração. A adição de óleos vegetais, residuais, óleo automotivo usado ou até mesmo solvente como adulterantes no óleo diesel provoca depósitos de carbono, bloqueio de injeção e combustão incompleta, devido a sua alta viscosidade e baixa volatilidade (LEUNG; KOO; GUO; 2006). Podem ocorrer alterações nas propriedades desse combustível durante o processo de transporte e armazenamento ou até mesmo através de adição de óleos vegetais e residuais no próprio posto, sendo difícil a identificação dessas adulterações ou contaminações devido às propriedades semelhantes desses óleos com o diesel (SOARES et al., 2008).

O tema “adulteração de combustíveis” vem despertando crescente interesse por parte de pesquisadores que a cada dia apresentam na literatura científica relevante contribuições como as citadas a seguir.

Kumar e Mishra 2012 quantificaram etanol em gasolina, biodiesel e misturas de biodiesel/diesel por espectrofluorimetria e calibração multivariada por regressão por componentes principais (PCR) e PLS. Os melhores valores de RMSEC (Erro Quadrático Médio de Calibração) foram de 0,79 e RMSEP (Erro Quadrático Médio de Previsão) de 0,52 para os modelos PLS.

Brandão e colaboradores (2012) em um dos seus trabalhos quantificaram óleos vegetais em diesel através da cromatografia líquida de alta eficiência aliada a métodos quimiométricos como: a análise de componentes principais (PCA), método do K-ésimo vizinho mais próximo (KNN) e PLS, sendo calculada a precisão média e a concentração mínima detectada. A cromatografia líquida de alta eficiência (HPLC) aliada a métodos quimiométricos mostrou grande potencial e confiabilidade para a determinação de óleos e gorduras usados presentes nas amostras analisadas de diesel.

Nas análises feitas pela ANP, as amostras são coletadas e levadas ao laboratório demandando tempo. Portanto, se faz necessário o uso de métodos de análises rápidos e eficientes em que os resultados sejam em tempo real. A espectroscopia no infravermelho é uma técnica rápida e de amostragem direta quando usada a reflectância total atenuada (ATR), que necessita de um pequeno volume de amostra (menos que 1,00 mL), não requerendo uso de solventes e minimizando a geração de resíduos. Esta técnica é utilizada nos métodos padrões para quantificar biodiesel em misturas com diesel de acordo com a Norma Europeia EN 14078 (2004), Associação Brasileira de Normas Técnicas (ABNT) NBR 15568 (2008) e Sociedade Americana de Testes e Materiais (ASTM) D7371 (2007).

### **1.1.3 Espectroscopia no infravermelho médio**

A espectroscopia estuda a interação da radiação eletromagnética com a matéria, sendo um dos seus principais objetivos o estudo dos níveis de energia de átomos ou moléculas. Normalmente, as transições eletrônicas são situadas na região do ultravioleta ou visível, as vibracionais na região do infravermelho e as rotacionais na região de micro-ondas e, em casos particulares, também na região do infravermelho distante (MASSART et al., 1998).

A região do infravermelho médio é muito utilizada e se estende de cerca de 4.000 a 200  $\text{cm}^{-1}$  em que os espectros são empregados tanto para análise quantitativa como qualitativa. A região do infravermelho próximo, de 12.800 a 4.000  $\text{cm}^{-1}$  é também usada em determinação quantitativa em rotina de certas espécies como água, dióxido de carbono, enxofre e outros compostos utilizados na indústria e agricultura. Já no infravermelho distante, de 200 a 10  $\text{cm}^{-1}$ , o uso tem sido na determinação de estruturas de espécies

inorgânicas e organometálicas, com base em medidas de absorção. A Tabela 1 apresenta os limites aproximados para cada região.

**Tabela 1.** Regiões espectrais no infravermelho.

Região	Número de onda (v), cm <sup>-1</sup>	Comprimento de onda (λ), nm	Frequência (v), Hz
<b>Próximo (NIR)</b>	12.800 a 4000	780 a 2.500	3,8 x 10 <sup>14</sup> a 1,2 x 10 <sup>14</sup>
<b>Médio (MIR)</b>	4.000 a 200	2.500 a 5.000	1,2 x 10 <sup>14</sup> a 6,0 x 10 <sup>14</sup>
<b>Distante (FIR)</b>	200 a 10	5.000 a 100.000	6,0x10 <sup>12</sup> a 3,0 x 10 <sup>11</sup>

Fonte: Skoog, 2002

A região no infravermelho médio (MIR) refere-se a transições vibracionais e rotacionais das ligações das moléculas e se estende de cerca de 4000 a 200 cm<sup>-1</sup> sendo considerada a mais importante para a caracterização de compostos orgânicos (BARBOSA, 2008).

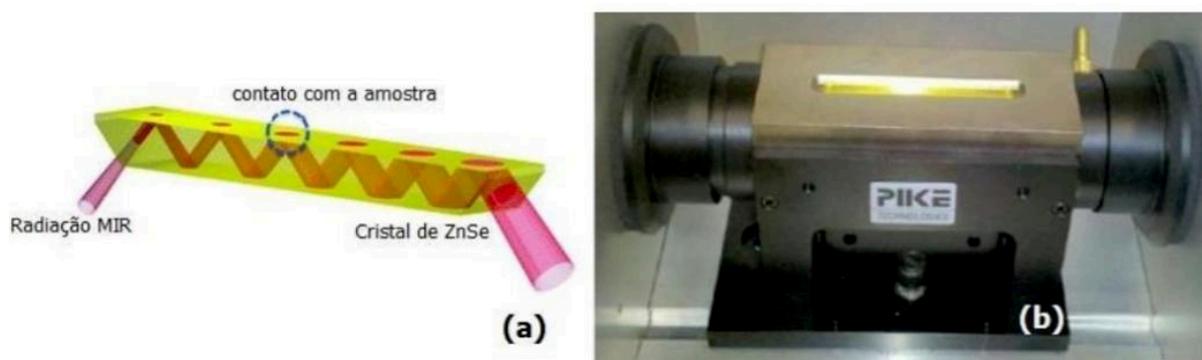
A primeira vantagem da espectroscopia MIR é o fato de se tratar de uma técnica quase universal: é possível analisar moléculas em diversos estados físicos (sólidos, semi-sólidos, líquidos, gases), pós, polímeros, substâncias orgânicas, inorgânicas, materiais biológicos, entre outras. A outra vantagem se refere ao fato que seus espectros são ricos em informação, pois as bandas nessa região fornecem as estruturas das moléculas em uma amostra, enquanto as intensidades estão relacionadas as concentrações. Existem diferentes tipos de reflectância podendo ser baseados na natureza da superfície refletora como: especular, difusa e reflexão total interna. Dentre os diferentes tipos de técnicas de amostragem para a espectroscopia no infravermelho médio a mais utilizada é a reflectância total atenuada (ATR, do inglês *attenuated total reflectance*) (SMITH, 2011).

As técnicas de reflectância apresentam vantagens como não ter o problema de opacidade, como ocorre nos métodos de transmissão. Portanto, não é necessário preparar as amostras como finos filmes, e, portanto, técnicas que usam a reflectância não são destrutivas. Isso faz com que suas análises sejam mais rápidas e fáceis de serem utilizadas. Assim, mesmo que os acessórios de reflectância sejam consideravelmente mais caros, a economia de tempo e dinheiro devido a eliminação da etapa de preparo de amostra cobre o maior custo (SMITH, 2011).

A técnica de ATR permite realizar análise direta em diversos materiais. Em casos de sólidos e semi-sólidos (linhas, tecidos, fibras, pós, pastas, suspensões, polímeros, entre outros), o material é depositado em cima do acessório e pressionado contra o cristal. Já no caso de amostras líquidas, este é adicionado diretamente no acessório ATR, como no caso das amostras de biodiesel/diesel. O cristal mais utilizado nesse tipo de análise é o de seleneto de zinco (ZnSe), pois é insolúvel em água, possui alto índice de refração (0,5 a 15 μm) e é

resistente a choques térmicos (COATES, J., 2006, BARBOSA, 2008). A técnica de ATR baseia-se no fato de que quando um feixe de luz é incidido na superfície da amostra, parte é absorvida e parte refletida, criando uma onda evanescente no seu interior e absorvendo energia, e em seguida sofre uma atenuação do feixe passando pelo detector e gerando o espectro (Figura 9). A desvantagem mais significativa do ATR é a baixa sensibilidade, sendo possível detectar substâncias que estão presentes apenas em concentrações maiores do que 0,1% (BARBOSA, 2008; SMITH, 2011).

**Figura 9.** (a) Reflectância de radiação no interior do cristal de ZnSe; (b) acessório de ATR.



Fonte: Souza, 2014.

Nos espectros de ATR enquanto as mesmas bandas são observadas, suas intensidades relativas diferem. As absorções são independentes da espessura da amostra, uma vez que a radiação penetra apenas alguns micrômetros na mesma (SKOOG; HOLLER; NIEMAN, 2002).

Assim, devido a esta vantagem, a espectroscopia MIR tem aumentado sua aplicabilidade. Apesar dos dados obtidos serem complexos visualmente, difíceis de serem interpretados, com o desenvolvimento de métodos quimiométricos, tornou-se comum o uso de métodos matemáticos e estatísticos multivariados para extrair informações úteis desse tipo de dado.

#### 1.1.4 Quimiometria

Atualmente, os instrumentos analíticos produzem uma grande quantidade de informações químicas (isto é, variáveis ou características) para um grande número de amostras. Assim, devido a necessidade de extrair o máximo de informação útil destes dados houve o surgimento e rápido desenvolvimento da Quimiometria (FERREIRA, 2015).

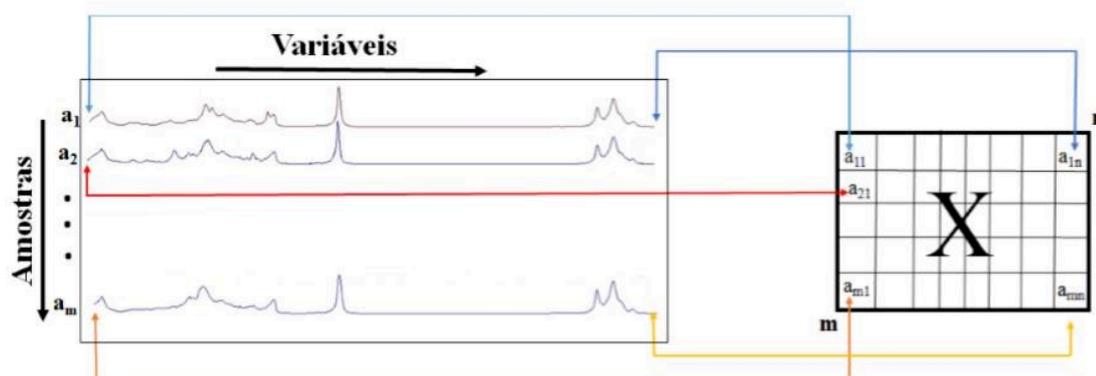
A Quimiometria é a área da química que utiliza métodos matemáticos, estatísticos e de lógica formal em dados de origem química (MASSART et al., 1998). Os métodos

quimiométricos podem ser aplicados em diferentes áreas, sendo as principais: processamento de sinais analíticos, planejamento e otimização de experimentos, conhecimento de padrões e classificação de dados, calibração multivariada, monitoramento e modelagem de processos multivariados, métodos de inteligência artificial, entre outros (VANDEGINSTE et al., 1998).

Neste trabalho foi dado foco na área de classificação de dados, isto é, em que a propriedade de interesse é categórica (também denominada de discreta). Nesse caso, cada uma das amostras é descrita por um conjunto de medidas experimentais (por exemplo, espectros MIR) e são divididas de acordo com a propriedade de interesse (bom/ruim; falso/verdadeiro; ativo/não ativo, etc.) (FERREIRA, 2015).

Antes do desenvolvimento do modelo, é necessário organizar os dados em questão numa matriz  $\mathbf{X}$  ( $n \times m$ ) a partir dos vetores das medidas instrumentais conforme mostra a Figura 11, em que  $n$  refere-se ao número de linhas e  $m$  o número de variáveis. As linhas da matriz referem-se aos espectros das amostras e as colunas representam as variáveis, valores de absorvância nos números de onda (Figura 10).

**Figura 10.** Organização dos dados para a construção da matriz  $\mathbf{X}$ .



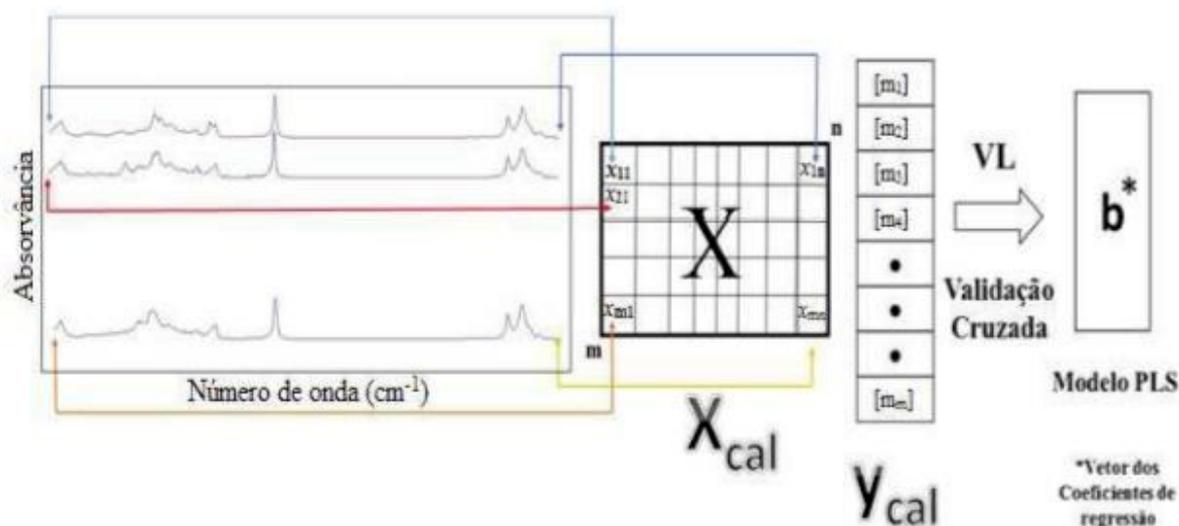
Fonte: Gontijo, 2016.

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

A PLS-DA faz parte dos métodos de classificação supervisionados, isto é, deve-se ter o conhecimento inicial sobre as categorias/classes das amostras que serão modeladas (FERREIRA, 2015). Existem diversos métodos, sendo os mais empregados atualmente a modelagem independente e flexível por analogia de classes (SIMCA, do inglês *soft independent modelling of class analogies*), K-ésimo vizinhos mais próximos (KNN, do inglês *K-nearest neighbor*) e PLS-DA.

A PLS-DA é uma variação do algoritmo de regressão por quadrados mínimos parciais (PLS, do inglês *partial least squares*), um método de calibração inversa que busca relacionar uma resposta instrumental (matriz  $\mathbf{X}$ ) com um vetor  $\mathbf{y}$  com a propriedade de interesse (por exemplo, valores de concentração). A diferença entre PLS-DA e PLS é que no primeiro o vetor  $\mathbf{y}$  conterá valores que representem a classe à qual uma amostra pertence (MASSART et al., 1998; BARKER; RAYENS, 2003). O vetor  $\mathbf{y}$  é constituído por valores inteiros (0 e 1), em que 0 indica que a amostra não pertence à classe de interesse, enquanto 1 refere-se a amostras que pertencem a classe de interesse, conforme é mostrado na Figura 11. Em casos em que há mais de uma classe de interesse para se distinguir simultaneamente, o vetor  $\mathbf{y}$  se torna uma matriz  $\mathbf{Y}$ , e o modelo é denominado de PLS-DA.

**Figura 11.** Representação da matriz de dados  $\mathbf{X}$  e o vetor  $\mathbf{y}$  usados no PLS-DA.



Fonte: Máquina, 2019.

Para a construção dos modelos de classificação supervisionados, são necessários 2 conjuntos de amostras: (i) conjunto de treinamento, o qual contém amostras representativas para a construção do modelo, e (ii) conjunto de teste, no qual novas amostras são utilizadas para testar o modelo, cujas classes já são conhecidas (BARKER, RAYENS, 2003). Embora idealmente os valores previstos pelo modelo PLS-DA sejam 0 e 1, na prática estes valores são próximos e nem sempre iguais. Portanto, torna-se necessário o cálculo de um valor limite para separar as classes (interesse e não interesse), denominado de *threshold*. Esse limite é normalmente representado graficamente por uma linha tracejada, cujo valor é baseado no teorema de Bayes (BARKER; RAYENS, 2003), o qual leva em consideração que a distribuição das amostras no conjunto de treinamento será semelhante a observada no conjunto de previsão. Ele é calculado de tal maneira que é minimizado o número de amostras previstas erroneamente pelo modelo PLS-DA (ALMEIDA et al., 2013). Quando a amostra tem um valor previsto acima do *threshold*, ela é considerada pertencente a classe 1 (classe de interesse) enquanto as amostras abaixo desse limite são pertencentes a classe 0 (classe não de interesse).

Um dos parâmetros importantes a ser escolhido durante a construção do modelo PLS-DA é a escolha do número de variáveis latentes (VL) que descreva a covariância nas amostras das matrizes e que tenham a correlação máxima com a classe de valores conhecidos, dando menos peso para a classe irrelevante ou a variância do ruído (BARKER; RAYENS, 2003). Quando o número de VL é abaixo do número ideal, informações importantes são descartadas do modelo (subajuste); o caso inverso, isto é, o número é acima do ideal, o ruído dos dados será modelado e informações não relevantes serão inseridas no modelo (sobreajuste).

Portanto, para encontrar o número ideal de VL é feita uma etapa de validação interna, isto é, usando as amostras do conjunto de treinamento. Neste trabalho optou-se pela validação cruzada do tipo *venetian blind* (venezianas, tradução livre). Nesse caso, parte das amostras é retirada do conjunto de treinamento e as demais amostras são usadas para construir o modelo PLS-DA. Utilizando esse modelo a classe das amostras retiradas é prevista. Esse processo é repetido até que todas as amostras tenham sido retiradas e previstas pelos modelos construídos. Assim, é possível calcular o erro quadrático médio de validação cruzada (RMSECV, do inglês *root mean squared error cross validation*) a partir da Equação 1.

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

em que  $\hat{y}_i$  é o valor de classe previsto,  $y_i$  é o valor da classe real e  $n$  é o número de amostras no conjunto de treinamento. O número ideal de VL é escolhido de acordo com o menor erro de RMSECV junto com a análise da variância capturada nos blocos  $\mathbf{X}$  e  $\mathbf{y}$  (LI et al., 2012).

Após a construção do modelo de classificação, pode-se fazer a análise da importância da variável na projeção (VIP, do inglês *variable importance in projection*), em que é estimada a importância de cada variável na projeção usada no modelo PLS-DA (ALMEIDA et al., 2013). As variáveis com escores maiores que do 1 são consideradas como as principais variáveis do modelo. O VIP expressa a importância de cada variável  $j$ , sendo refletida por  $w$  em cada variável latente e é dado pela Equação 2 (MEHMOOD et al., 2012):

$$v_j = \sqrt{\frac{p \sum_{a=1}^A [(q^2 at' ata)(waj/||wa||)^2]}{\sum_{a=1}^A (q^2 at' ata)}} \quad (2)$$

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

#### 1.1.5.1 Parâmetros de avaliação do modelo PLS-DA

A avaliação dos parâmetros de desempenho do modelo de classificação é feita através da Tabela de Confusão, a qual fornece a base para descrever os resultados do modelo PLS-DA e para caracterizar os erros. Os parâmetros utilizados foram falso positivo (FP), falso negativo (FN), verdadeiro positivo (VP), verdadeiro negativo (VN), sensibilidade e especificidade. Deve-se ressaltar que alguns parâmetros têm o mesmo nome, mas diferem da definição utilizada em química analítica e metrologia.

FP é o número de amostras que não pertencem a uma classe que foram classificadas como pertencentes (por exemplo, uma amostra adulterada que foi classificada como uma amostra autêntica). O inverso, isto é, as amostras pertencentes a uma classe que foram classificadas como não pertencentes é a taxa FN (por exemplo, uma amostra autêntica classificada como adulterada). VP é o número de amostras da classe que foram classificadas corretamente como pertencentes a classe (amostra autêntica classificada como autêntica) e VN é o número de amostras não-pertencentes à classe que foram classificadas corretamente como não-pertencentes (amostra adulterada classificada como adulterada).

A sensibilidade é definida como a capacidade do modelo de classificar corretamente as amostras como VP, e é dada pela Equação 3 (XU et al., 2012):

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (3)$$

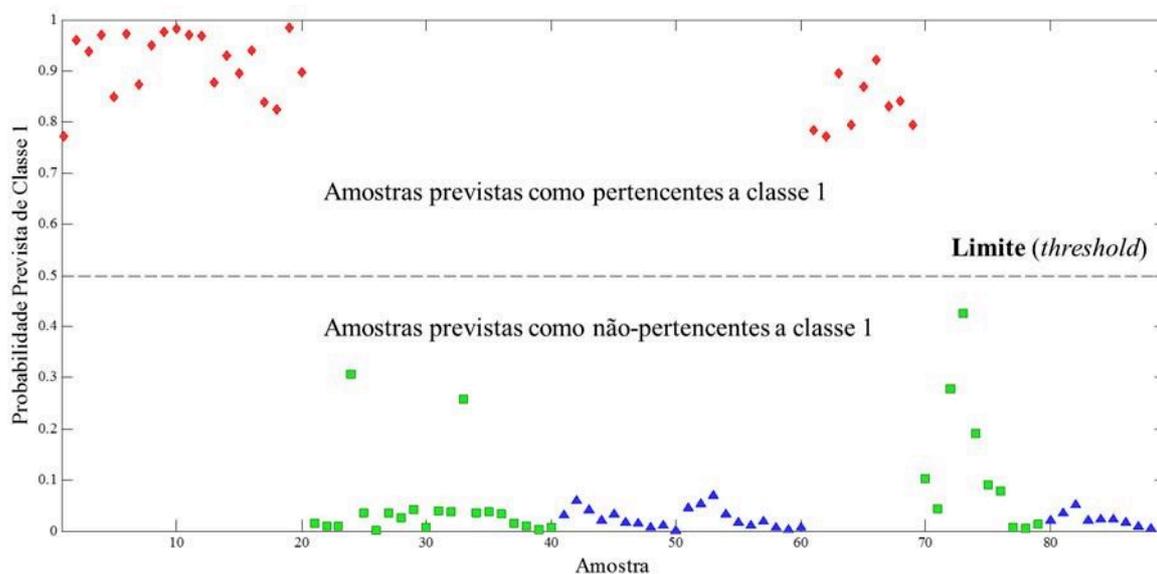
A especificidade, por outro lado, é a capacidade do modelo de classificar corretamente as amostras VN, e é calculada pela Equação 4 (XU et al., 2012):

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (4)$$

### 1.1.5.2 Representação gráfica do modelo PLS-DA

O resultado do modelo PLS-DA geralmente é expresso através de um gráfico de estimativas de classes, onde as amostras posicionadas acima da linha tracejada (*threshold*) são classificadas como pertencentes à classe 1 (classe de interesse) e, as amostras posicionadas abaixo do valor de *threshold* são classificadas como pertencentes à classe 0, ou seja, a classe que não é de interesse. A Figura 12 mostra um exemplo do gráfico de estimativas para um modelo PLS-DA.

**Figura 12.** Resultado gráfico do modelo PLS-DA



Fonte: O autor, 2018.

Através desse gráfico, verifica-se que as amostras da classe 1, tanto do conjunto de treinamento (amostras 1-20) quanto de teste (amostras 61-69) (◆ na Figura 12) foram classificadas corretamente, isto é, estão todas acima do *threshold*. Por outro lado, as amostras que não pertenciam à classe 1 (■ e ▲) ficaram abaixo do *threshold* para ambos os conjuntos de treinamento e validação.

### 1.1.6 Aplicações da espectroscopia no infravermelho aliada a métodos quimiométricos para a análise de biodiesel/diesel

O primeiro trabalho, para a quantificação de biodiesel em diesel foi publicado em 2001 por Knothe. Nesse caso foi empregada a espectroscopia no infravermelho próximo (NIR, do inglês *near infrared*) para a quantificação de biodiesel metílico de Soja na faixa de 1 a 95% (v/v). Foram utilizadas 38 amostras para o conjunto de calibração e 9 para o conjunto de previsão, empregando PLS. Os desvios máximos entre os valores previstos pelo modelo e os experimentais variaram entre 1-1,5. Além disso, foi destacado que o uso apenas da região vibracional da carbonila no MIR apresenta a desvantagem do fato que os espectros de óleo vegetal e biodiesel serem muito parecidos nessa região (KNOTHE, 2001).

Santos e colaboradores (2016) empregaram a espectroscopia MIR aliada a métodos quimiométricos para a discriminação entre biodieseis preparados a partir de óleos de Soja, milho, canola, girassol e oliva. Dos métodos testados, os modelos de PLS-DA e máquina de suporte de vetor (SVM, do inglês *support vector machine*) se mostraram os melhores para a discriminação das amostras, com sensibilidade e especificidade acima de 81,3%. As regiões espectrais usadas para a construção do modelo foram  $3699-2974\text{ cm}^{-1}$  e  $1841-925\text{ cm}^{-1}$  e foram usadas 7 VL.

Vasconcelos e colaboradores (2012) utilizaram a espectroscopia NIR junto aos métodos PLS e regressão linear múltipla (MLR, do inglês *multiple linear regression*) junto ao algoritmo das projeções sucessivas para quantificar biodiesel em diesel no intervalo de 0,0 a 10,0% (v/v), mesmo na presença de óleo adulterante. Seus resultados usando esse algoritmo de seleção de variáveis mostraram-se mais promissores do que os obtidos pelos modelos usando as bandas de absorção de combinação do NIR e regiões do MIR.

Feng, Wu e Zeng (2015) utilizaram 441 espectros NIR de amostras de diesel para construir modelos PLS e SVM, sendo o primeiro usado para quantificar e o segundo para classificar. PLS foi usado como análise multi-parâmetro e foi usado para prever 6 parâmetros de qualidade (ponto de fulgor, índice de cetano, densidade, ponto de entupimento de filtro a frio, total de aromáticos e viscosidade). O modelo SVM permitiu classificar com exatidão de 93,75% os biodieseis em quatro diferentes classes, dependendo da sua qualidade.

Garmarudi e colaboradores (2019) usaram MIR aliado à análise de componentes principais (PCA, do inglês *principal component analysis*) e SIMCA para classificar 251 espectros de petróleo bruto de acordo com a origem geográfica. Para tal, as amostras foram divididas em 10 classes (7 campos diferentes e 3 misturas de combinações). O modelo SIMCA se mostrou efetivo, com apenas 4 amostras previstas erroneamente.

Balabin e Safieva (2011) compararam diversos métodos quimiométricos como: Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA), K vizinho Mais Próximo (KNN) e SVM para a classificação de biodiesel em relação ao óleo utilizado para a produção (canola, girassol, coco, palma, Soja, algodão, mamona, Pinhão Manso, linhaça e óleo de fritura). Foram obtidos espectros NIR de 403 amostras, e o desempenho dos modelos foi comparado usando a Tabela de Confusão, em que o modelo de SVM se mostrou mais eficiente na separação, com erro de classificação em torno de 5%.

Guimarães e colaboradores (2014) utilizaram o PLS para quantificar etanol em amostras de biodieseis etílicos de óleo de Soja e residual. Os modelos obtidos foram eficientes na determinação da concentração de etanol com erros abaixo de 0,2% de adulterante. Além disso foram validados de acordo com orientações e figuras de méritos. Os mesmos autores utilizaram o PLS para quantificar óleo de Soja em biodieseis metílicos e etílicos de óleo de Soja em suas misturas com o diesel (B5). Também obtiveram baixos valores de erros de quadráticos de calibração (RMSEC) abaixo de 0,21%, alta linearidade e ausência de erros sistemático (GUIMARÃES et al., 2014).

Souza e colaboradores (2014), na quantificação de óleo automotivo usado como adulterante em diesel brasileiro S-10, analisaram amostras usando PLS e o modelo apresentou baixo valor de erros de previsão 0,04% (v/v) sendo validado pelas estimativas de figuras de mérito e orientações da norma ASTM E 1655.

Também foram publicados artigos usando PLS-DA, método de classificação na identificação do tipo de óleo e álcool usado na produção de biodiesel (MAZIVILA et al., 2015a), na identificação da presença de adulterantes como óleo de fritura usado, automotivo e gasolina (MAZIVILA et al., 2015b).

Um método de monitoramento foi utilizado por Buiatte e colaboradores (2016). Cartas de controle multivariadas baseadas no sinal analítico líquido (NAS) juntamente com o método quantitativo PLS na análise de controle de qualidade de biodiesel metílico de algodão em mistura com o diesel (B7). No método qualitativo foi analisado o teor de biodiesel no diesel e no método quantitativo foi analisado o conteúdo de biodiesel no diesel com baixos valores de erros de previsão 0,05% (v/v), boa correlação e ausência de erro sistemático. Os dois métodos conseguiram resultados satisfatórios tanto na análise qualitativa quanto quantitativa.

Para analisar o teor de biodiesel metílico de algodão no diesel (B10) tanto no aspecto qualitativo quanto quantitativo, Máquina e colaboradores (2019), utilizaram PLS e PLS-DA em que ambos os métodos obtiveram bons resultados. No PLS obtiveram baixos valores de erros e o modelo foi validado por figuras de mérito, orientações internacionais e norma ASTM E 1655. No PLS-DA, os parâmetros utilizados na classificação das amostras como sensibilidade e especificidade tiveram 100% de conformidade.

## 2 OBJETIVOS

### 2.1 Objetivo geral

Desenvolver modelos PLS-DA usando a espectroscopia no infravermelho médio para o controle de qualidade de misturas biodiesel/diesel (B10), com biodiesel produzido via rota metálica e óleos de Crambe e Pinhão Manso.

### 2.2 Objetivos específicos

- Construir modelos utilizando o algoritmo PLS-DA para classificar amostras de biodiesel/diesel dentro da conformidade em relação ao teor de biodiesel;
- Construir modelos PLS-DA para classificar as amostras de biodiesel/diesel que estão em conformidade em relação à presença/ausência de adulterantes como: óleo de Soja, óleo residual, óleo automotivo usado e gasolina;
- Validar os modelos PLS-DA construídos usando os parâmetros de sensibilidade e especificidade.

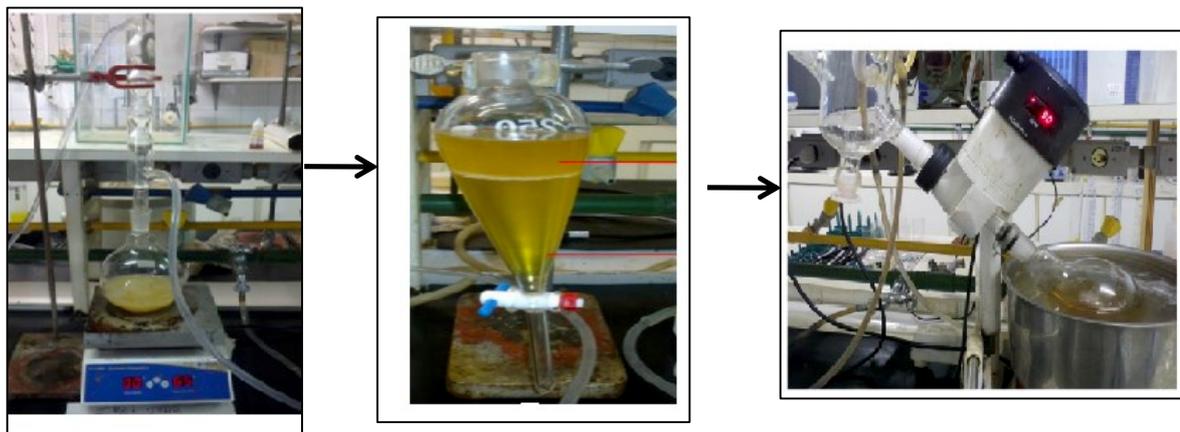
### 3 PROCEDIMENTO EXPERIMENTAL

#### 3.1 Amostras de dieseis e biodieseis

O diesel puro utilizado para o desenvolvimento deste trabalho foi cedido pela TRANSPETRO S/A (Uberlândia, Minas Gerais, Brasil). Os biodieseis metílicos de Pinhão Manso e Crambe foram sintetizados por mim no Laboratório de Quimiometria do Triângulo (LQT) da Universidade Federal de Uberlândia, (Minas Gerais, Brasil).

Para a produção do biodiesel de Crambe foi necessária à catálise dupla, uma vez que o óleo bruto possui a um índice de acidez acima do máximo permitido (0,5mg KOH/g, P.A. merck). Durante a catálise ácida (esterificação), utilizou-se o óleo e o álcool na proporção 1:1 (m/m) e 3% do ácido sulfúrico concentrado (P.A. A.C.S. marca Synth) em relação à massa do óleo. Em um balão de fundo redondo adicionou-se o ácido ao álcool metílico (P.A. A.C.S. marca Synth) e, em seguida, o óleo, permanecendo sobre fluxo em ebulição durante 3 horas a 60 °C, com agitação de 300 rpm, conforme Figura 13.

**Figura 13.** Esquema da reação de esterificação do biodiesel

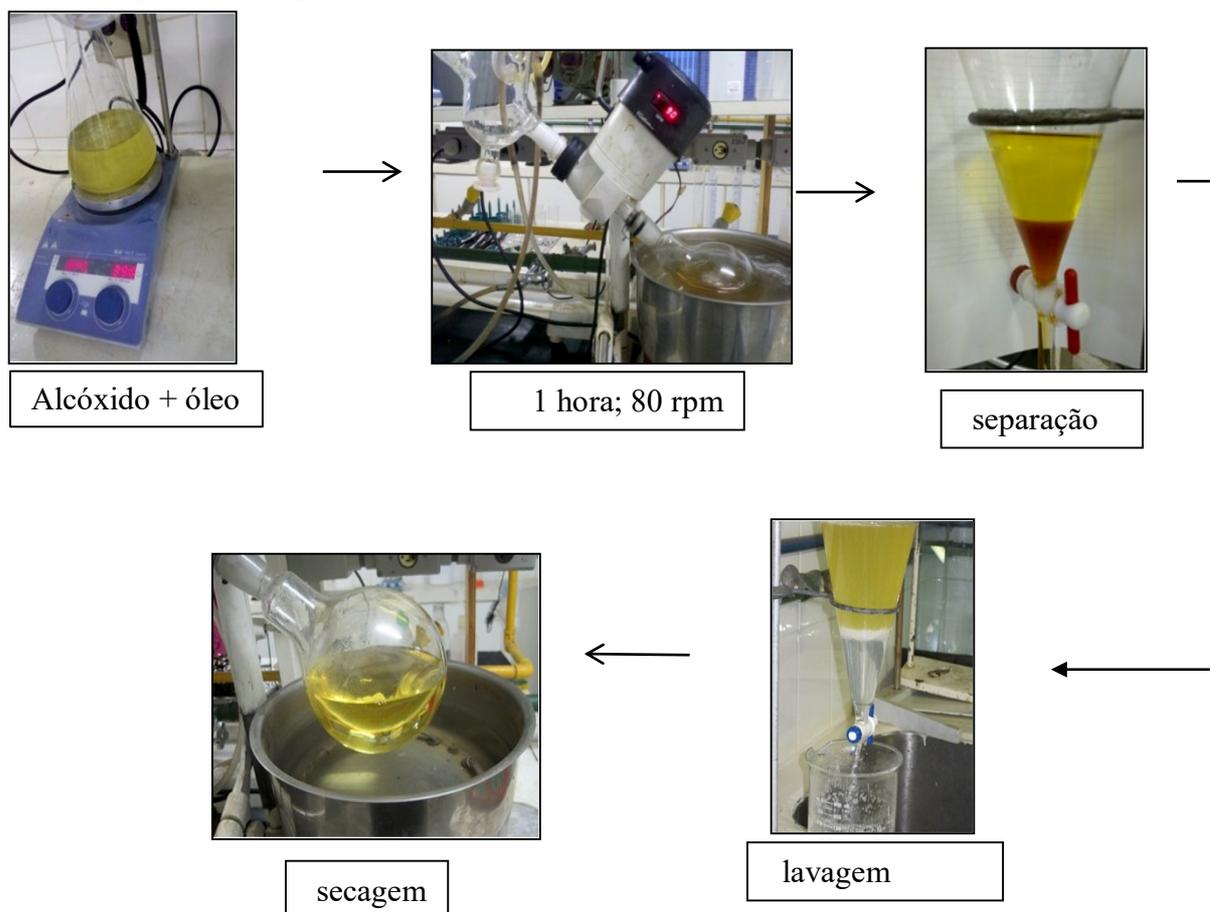


**Fonte:** O autor, 2018.

Após a esterificação a mistura heterogênea resultante foi transferida para um funil de decantação permanecendo em repouso até a separação das fases em que na parte inferior encontrava-se a mistura de ésteres e na fase superior o álcool. A fase inferior foi submetida a uma catálise básica (transesterificação) utilizando-se a proporção de 100,0 g do óleo para 30 g do álcool e 1,0 g de KOH. Misturou-se o álcool como catalisador formando o alcóoxido.

Em seguida, adicionou-se o óleo à mistura, permanecendo sob agitação a 300 rpm, temperatura ambiente e durante 80 min. Na etapa de purificação, a mistura permaneceu em decantação em um funil de separação contendo biodiesel e glicerol. Este foi removido da mistura, e o biodiesel lavado com água destilada a 85°C para retirada de impurezas. A secagem do biodiesel foi realizada utilizando o rotaevaporador por 1 hora a 80 rpm e 80 °C. O fluxograma com as etapas de processo é mostrado na Figura 14.

**Figura 14.** Esquema da reação de transesterificação do biodiesel



Fonte: O autor, 2018.

### 3.2 Preparo de amostras de biodiesel/diesel

As amostras foram pesadas à temperatura ambiente ( $25 \pm 2$  °C) utilizando-se uma balança analítica (Sartorius, modelo BP211D).

Para os modelos PLS-DA para teor de biodiesel metílico de Pinhão Manso e Crambe, foram preparadas 80 e 90 amostras, respectivamente, com o teor de biodiesel variando entre 1,0 a 30,0% (v/v). Destas, 40 e 45 referem-se às amostras B7 ou B10 (classe de interesse) e as demais são amostras com teores abaixo ou acima do permitido.

Para os modelos PLS-DA na adulteração com óleo de Soja, foram preparadas 45 amostras sem adulteração e 45 amostras adulteradas na faixa de concentração de 1,0 a 30,0% (v/v).

Em todos os modelos, 2/3 das amostras foram utilizadas para o conjunto de treinamento, enquanto 1/3 foi utilizado como conjunto de validação externa.

### 3.3. Obtenção dos espectros MIR.

Os espectros MIR das amostras de todos os modelos foram obtidos sem quintuplicatas, na região de 4.000 a 600  $\text{cm}^{-1}$  resolução espectral de  $4\text{cm}^{-1}$  e 16 varreduras. Foi utilizado um espectrômetro FT-MIR Spectrum Two (Perkin Elmer, Waltham, MA, EUA) com Transformada de Fourier, equipado com um acessório de amostragem ATR horizontal feito de ZnSe (Pike Technologies, Fitchburg, WI, EUA). Após a obtenção de cada espectro, o ATR horizontal foi limpo com álcool isopropílico (marca Sint 95% v/v)

### 3.4. Construção dos modelos PLS-DA

Foram construídos 10 modelos PLS-DA para classificar amostras autênticas e não autênticas com relação ao teor de biodiesel no diesel e quanto à presença de adulterantes.

Quanto ao teor foram construídos 2 modelos: biodiesel metílico de Pinhão Manso com 7% de biodiesel no diesel (B7) e biodiesel metílico de Crambe com 10% de biodiesel no diesel (B10). Em relação às adulterações foram construídos 4 modelos para a mistura de biodiesel/diesel de Pinhão Manso adulterados com: óleo de Soja, óleo residual, óleo automotivo usado e gasolina (B7) e 4 modelos com a mistura de biodiesel/diesel de Crambe também adulterados com: óleo de Soja, óleo residual, óleo automotivo usado e gasolina (B10).

Os pré-processamentos e a construção dos modelos PLS-DA foram realizados usando os programas Matlab v. 9.0 (Mathworks Inc., Natick, MA, EUA) e PLS\_toolbox v. 7.4 (Eigen vectorResearch Inc., Wenatchee, EUA).

Para o desenvolvimento dos modelos de classificação PLS-DA, inicialmente a linha de base dos espectros foi corrigida usando a função *baseline*, em que ajusta um polinômio na região sem sinal/bandas no MIR, a qual é então subtraída dos espectros originais. A região espectral escolhida para a construção dos modelos foi de 3.099-659  $\text{cm}^{-1}$ . Além do *baseline*, outro pré-processamento utilizado foi o *Savitzky-Golay* que é um método de suavização e serve para diminuir a amplificação de ruído (SAVITZKY; GOLAY, 1964).

Os dados foram centrados na média e a validação cruzada *venetian blind* foi usada para a validação interna dos modelos. O número ideal de VL foi escolhido usando como critérios o valor de RMSECV e em análise conjunto com a variância acumulada em  $\mathbf{X}$  e  $\mathbf{y}$ .

Após a construção do modelo PLS-DA, foi verificada a presença de *outliers*, isto é, amostras anômalas (amostras que possuem comportamento diferente comparado com o restante do modelo). Geralmente amostras anômalas são avaliadas com base no *leverage* (o grau que uma amostra está distante da média do conjunto de dados) e nos resíduos não modelados nos dados espectrais ou nos valores de concentração analítica.

A validação externa do modelo foi realizada usando o conjunto de validação, as quais fornecem o valor de erro quadrático médio de previsão (RMSEP, do inglês *root mean square error of prediction*). A qualidade dos modelos também foi avaliada através da Tabela de Confusão e os parâmetros erro quadrático médio de calibração (RMSEC, do inglês *root mean square of calibration*), sensibilidade e especificidade. O parâmetro VIP também foi analisado e associado aos grupos funcionais responsáveis pela separação das classes de amostras autênticas e adulteradas.

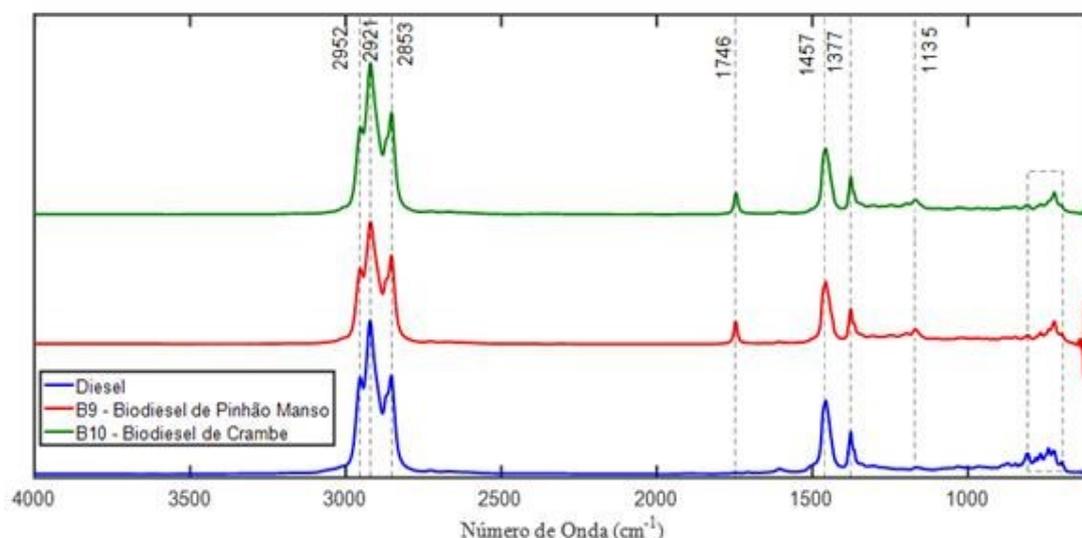
## 4 RESULTADOS E DISCUSSÃO

### 4.1 Caracterização dos espectros MIR

A Figura 15 mostra os diferentes espectros das misturas biodiesel/diesel autênticas e do diesel na faixa de 4.000 a 600  $\text{cm}^{-1}$  empregados neste estudo. O modelo de biodiesel/diesel de Pinhão Manso foi construído como B7 (7% de biodiesel no diesel) e B9 (9% de biodiesel no diesel) de acordo com a legislação vigente no momento em que foram desenvolvidos. Já os modelos com o Crambe foram desenvolvidos após a alteração da lei de março/2018 em que o teor de biodiesel no diesel passou a ser 10% (B10).

É possível verificar que os espectros dos biodieseis apresentam grande semelhança entre si. O espectro do diesel apresenta, essencialmente absorções correspondentes aos modos vibracionais característicos de alcanos normais. As bandas em 2.952, 2.921 e 2.853 referem-se aos estiramentos assimétricos de  $-\text{CH}_3$ , simétrico e assimétrico de  $-\text{CH}_2$ , respectivamente. A banda característica das vibrações de deformação axial carbonila,  $\text{C}=\text{O}$ , aparece em 1.746  $\text{cm}^{-1}$  e bandas na região de 1.135  $\text{cm}^{-1}$  são atribuídas a deformação axial da ligação  $\text{C}-\text{O}$ , presente apenas nos espectros de biodieseis. As bandas formadas em 1.746, 1.456 e 1.376  $\text{cm}^{-1}$  são referentes a vibrações de deformação axial  $\text{C}=\text{C}$ , deformação angular de  $\text{C}-\text{H}$  em  $\text{CH}_3$  e  $\text{CH}_2$ , respectivamente (HOLLER; SKOOG; CROUCH, 2009). As bandas na região 750-800  $\text{cm}^{-1}$  (destacados no retângulo na Figura 15) são consideradas características de derivados de petróleo e, portanto, podem ser atribuídos ao diesel (GAYDOU; KISTER; DUPUY, 2011).

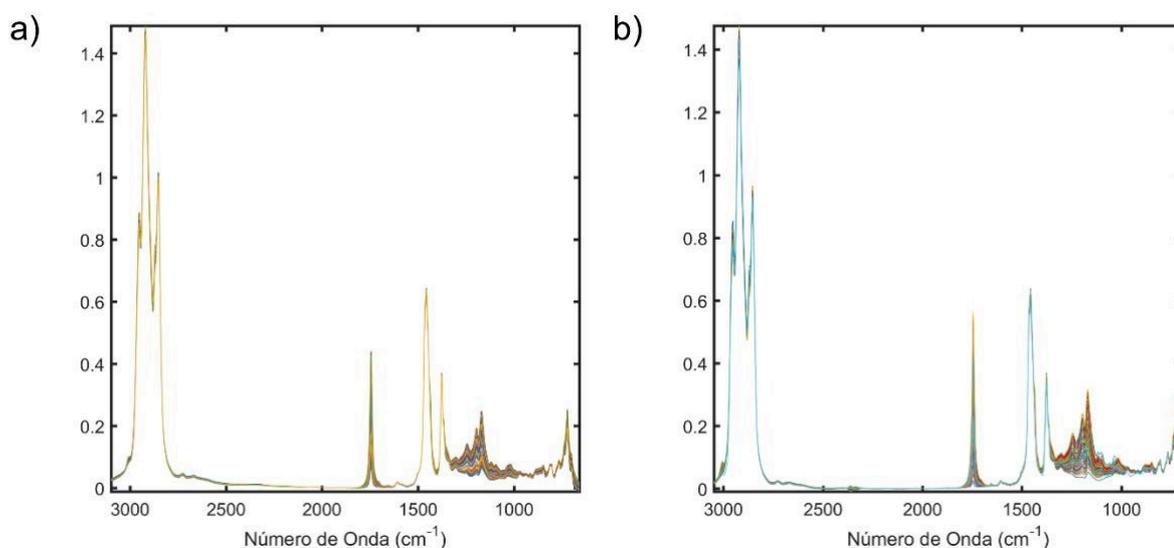
**Figura 15.** Espectros MIR de diesel puro (azul), B9 de Pinhão Manso (vermelho) e B10 de Crambe (verde)



Fonte: O autor, 2018.

A Figura 16 apresenta os espectros obtidos para as amostras de treinamento e teste para a construção de cada modelo PLS-DA. A Figura 16a está relacionada ao modelo de biodiesel metílico de Pinhão Manso enquanto a Figura 16b ao modelo de biodiesel metílico de Crambe em diferentes proporções com o diesel. Para os dois modelos, as regiões de maiores variações ocorrem na região de impressão digital (entre  $1.300-1.000\text{ cm}^{-1}$ ) e na região da carbonila, em torno de  $1.740$  a  $1.750\text{ cm}^{-1}$ . Isso porque estas regiões correspondem a ligações de carbono e oxigênio, presentes tanto no biodiesel de Pinhão Manso quanto no biodiesel de Crambe. No entanto, também são verificadas pequenas variações nas regiões do diesel, uma vez que com o aumento na concentração de biodiesel, ocorre uma diminuição na concentração do diesel. Deve-se ressaltar que, devido à grande sobreposição de sinais, uma análise visual ou mesmo um método univariado (como a intensidade de uma banda em específico) se torna praticamente impossível. Portanto, o uso de métodos multivariados para extrair a informação é extremamente útil e necessário nessas situações.

**Figura 16.** Espectros MIR das amostras usadas para a construção do (a) teor de biodiesel metílico de Pinhão Manso e (b) teor de biodiesel de Crambe.



Fonte: O autor, 2017

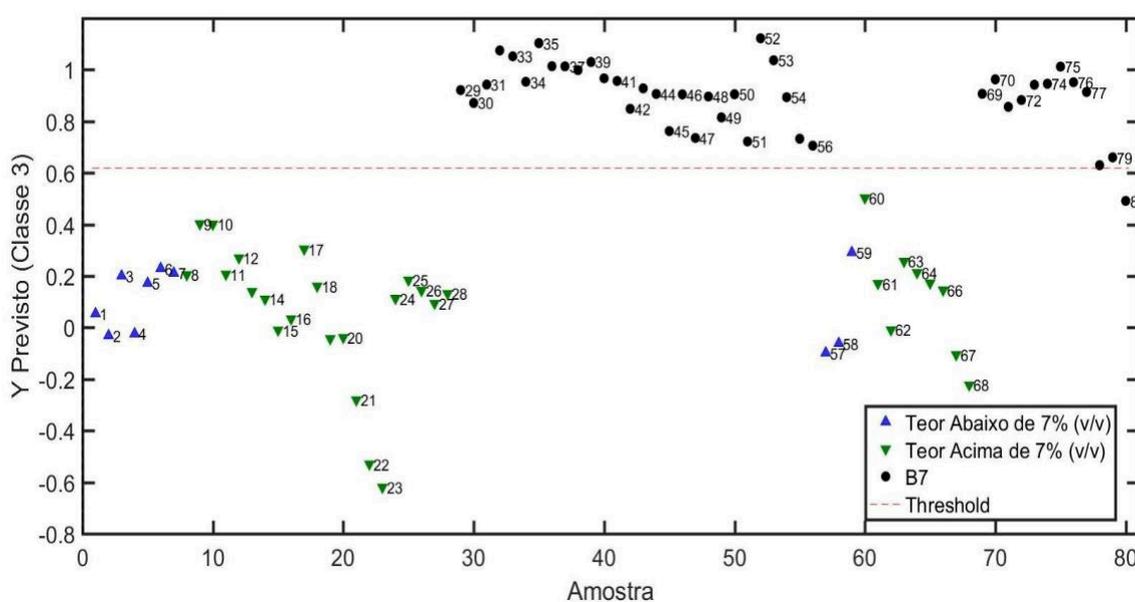
## 4.2 Construção do modelo PLS-DA

### 4.2.1 Teor de biodiesel de Pinhão Manso e Crambe

Foram construídos dois modelos PLS-DA para a distinção das amostras de biodiesel/diesel em relação ao teor. Para o modelo de biodiesel/diesel de Pinhão Manso a classe de interesse refere-se ao B7, isto é, 7% (v/v) de biodiesel na mistura. E, para o modelo de biodiesel/diesel de Crambe a classe de interesse refere-se ao B10, isto é, 10% (v/v) de biodiesel

na mistura. No pré-processamento usado, além da correção da linha de base, foi também utilizado o método de suavização *Savitzky-Golay* que serve para calcular um polinômio de tamanho e grau determinado que melhor ajuste ao conjunto de dados que foram utilizados. Neste trabalho os dados foram alisados com 11 pontos, polinômio de segundo grau seguido de primeira derivada (11, 2, 1) para melhorar a resolução de sinal. O modelo PLS-DA com relação ao teor de biodiesel de Pinhão Manso no diesel foi utilizado com mais de uma classe de interesse. A classe 1 refere-se às misturas BX cujo teor de biodiesel é abaixo de 7,00% (v/v), a classe 2 são as misturas BX com teor de biodiesel acima de 7,00% (v/v) e a classe 3 são as amostras B7 (classe de interesse). A opção de usar duas classes para as amostras que não são de interesse é devido à natureza linear do modelo PLS-DA, o que poderia dificultar ou mesmo atrapalhar o desempenho do mesmo caso estas fossem agrupadas em uma única classe. O modelo foi construído com 5 variáveis latentes (VL) e capturou 99,60 e 87,13% de variância nos blocos **X** e **Y**, respectivamente. Nenhuma amostra anômala (*outlier*) foi detectada nesse modelo. As amostras de 1 a 56 são as amostras de treinamento, isto é, as amostras que foram utilizadas para a construção do modelo. As amostras de 57 a 80 (totalizando 24) são as amostras do conjunto de teste para avaliar se o modelo construído consegue prever bem amostras desconhecidas. Nesse caso, apenas uma amostra do conjunto teste (amostra 80) da classe 3 (B7) foi prevista erroneamente como não pertencente a esta classe, uma vez que o valor previsto para esta amostra está abaixo do valor de *threshold* (Figura 17).

**Figura 17.** Estimativas de B7 de biodiesel metílico de Pinhão Manso.

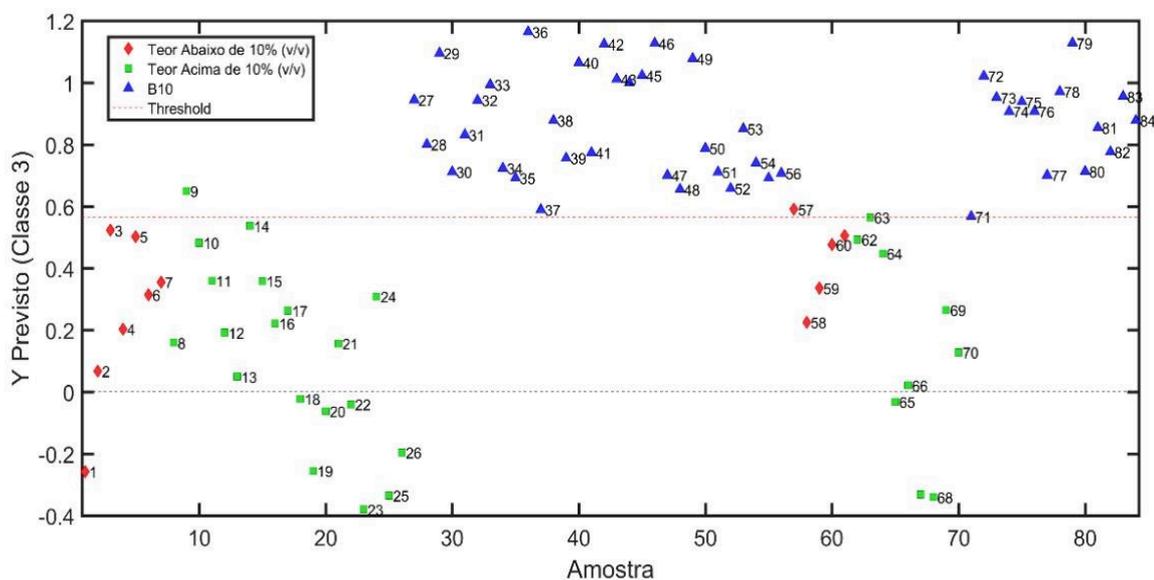


Fonte: O autor, 2018.

No modelo PLS-DA para a distinção das amostras de biodiesel de Crambe/diesel em relação ao teor, a classe de interesse refere-se ao B10, isto é, 10,00% (v/v) de biodiesel na

mistura. Foi realizada a correção da linha de base e os dados foram centrados na média. Assim como desenvolvido no caso do teor de Pinhão Manso foi utilizado um modelo PLS-DA com mais de uma classe de interesse. A classe 1 refere-se às misturas BX cujo teor de biodiesel é abaixo de 10,00% (v/v), a classe 2 são as misturas BX com teor de biodiesel acima de 10,00% (v/v) e a classe 3 são as amostras B10 (classe de interesse). Como pré-processamento foi utilizada a normalização pelo quociente probabilístico (PQN, do inglês *probabilistic quotient normalization*) para corrigir problemas de diferenças na escala devido a uma superfície não homogênea, e posteriormente os dados foram centrados na média (SACRÉ et al., 2014). O modelo foi construído com 6 variáveis latentes (VL) e capturou 99,84 e 73,04% de variância nos blocos X e Y, respectivamente. O gráfico das estimativas do PLS-DA é mostrado na Figura 18. Da mesma maneira que no modelo de biodiesel metílico de Pinhão Manso, as amostras de 1 a 56 são as amostras de treinamento e as amostras de 57 a 84 (totalizando 28) são as amostras do conjunto teste usadas para a validação do modelo. Nesse caso, apenas uma amostra do conjunto de treinamento (amostra 9) da classe 2 (acima de B10) foi prevista erroneamente como pertencente a esta classe, uma vez que o valor previsto para esta amostra está acima do valor de *threshold*. Na validação, as amostras do conjunto de teste, 57 (abaixo de B10) e 63 (acima de B10) foram previstas erroneamente como pertencentes a classe B10.

**Figura 18.** Estimativas de B10 de biodiesel metílico de Crambe.



Fonte: O autor, 2018.

Os parâmetros: verdadeiro positivo (VP), falso positivo (FP), verdadeiro negativo (VN), falso negativo (FN) sensibilidade e especificidade para os biodieseis metílico de Pinhão Manso

e Crambe foram calculados de acordo com a Tabela 2. Na tabela 2a estão os resultados dos parâmetros citados para o modelo de biodiesel metílico de Pinhão Manso.

**Tabela 2a** Tabela de confusão para os resultados do modelo PLS-DA para teor de biodiesel metílico de Pinhão Manso

Modelo	Biodiesel metílico de Pinhão Manso					
	Treinamento			Teste		
	Real Classe			Real Classe		
	Classe 1	Classe 2	Classe 3	Classe 1	Classe 2	Classe 3
<b>Previsto como Classe 1</b>	7	0	0	3	0	1
<b>Previsto como Classe 2</b>	0	21	0	0	9	0
<b>Previsto como Classe 3</b>	0	0	28	0	0	11

Fonte: O autor, 2018.

Para a classe 1 do modelo de teor de biodiesel metílico de Pinhão Manso abaixo de B7:

- VP são as amostras pertencentes a esta classe que foram previstas corretamente como pertencentes a ela ( $VP_{\text{treinamento}} = 7$  e  $VP_{\text{teste}} = 3$ );
- FP são as amostras não pertencentes a esta classe que foram previstas erroneamente como pertencentes a ela ( $FP_{\text{treinamento}} = 0$  e  $FP_{\text{teste}} = 1$ );
- VN são as amostras não pertencentes a esta classe que foram previstas corretamente como não pertencentes a mesma ( $VN_{\text{treinamento}} = 21 + 28 = 49$  e  $VN_{\text{teste}} = 9 + 11 = 20$ );
- FN são as amostras pertencentes a esta classe que foram previstas erroneamente como não pertencentes a esta classe ( $FN_{\text{treinamento}} = 0$  e  $FN_{\text{teste}} = 0$ ).

Portanto, pelas Equações 3 e 4 (secção 1.1.5), os resultados para o modelo de biodiesel metílico de Pinhão Manso são:

$$Sensibilidade_{\text{treinamento}} = \frac{VP_{\text{treinamento}}}{VP_{\text{treinamento}} + FN_{\text{treinamento}}} = \frac{7}{7 + 0} = 1$$

$$Sensibilidade_{teste} = \frac{VP_{teste}}{VP_{teste} + FN_{teste}} = \frac{3}{3 + 0} = 1$$

$$Especificidade_{treinamento} = \frac{VN_{treinamento}}{VN_{treinamento} + FP_{treinamento}} = \frac{49}{49 + 0} = 1$$

$$Especificidade_{teste} = \frac{VN_{teste}}{VN_{teste} + FP_{teste}} = \frac{20}{20 + 1} = 0,95$$

Os parâmetros VP, FP, VN, FN sensibilidade e especificidade para o biodiesel metílico de Crambe foram calculados de acordo com a Tabela 2b.

**Tabela 2b.** Tabela de confusão para os resultados do modelo PLS-DA para teor de biodiesel metílico de Crambe

Modelo	Biodiesel metílico de Crambe					
	Treinamento			Teste		
	Real Classe			Real Classe		
	Classe 1	Classe 2	Classe 3	Classe 1	Classe 2	Classe 3
Previsto como Classe 1	7	0	0	4	0	1
Previsto como Classe 2	0	18	1	0	8	1
Previsto como Classe 3	0	0	30	0	0	14

Fonte: O autor, 2018.

Para a classe 1 do teor de biodiesel metílico de Crambe abaixo de B10:

- VP são as amostras pertencentes a esta classe que foram previstas corretamente como pertencentes a ela ( $VP_{treinamento} = 7$  e  $VP_{teste} = 4$ );
- FP são as amostras não pertencentes a esta classe que foram previstas erroneamente como pertencentes a ela ( $FP_{treinamento} = 0$  e  $FP_{teste} = 1$ );

- VN são as amostras não pertencentes a esta classe que foram previstas corretamente como não pertencentes a mesma ( $VN_{\text{treinamento}} = 18+30 = 48$  e  $VN_{\text{teste}} = 8+14 = 22$ );
- FN são as amostras pertencentes a esta classe que foram previstas erroneamente como não pertencentes a esta classe ( $FN_{\text{treinamento}} = 0$  e  $FN_{\text{teste}} = 0$ ).

E os resultados para o modelo de biodiesel metílico de Crambe são:

$$\text{Sensibilidade}_{\text{treinamento}} = \frac{VP_{\text{treinamento}}}{VP_{\text{treinamento}} + FN_{\text{treinamento}}} = \frac{7}{7 + 0} = 1$$

$$\text{Sensibilidade}_{\text{teste}} = \frac{VP_{\text{validação}}}{VP_{\text{validação}} + FN_{\text{validação}}} = \frac{4}{4 + 0} = 1$$

$$\text{Especificidade}_{\text{treinamento}} = \frac{VN_{\text{treinamento}}}{VN_{\text{treinamento}} + FP_{\text{treinamento}}} = \frac{48}{48 + 0} = 1$$

$$\text{Especificidade}_{\text{teste}} = \frac{VN_{\text{validação}}}{VN_{\text{validação}} + FP_{\text{validação}}} = \frac{22}{22 + 1} = 0,95$$

Os resultados para todas as 3 classes, tanto no conjunto de treinamento quanto teste para os dois modelos são mostrados na Tabela 3. Esses resultados com baixos valores de erros: erro quadrático médio de calibração-RMSEC (inferior a 0,27%), erro quadrático médio de validação cruzada- RMSECV (inferior a 0,33%) e erro quadrático médio de previsão RMSEP (inferior a 0,30%) sensibilidade acima de 0,92 e especificidade acima de 0,95) atestam a boa eficiência do modelo PLS-DA de biodiesel metílico de Pinhão Manso construído para a classificação das amostras em relação ao teor de biodiesel estar de acordo ou não com o especificado pela ANP.

Também apresenta os resultados obtidos para o modelo de biodiesel metílico de Crambe, onde baixos valores de erros de calibração (RMSEC), validação cruzada (RMSECV) e previsão (RMSEP), sensibilidade acima de 0,91 e especificidade acima de 0,856 atestam a boa eficiência do modelo PLS-DA construído para a classificação das amostras em relação ao teor de biodiesel estar de acordo ou não com o especificado pela ANP.

**Tabela 3.** Parâmetros de classificação obtidos pelos modelos PLS-DA para o teor de biodiesel de Pinhão Manso e Crambe na mistura BX.

Parâmetros	Valores dos modelos PLS-DA					
	BMPM			BMC		
	Classe 1	Classe 2	Classe 3	Classe 1	Classe 2	Classe 3
RMSECV (% v/v)	0,15	0,18	0,25	0,23	0,25	0,33
RMSEC (% v/v)	0,12	0,15	0,20	0,19	0,22	0,27
RMSEP (% v/v)	0,15	0,17	0,22	0,23	0,22	0,30
Especificidade <sub>treinamento</sub>	1,00	1,00	1,00	0,96	1,00	0,97
Especificidade <sub>teste</sub>	0,95	1,00	1,00	1,00	1,00	0,86
Sensibilidade <sub>treinamento</sub>	1,00	1,00	1,00	1,00	1,00	1,00
Sensibilidade <sub>teste</sub>	1,00	1,00	0,92	1,00	1,00	1,00

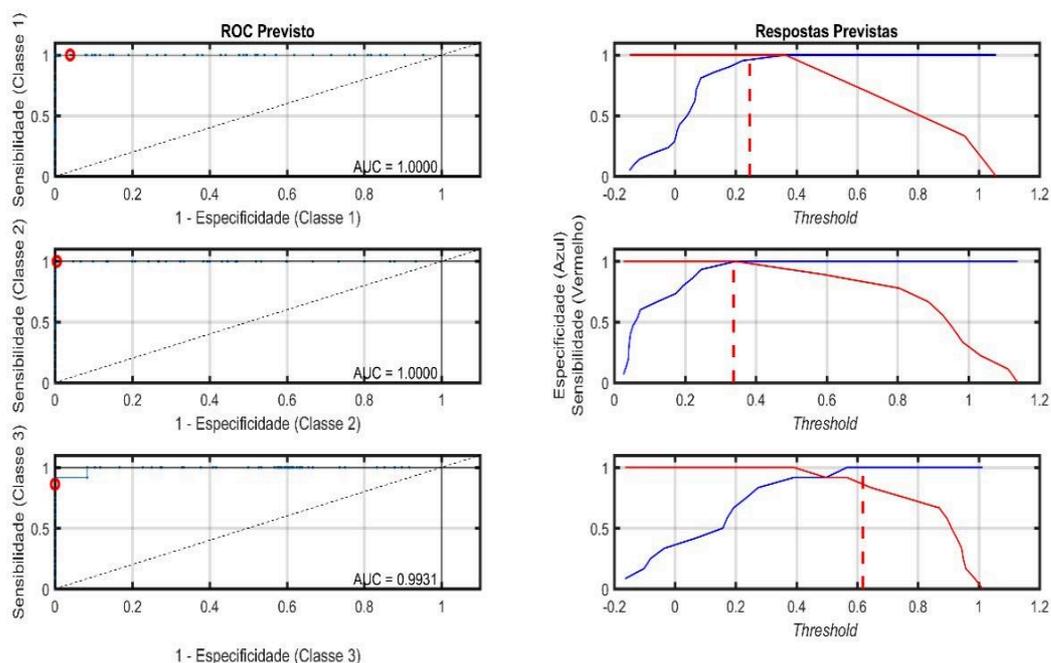
Fonte: O autor, 2018.

Além desses parâmetros numéricos apresentados na Tabela 3, ferramentas gráficas como as curvas característica de operação do receptor (ROC, do inglês *receiver operating characteristics*) podem ser usadas para avaliar a capacidade de classificação de um modelo (Figura 19), no caso de biodiesel metílico de Pinhão Manso. Estas curvas, calculadas para cada classe separadamente, são gráficos da especificidade vs. a sensibilidade. Como o eixo  $x$  é '1 – especificidade', o valor ideal deste é igual a 0, enquanto para o eixo  $y$  o valor ideal é igual a 1 (sensibilidade). Portanto, um método de classificação perfeito forneceria um ponto no canto esquerdo superior do espaço ROC, o que representa máxima sensibilidade e especificidade. Esse valor foi alcançado para a classe 2 e 1, enquanto para a classe 3 houve um pequeno desvio (uma vez que uma amostra da classe 3 foi prevista como pertencente a classe 1). A área sob a curva ROC (AUC, do inglês *area under the ROC curve*) mede a área bidimensional inteira debaixo da curva ROC (a integral) de (0,0) a (1,1) cujo valor varia entre 0 e 1. Em modelos cuja previsão é 100% errada, AUC é igual a 0 e 100% correta, AUC é igual a 1.

Além disso, no caso de uma classificação aleatória, existiriam pontos ao longo da diagonal do canto inferior esquerdo ao canto superior direito. Outro gráfico a ser observado é a variação do *threshold* para cada classe (Figura 19). De acordo com o teorema de Bayes, este valor limite deve ser escolhido de forma a minimizar os falsos negativos e falsos positivos, isto é, sensibilidade e especificidade devem ser idealmente 1. A sensibilidade para valores de *threshold* baixos é sempre igual a 1, isso porque as amostras VP estarão acima deste limite. No entanto, *threshold* muito baixo diminui a especificidade, isto é, aumenta o número de FP. Em casos de limites muito elevados, ocorre o inverso: as amostras VN estão abaixo do limite, mas aumentam-se os casos de FN (amostras pertencentes a uma classe previstas como não pertencentes) e, portanto, diminui-se a sensibilidade. Assim, o valor de *threshold* ideal corresponde ao ponto onde a linha da especificidade (em azul) cruza com a linha da sensibilidade (em vermelho) (BALLABIO; CONSONNI, 2013). O limite de *threshold* para a

classe 3 (amostras de B7) é igual a 0,618. Isso implica que todas as amostras em que o valor Y previsto for maior do que 0,618 será classificado como pertencente a classe 3.

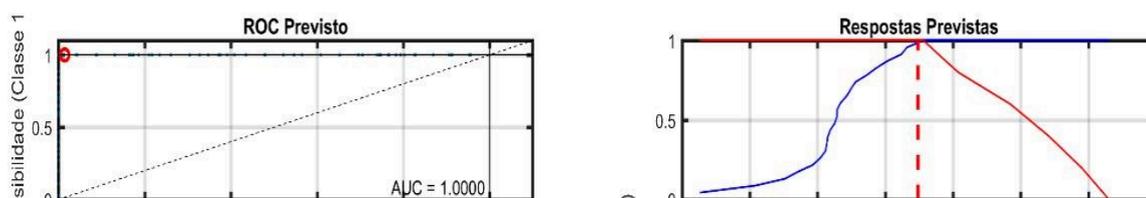
**Figura 19.** ROC dos modelos PLS-DA para teor de biodiesel metílico de Pinhão Manso



Fonte: O autor, 2018.

As curvas ROC e a relação de especificidade, sensibilidade e *threshold* do modelo de biodiesel metílico de Crambe são mostradas na Figura 20. Como a taxa de classificação correta não foi 100% para a classe 3, AUC foi menor do que 1,0 (0,9949) e o máximo de sensibilidade e especificidade foi alcançado em um ponto um pouco deslocado no canto superior esquerdo. A partir do gráfico de *threshold*, sensibilidade e especificidade, o limite para as classes 1, 2 e 3 foram de 0,29, 0,28 e 0,56 respectivamente.

**Figura 20.** ROC dos modelos PLS-DA para teor de biodiesel metílico de Crambe

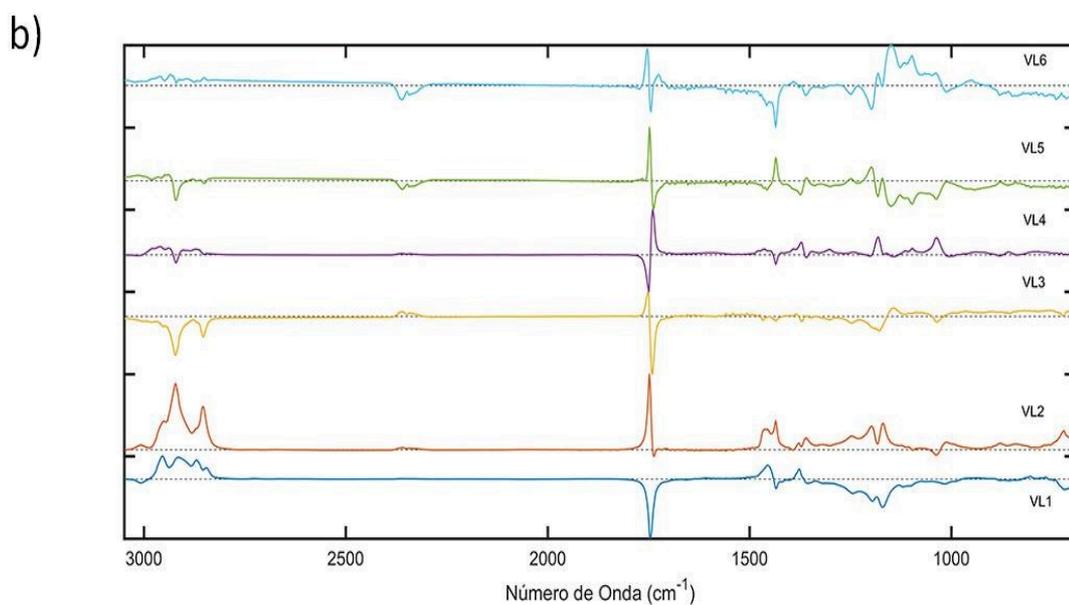
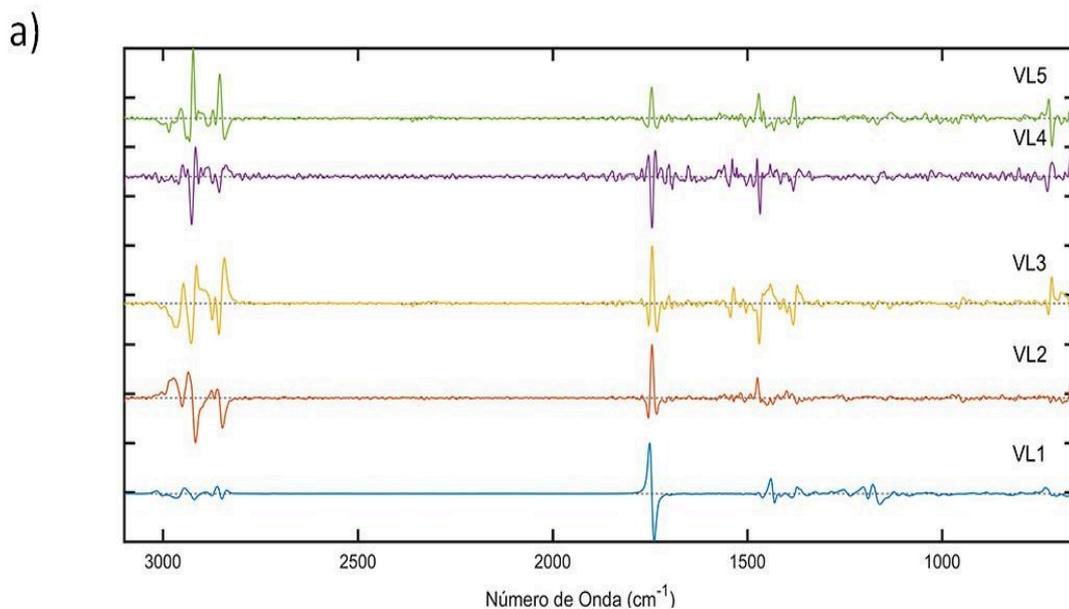


**Fonte:** O autor,2018.

Os pesos, isto é, a importância das variáveis originais em cada variável latente (VL) são mostrados na Figura 21a (biodiesel metílico de Pinhão Manso). As bandas situadas entre 2.940 a 2.840  $\text{cm}^{-1}$  se mostraram importantes a partir da VL2, enquanto não foi tão importante para a VL1. A banda de carbonila (em torno de 1.740  $\text{cm}^{-1}$ ) foi importante em todas as VL. As bandas entre 800-700  $\text{cm}^{-1}$ , identificadas como características do diesel, foram importantes na VL3 a VL5.

Os pesos, para o modelo de biodiesel metílico de Crambe são mostrados na Figura 21b. As bandas situadas entre 2.940 a 2.840  $\text{cm}^{-1}$  se mostraram importantes a partir da VL1 e VL2. A banda de carbonila (em torno de 1.740  $\text{cm}^{-1}$ ) foi importante em todas as VL. As bandas entre 800-700  $\text{cm}^{-1}$ , identificadas como características do diesel, foram importantes na VL1 e VL2.

**Figura 21.** Pesos dos modelos PLS-DA para teor de biodiesel metílico de (a) Pinhão Manso e (b) Crambe



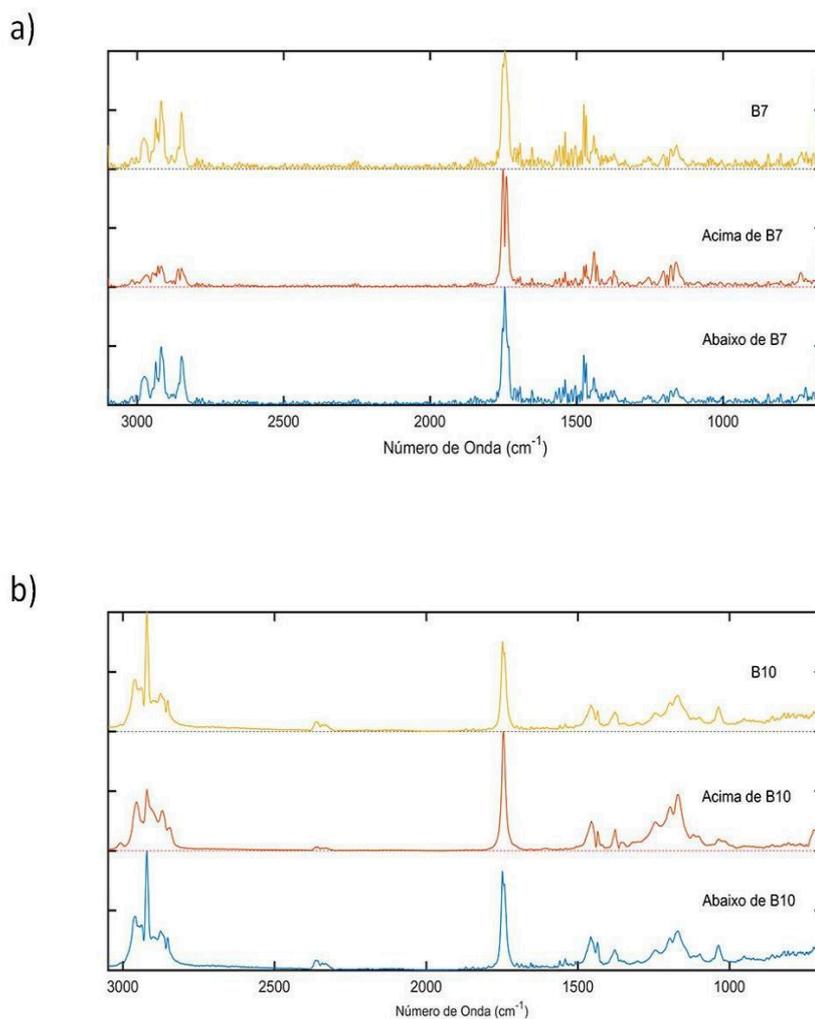
Fonte: O autor, 2018.

Para avaliar quais as variáveis tiveram mais influência na separação das amostras, foi usada a importância da variável na projeção (VIP), cujo gráfico é mostrado na Figura 22. Para o modelo de biodiesel metílico de Pinhão Manso (Figura 22a), os números de onda cujos escores VIP são superiores a 1 são importantes para o modelo PLS-DA (ALMEIDA et al., 2013). A região de carbonila se mostrou a mais importante para esta discriminação, embora algumas regiões relacionadas com o diesel também se mostraram importantes para o modelo.

A VIP para o modelo de biodiesel metílico de Crambe é mostrada no gráfico (Figura 22b). A região situada entre  $2940$  a  $2840 \text{ cm}^{-1}$  mostrou-se importante nos espectros abaixo de B10 e no B10. No entanto, a região de carbonila continua sendo a mais importante para esta

discriminação, além de algumas regiões relacionadas com o diesel também apareceram para este modelo.

**Figura 22.** VIP dos modelos PLS-DA para teor de biodiesel metílico de (a) Pinhão Manso e (b) Crambe



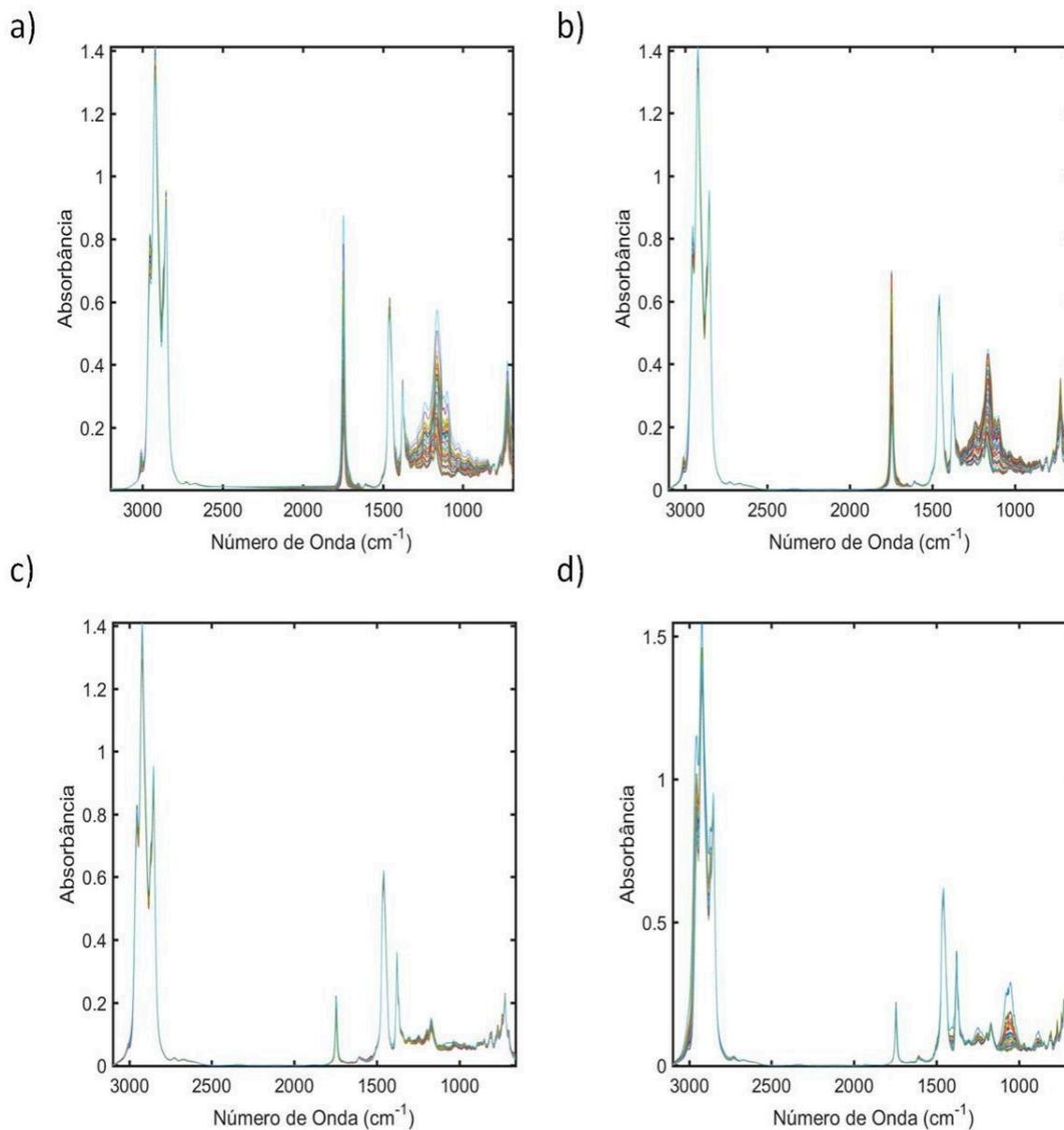
Fonte: O autor, 2018.

#### 4.2.2 Adulteração da mistura biodiesel de Pinhão Manso/diesel usando óleo de Soja, óleo residual, óleo automotivo usado e gasolina

A Figura 23 mostra os espectros MIR das amostras de biodiesel/diesel de pinhão adulterado com óleos e solventes. Analisando as figuras 23a (adulterada com óleo de Soja) e 23b (adulterada com óleo residual), os padrões de absorção dos espectros de biodieseis e óleos diferem em maior grau na região de impressão digital, inferior a  $1.500\text{ cm}^{-1}$ .

Como foi discutida anteriormente, a banda ao redor de  $1.655\text{ cm}^{-1}$  pode ser atribuída às vibrações de deformação axial da ligação C=C. Observa-se a presença de uma banda de absorbância na faixa de  $1.750\text{ cm}^{-1}$  a  $1.700\text{ cm}^{-1}$ , proveniente da presença de carbonila (C = O) e a presença de bandas na região de  $1.300\text{ cm}^{-1}$  a  $1.000\text{ cm}^{-1}$  atribuída às vibrações de deformação axial da ligação C-O. Na figura 23c (adulterado com óleo automotivo usado) e 23d (adulterado com gasolina), observa-se a presença de um pequeno banda de absorbância na faixa de  $1.750\text{ cm}^{-1}$  a  $1.700\text{ cm}^{-1}$ , proveniente da presença de carbonila (C = O). Esta banda é devida somente ao biodiesel, pois os adulterantes são derivados do petróleo não constituindo ligações (C=O). Além disso, na figura 23d observa-se um banda maior na região entre  $3000$  e  $3500\text{ cm}^{-1}$  referente à banda de OH presente no álcool constituinte da gasolina comercial.

**Figura 23.** Espectros MIR de amostras biodiesel/diesel de Pinhão Manso adulterado com (a) óleo Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina.

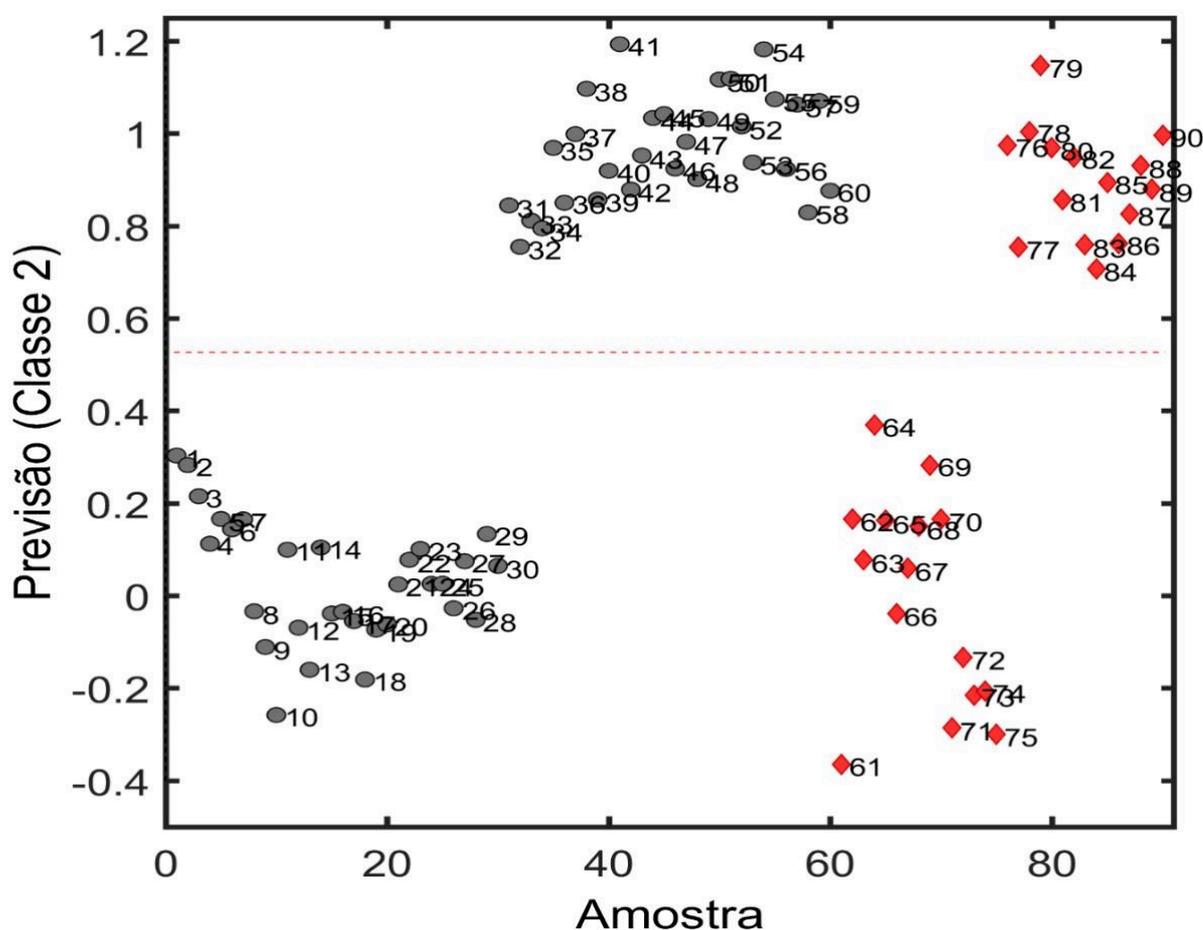


**Fonte:** O autor, 2018.

No caso de adulterações existem apenas duas classes: as amostras autênticas (B7 sem adulteração) e as amostras adulteradas. Nos modelos de biodiesel metílico de Pinhão Manso não foi realizado nenhum outro pré-processamento adicional, além da linha de base e centrar os dados na média. Além disso, não foi detectada amostra anômala em nenhum dos quatro modelos. Para todos os modelos, as amostras de 1 a 60 foram consideradas amostras de treinamento e as amostras de 61 a 90 (totalizando 30 amostras) pertencentes ao conjunto de teste.

O modelo de biodiesel metílico de Pinhão Manso adulterado com óleo de Soja foi construído com 7 variáveis latentes (VL) e capturou 99,98 e 93,60% de variância nos blocos X e Y, respectivamente. A Figura 24 mostra as estimativas para a classe das amostras autênticas. As amostras adulteradas (1-30 no treinamento e 61-75 no teste) foram previstas corretamente como não pertencentes à classe das autênticas. Portanto  $VN_{\text{treinamento}}$ ,  $VN_{\text{teste}}$ ,  $FP_{\text{treinamento}}$  e  $FP_{\text{teste}}$  é igual a 30, 15, 0 e 0, respectivamente. Já as amostras autênticas (31-60 no treinamento e 76-90 no teste) foram previstas corretamente como pertencentes a esta classe e, assim,  $VP_{\text{treinamento}}$ ,  $VP_{\text{teste}}$ ,  $FN_{\text{treinamento}}$  e  $FN_{\text{teste}}$  é igual a 30, 15, 0 e 0, respectivamente. Em outras palavras, todas as amostras tanto de treinamento quanto de teste foram previstas corretamente como pertencentes as suas classes.

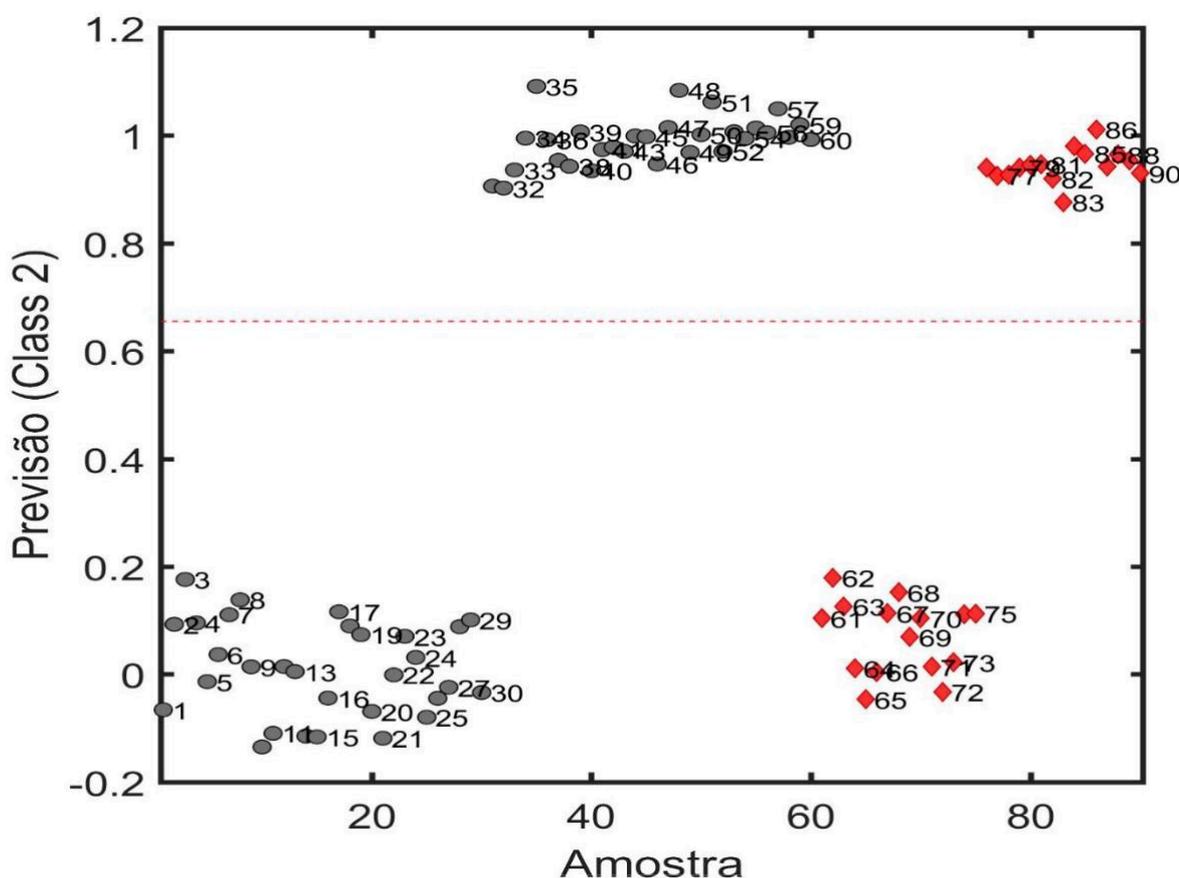
**Figura 24.** Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com óleo Soja



Fonte: O autor, 2018.

O modelo de biodiesel metílico de Pinhão Manso adulterado com óleo residual foi construído com 5 variáveis latentes (VL) e capturou 99,92, e 98,09% de variância nos blocos **X** e **Y**, respectivamente. Na Figura 25 as amostras adulteradas (1-30 no treinamento e 61-75 no teste) foram previstas corretamente como não-pertencentes à classe das autênticas. Portanto  $VN_{\text{treinamento}}$ ,  $VN_{\text{teste}}$ ,  $FP_{\text{treinamento}}$  e  $FP_{\text{teste}}$  é igual a 30, 15, 0 e 0, respectivamente. No caso das amostras autênticas (31-60 no treinamento e 76-90 no teste), elas foram previstas corretamente como pertencentes a esta classe e, assim,  $VP_{\text{treinamento}}$ ,  $VP_{\text{teste}}$ ,  $FN_{\text{treinamento}}$  e  $FN_{\text{teste}}$  é igual a 30, 15, 0 e 0, respectivamente. Assim, todas as amostras tanto de treinamento quanto teste foram previstas corretamente como pertencentes as suas classes assim como no primeiro modelo.

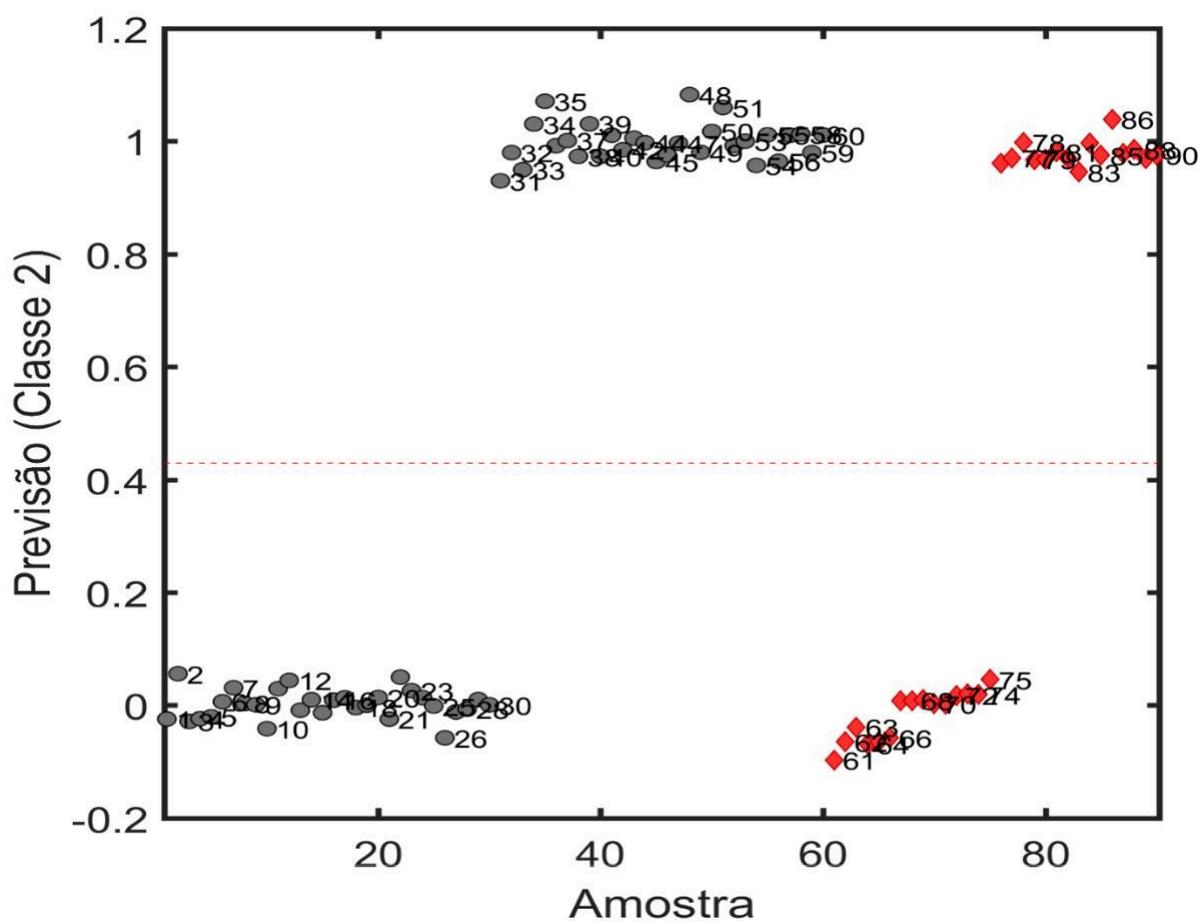
**Figura 25.** Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com óleo residual.



Fonte: O autor, 2018.

Para o modelo de biodiesel metílico de Pinhão Manso adulterado com óleo automotivo usado foi construído com 5 variáveis latentes (VL) e capturou 97,13 e 99,65% de variância nos blocos X e Y, respectivamente. A Figura 26 mostra as estimativas para a classe das autênticas. As amostras adulteradas (1-30 no treinamento e 61-75 no teste) foram previstas corretamente como não pertencentes à classe das autênticas e as amostras autênticas (31-60 no treinamento e 76-90 no teste) foram previstas corretamente como pertencentes a esta classe em que todas as amostras tanto de treinamento quanto de teste foram previstas corretamente como pertencentes as suas classes.

**Figura 26.** Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com óleo automotivo usado.

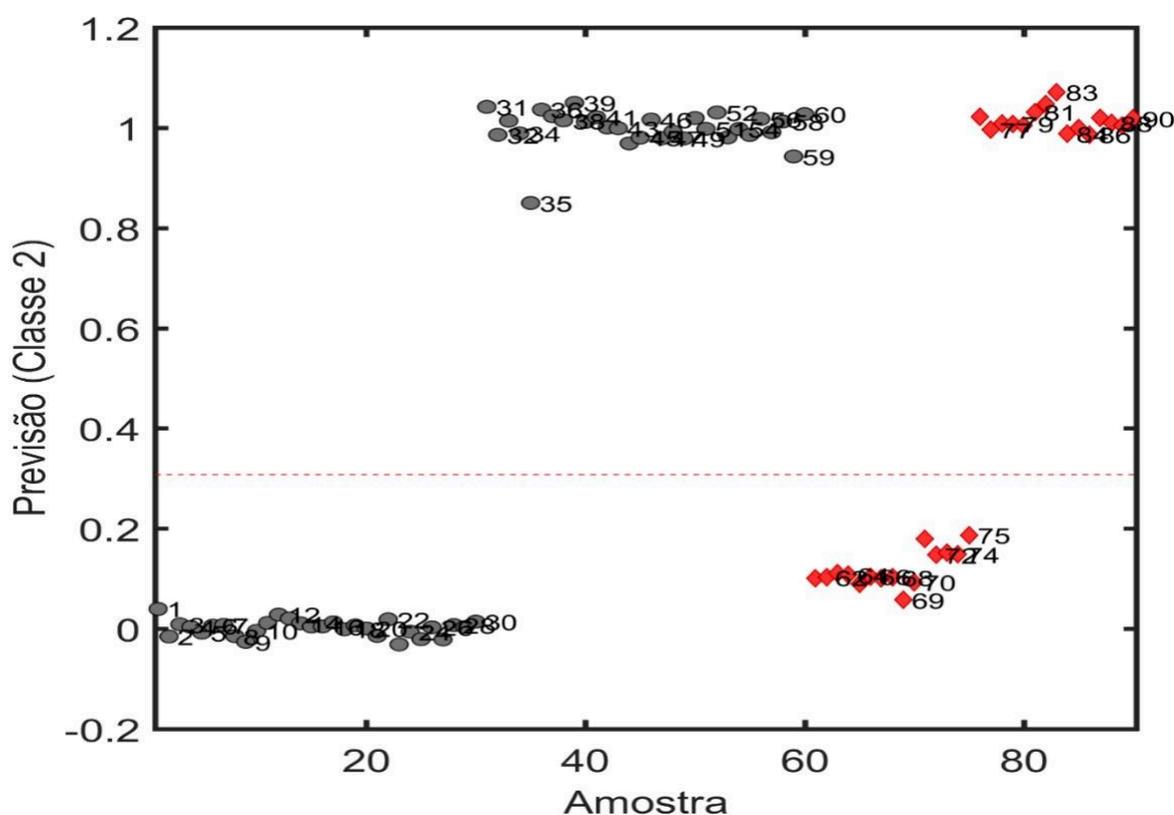


Fonte: O autor, 2018.

O modelo de biodiesel metílico de Pinhão Manso adulterado com gasolina foi construído com 5 variáveis latentes (VL) e capturou 99,79 e 99,69% de variância nos blocos X e Y, respectivamente. As amostras adulteradas (1-30 no treinamento e 61-75 no teste) foram previstas corretamente como não pertencentes à classe das autênticas na Figura 27.

Portanto  $VN_{\text{treinamento}}$ ,  $VN_{\text{teste}}$ ,  $FP_{\text{treinamento}}$  e  $FP_{\text{teste}}$  é igual a 30, 15, 0 e 0, respectivamente. Já as amostras autênticas (31-60 no treinamento e 76-90 no teste) foram previstas corretamente como pertencentes a esta classe e, assim,  $VP_{\text{treinamento}}$ ,  $VP_{\text{teste}}$ ,  $FN_{\text{treinamento}}$  e  $FN_{\text{teste}}$  é igual a 30, 15, 0 e 0, respectivamente. Desta maneira, todas as amostras tanto de treinamento quanto de teste foram previstas corretamente como pertencentes as suas classes.

**Figura 27.** Estimativas das amostras autênticas de biodiesel/diesel de Pinhão Manso para o conjunto de adulteração com gasolina.



Fonte: O autor, 2018.

A tabela de confusão obtida para os modelos de biodiesel metílico de Pinhão Manso adulterado com óleo de Soja, residual na Tabela 4. Como apresentaram os gráficos de estimativas, todas as amostras autênticas foram classificadas corretamente como autênticas e todas as amostras adulteradas foram classificadas como adulteradas para todos os modelos.

**Tabela 4.** Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Pinhão Manso com óleo de Soja e óleo residual.

Adulterante	Óleo de Soja				Óleo residual			
	Treinamento		Teste		Treinamento		Teste	
	Real Classe		Real Classe		Real Classe		Real Classe	
Classe Adultera da	Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe
	Autêntica	Adulterada	Autêntica	Adulterada	Autêntica	Adulterada	Autêntica	Adulterada
Previsto como Classe Adulterada	30	0	15	0	30	0	15	0
Previsto como Classe Autêntica	0	30	0	15	0	30	0	15

Fonte: O autor, 2018.

A tabela de confusão obtida para os modelos de biodiesel metílico de Pinhão Manso adulterados com óleo automotivo usado (Tabela 5) é mostrada abaixo.

**Tabela 5.** Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Pinhão Manso com óleo automotivo usado e gasolina.

adulterante	Óleo automotivo usado				gasolina			
	Treinamento		Teste		Treinamento		Teste	
	Real Classe		Real Classe		Real Classe		Real Classe	
Classe Adulterada	Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe
	Autêntica	Adulterada	Autêntica	Adulterada	Autêntica	Adulterada	Autêntica	Adulterada
Previsto como Classe Adulterada	30	0	15	0	30	0	15	0
Previsto como Classe Autêntica	0	30	0	15	0	30	0	15

Fonte: O autor, 2018.

Os resultados para as 2 classes, tanto no conjunto de treinamento quanto de teste são mostrados na Tabela 6. Esses resultados com baixos valores de RMSEC (inferior a 0,127), RMSECV (inferior a 0,190), RMSEP (inferior a 0,193) sensibilidade e especificidade igual a 1,000, mostram que os modelos construídos para a classificação das amostras adulteradas e autênticas foram eficientes na separação destas.

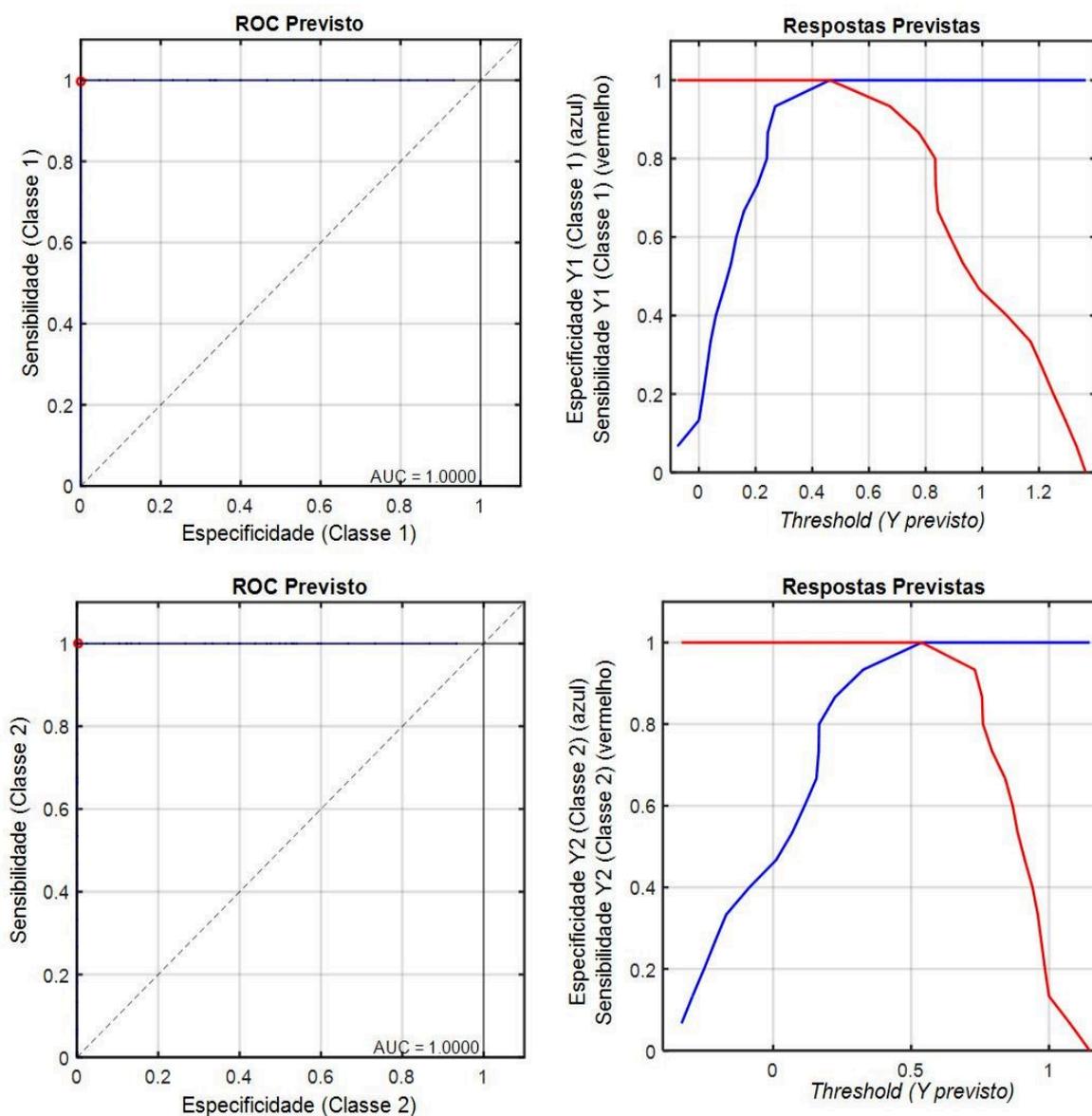
**Tabela 6.** Parâmetros de classificação obtidos pelo modelo PLS-DA para adulteração de biodiesel metílico de Pinhão Manso adulterado com: óleo de Soja, óleo residual, óleo automotivo usado e gasolina

Adulterante	Óleo de Soja		Óleo residual		Óleo automotivo usado		Gasolina	
	Classe 1	Classe 2	Classe 1	Classe 2	Classe 1	Classe 2	Classe 1	Classe 2
RMSECV (% v/v)	0,19	0,19	0,11	0,11	0,04	0,04	0,06	0,06
RMSEC (% v/v)	0,13	0,13	0,07	0,07	0,03	0,03	0,03	0,03
RMSEP (% v/v)	0,19	0,19	0,08	0,08	0,04	0,04	0,09	0,09
Especificidade <sub>treinamento</sub>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Especificidade <sub>teste</sub>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Sensibilidade <sub>treinamento</sub>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Sensibilidade <sub>teste</sub>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

**Fonte:** O autor, 2018.

As curvas ROC e a relação de especificidade, sensibilidade e *threshold* no modelo de biodiesel metílico de Pinhão Manso adulterado com óleo de Soja são mostradas na Figura 28. Como a taxa de classificação correta foi de 100%, AUC em ambos os casos foi igual a 1,0 e o máximo de sensibilidade e especificidade foi alcançado no canto superior esquerdo. A partir do gráfico de *threshold*, sensibilidade e especificidade (Figura 23a), o limite para as classes 1 e 2 foram de 0,46 e 0,54, respectivamente, para o modelo de biodiesel metílico de Pinhão Manso adulterado com óleo de Soja. Portanto, as amostras cujo Y1 é previsto acima de 0,46 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,54 são consideradas como autênticas.

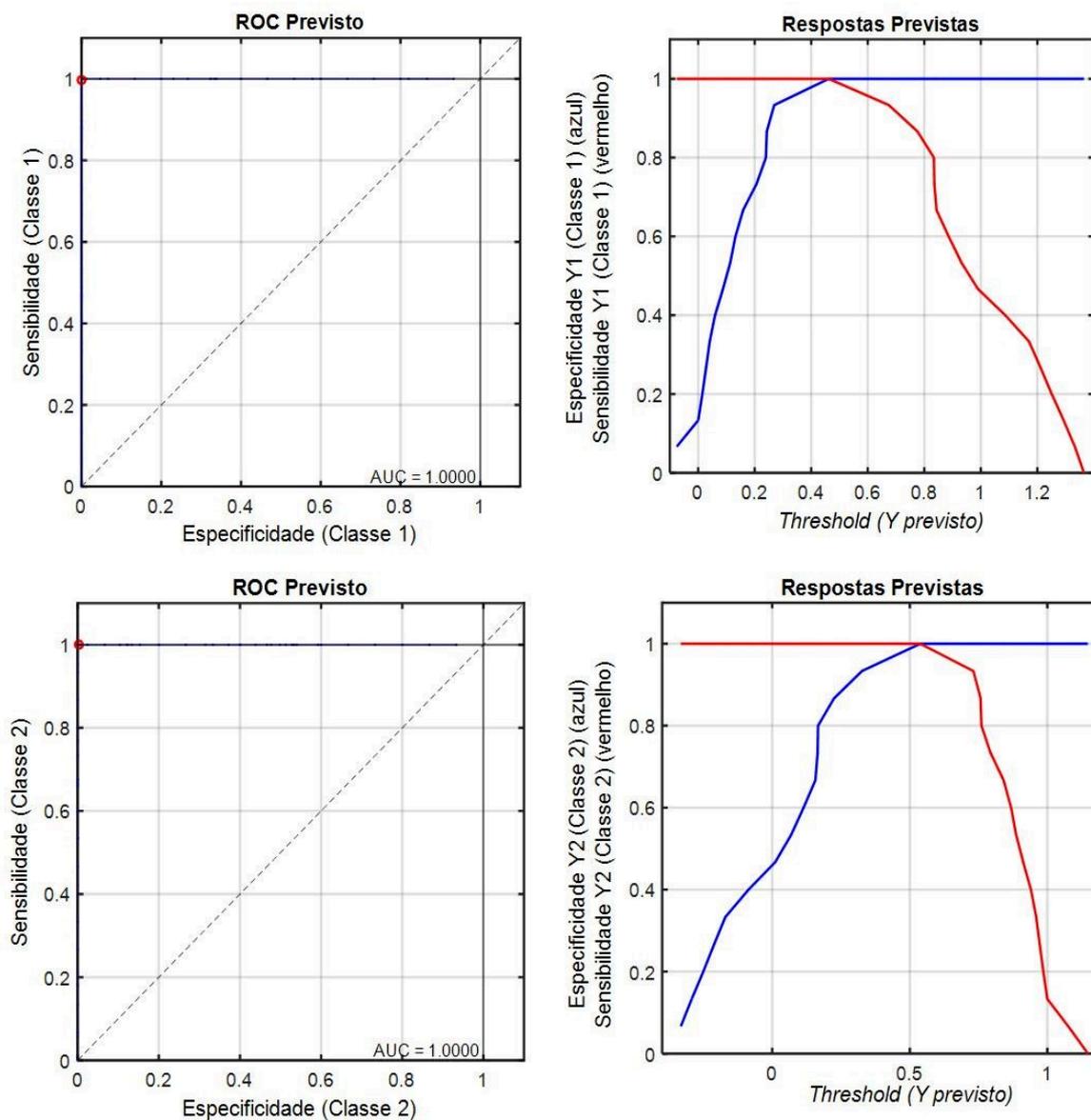
**Figura 28.** ROC das amostras de biodiesel/biodiesel de Pinhão Manso para as classes adulteradas (classe1) e autênticas (classe 2) para o conjunto de adulteração óleo de Soja



Fonte: O autor, 2018.

Para o modelo de biodiesel metílico adulterado com óleo residual, a sensibilidade e especificidade (Figura 29), o limite para as classes 1 e 2 foram de 0,45 e 0,52, respectivamente. Portanto, as amostras cujo Y1 é previsto acima de 0,45 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,52 são consideradas como autênticas.

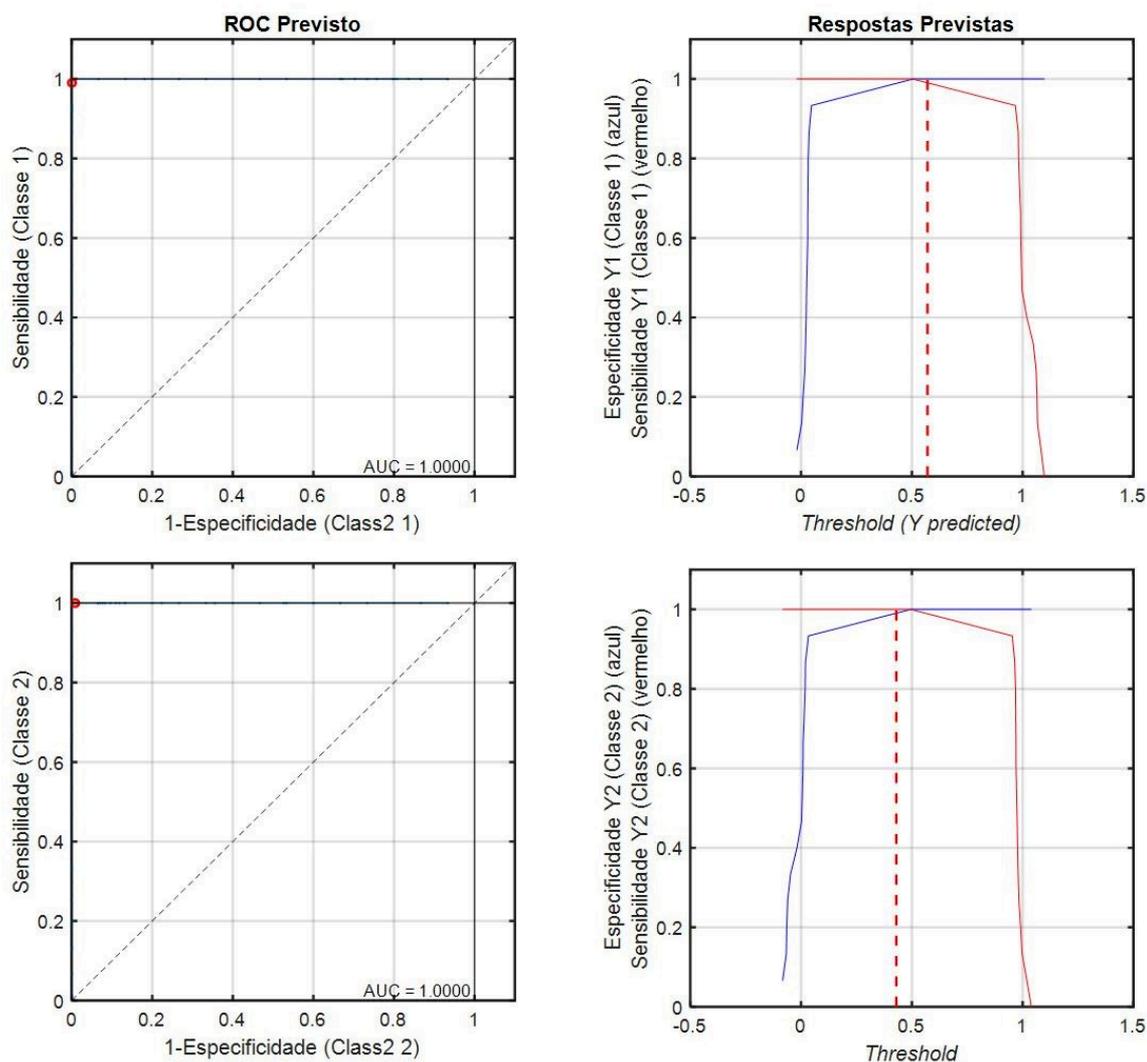
**Figura 29.** ROC das amostras biodiesel/diesel de Pinhão Manso para as classes adulteradas (classe1) e autênticas (classe2) para o conjunto de adulteração com óleo residual



Fonte: O autor, 2018.

O modelo de biodiesel metílico adulterado com óleo automotivo usado, a sensibilidade e especificidade (Figura 30), teve o seguinte resultado: o limite para as classes 1 e 2 foram de 0,58 e 0,45, respectivamente. Portanto, as amostras cujo Y1 é previsto acima de 0,58 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,45 são consideradas como autênticas.

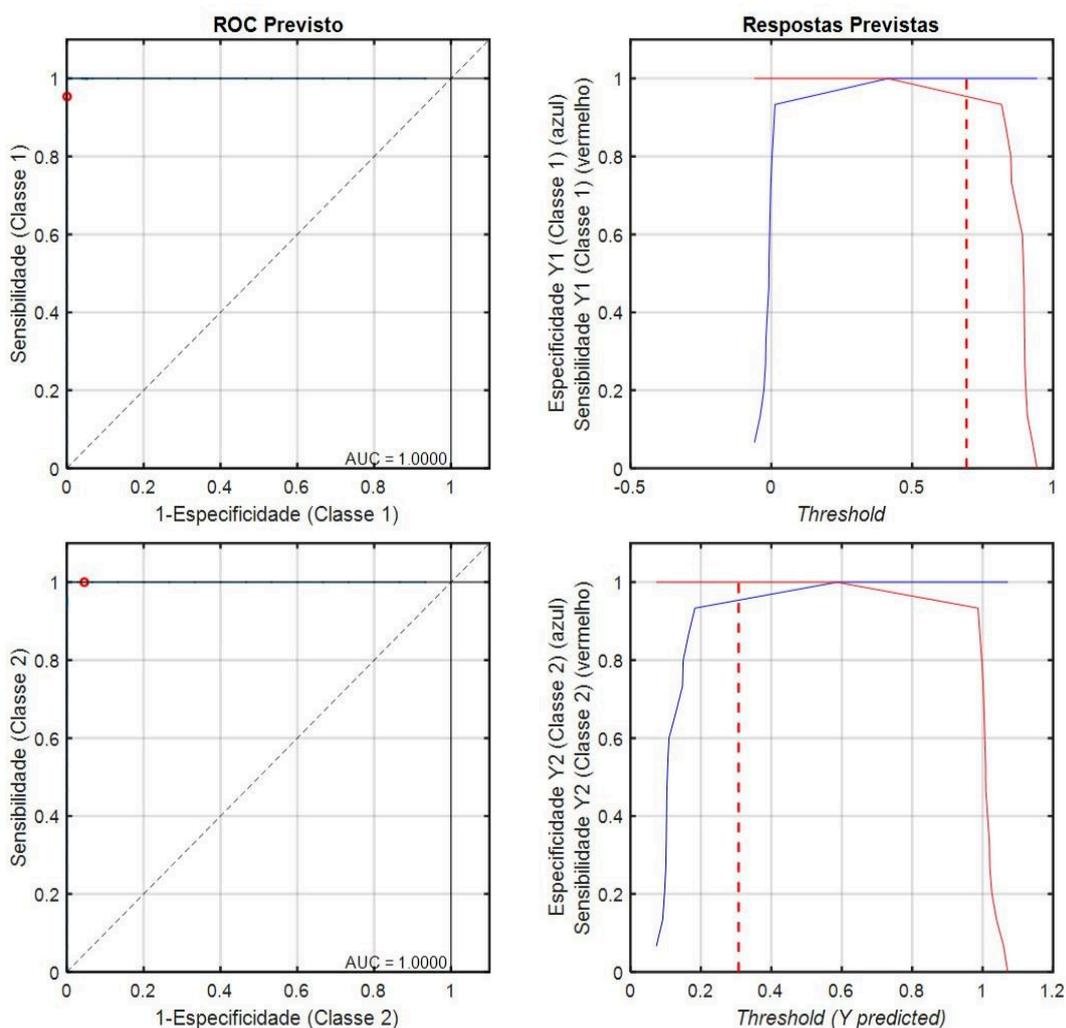
**Figura 30.** ROC das amostras biodiesel /diesel de Pinhão Manso para as classes adulteradas (classe1) e autênticas (classe2) para o conjunto de adulteração com óleo automotivo usado.



Fonte: O autor, 2018

Já o modelo de biodiesel metílico adulterado com gasolina, a sensibilidade e especificidade (Figura 31), o limite para as classes 1 e 2 foram de 0,70 e 0,32, respectivamente. Portanto, as amostras cujo Y1 é previsto acima de 0,70 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,32 são consideradas como autênticas.

**Figura 31.** ROC das amostras biodiesel/diesel de Pinhão Manso para as classes adulteradas (classe1) e autênticas (classe2) para o conjunto de adulteração com gasolina

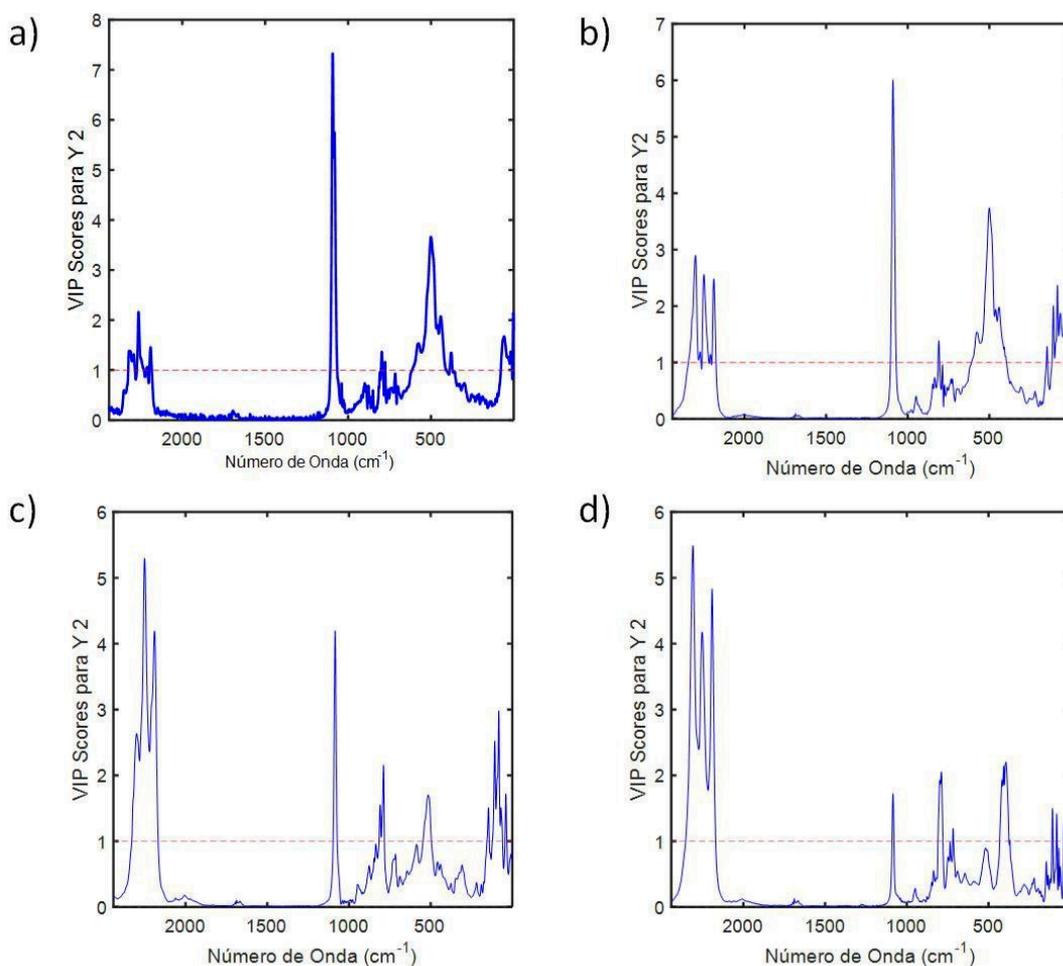


**Fonte:** O autor, 2018.

Assim como nos modelos anteriores, os escores VIP foram usados para avaliar quais as variáveis tiveram mais influência na separação das amostras. Os escores VIP para o modelo com adulteração são significativamente diferentes do que o obtido para os modelos de teor de biodiesel de Pinhão Manso e Crambe. A região da carbonila foi a mais importante em ambos os modelos, mas no caso da adulteração com óleo de Soja (Figura 32a) e com óleo residual (Figura 32b), a região em torno de  $1.300$  a  $1.000\text{ cm}^{-1}$  desempenhou um papel significativo nessa distinção. Essa parte do espectro MIR está associada com estiramentos da ligação C-O e a variação nessa região pode ser associada a maior quantidade de óleo e consequente diminuição de diesel, aumentando assim os compostos com oxigênio nas misturas adulteradas.

Para o modelo de biodiesel metílico de Pinhão Manso adulterado com óleo automotivo usado, destaca-se a região conhecida como impressão digital correspondendo à faixa de 1.200-600  $\text{cm}^{-1}$  (Figura 32c), observa-se a presença de um banda bem menor de absorbância na faixa de 1.750  $\text{cm}^{-1}$  a 1.700  $\text{cm}^{-1}$ , proveniente da presença de carbonila ( $\text{C}=\text{O}$ ). Esta banda é devida somente ao biodiesel, pois os adulterantes são derivados do petróleo não constituindo ligações ( $\text{C}=\text{O}$ ). Além disso, a região significativa de derivados do petróleo como o óleo automotivo usado (Figura 33c) e gasolina (Figura 32d) é a região entre 3000 e 2840  $\text{cm}^{-1}$  atribuída às vibrações de estiramento das ligações C-H dos grupamentos de metila e metileno (SKOOG, 2002). Além disso, nota-se na Figura 32d um banda maior na região entre 3500 e 3000  $\text{cm}^{-1}$  referente à banda de OH presente no álcool constituinte da gasolina comercial.

**Figura 32.** VIP para PLS-DA das amostras biodiesel/diesel de Pinhão Manso para os conjuntos de adulteração com (a) óleo de Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina.

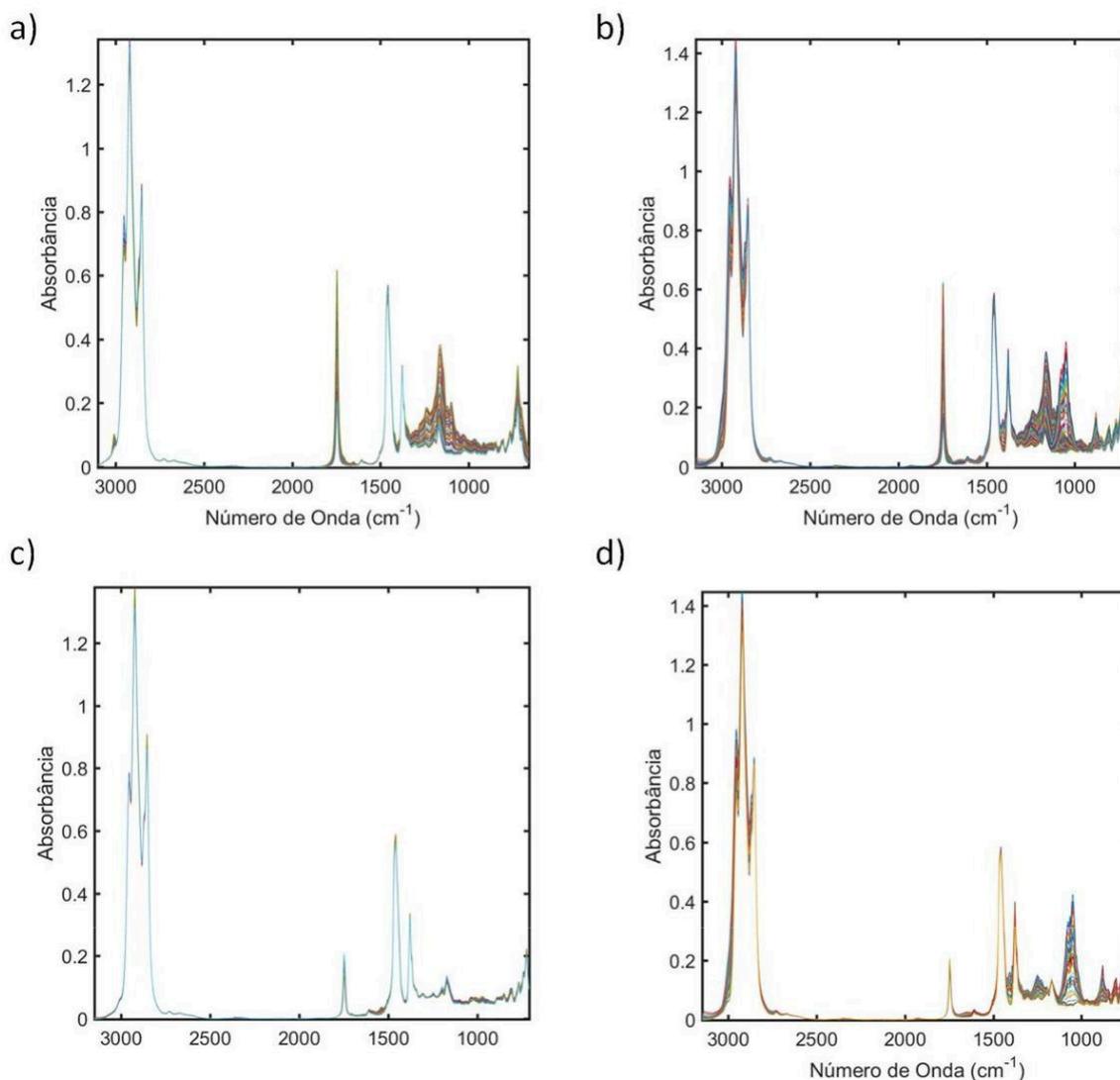


Fonte: O autor, 2018.

### 4.2.3 Adulteração da mistura biodiesel de Crambe/diesel usando óleo de Soja, óleo residual, óleo automotivo usado e gasolina

A Figura 33 mostra os espectros MIR das amostras de biodiesel/diesel de Crambe adulterados com óleos e solventes. De acordo com as figuras 33a (adulterada com óleo de Soja) e 33b (adulterada com óleo residual), os padrões de absorção dos espectros de biodieseis e óleos diferem em maior grau na região de impressão digital, inferior a  $1.500\text{ cm}^{-1}$ . Além disso, se diferem as bandas ao redor de  $1.655\text{ cm}^{-1}$  atribuída às vibrações de deformação axial da ligação C=C, de  $1.750\text{ cm}^{-1}$  a  $1.700\text{ cm}^{-1}$ , proveniente da presença de carbonila (C = O) e na região de  $1.300\text{ cm}^{-1}$  a  $1.000\text{ cm}^{-1}$  atribuída às vibrações de deformação axial da ligação C-O. Na figura 33c (adulterado com óleo automotivo usado) e 33d (adulterado com gasolina), destaca-se a região entre  $3000$  a  $2840\text{ cm}^{-1}$  referentes aos derivados do petróleo e na Figura 33d uma banda maior na região entre  $3500$  e  $3000\text{ cm}^{-1}$  referente à OH presente no álcool misturado com a gasolina.

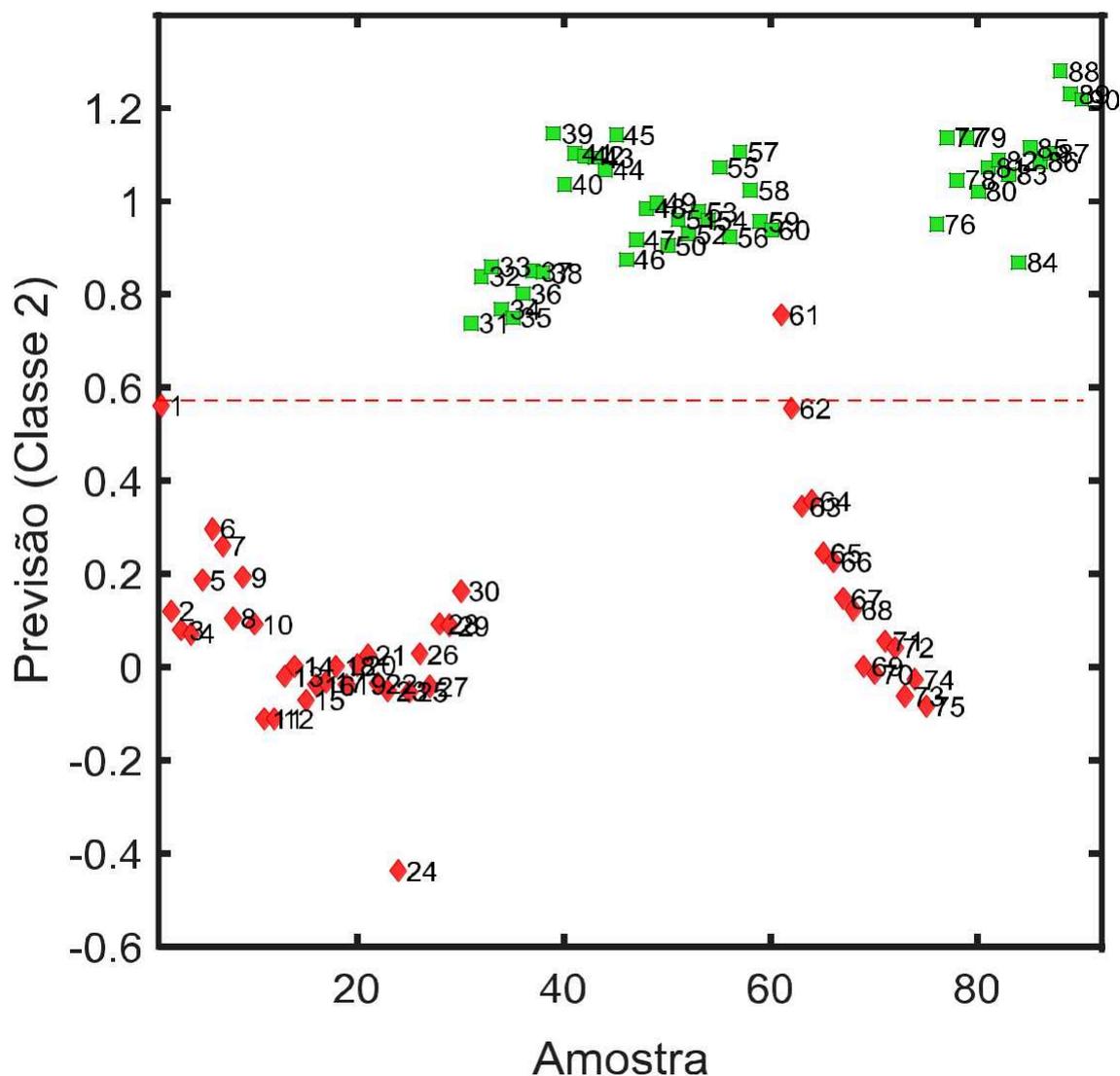
**Figura 33.** Espectros MIR de amostras biodiesel/diesel de Crambe adulteradas com: a) óleo Soja, b) óleo residual, c) óleo automotivo usado e d) gasolina.



Fonte: O autor, 2018.

O modelo de biodiesel metílico de Crambe adulterado com óleo de Soja foi construído com 5 variáveis latentes (VL) e capturou 99,90 e 91,14% de variância nos blocos X e Y, respectivamente. Nenhuma amostra *outlier* foi detectada nesse modelo. As amostras de 1 a 60 são as amostras de treinamento e as amostras de 61 a 90 (totalizando 30 amostras) pertencem ao conjunto de teste. A Figura 34 mostra as estimativas para a classe das autênticas. Uma das amostras adulteradas (amostra 61, com, 1,50% (v/v) de óleo de Soja) foi prevista erroneamente como pertencente à classe das amostras autênticas pois ficou acima do *threshold*. As amostras 1 e 60, embora próximas do limite, foram previstas corretamente como pertencentes à classe 1 (amostras adulteradas com óleo de Soja).

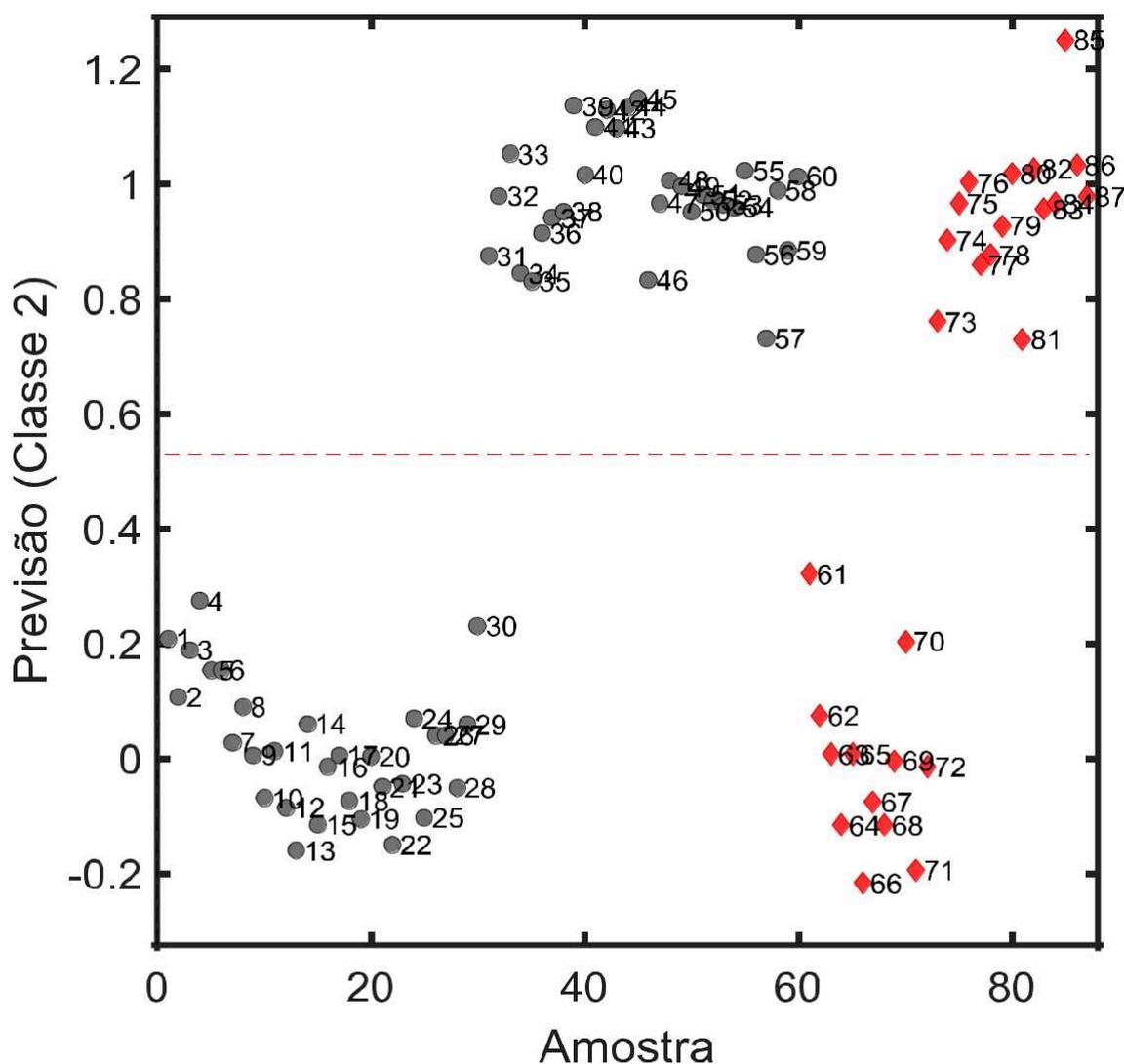
**Figura 34.** Estimativas das amostras autênticas de biodiesel/diesel de Crambe para o conjunto de adulteração com óleo de Soja.



Fonte: O autor, 2018.

O modelo de biodiesel metílico de Crambe adulterado com óleo residual foi construído com 5 variáveis latentes (VL) e capturou 99,91 e 95,23% de variância nos blocos X e Y, respectivamente. Nenhuma amostra anômala (*outlier*) foi detectada nesse modelo. As amostras de 1 a 60 são as amostras de treinamento e as amostras de 61 a 90 (totalizando 30 amostras) pertencem ao conjunto de teste. A Figura 35 mostra as estimativas para a classe das autênticas. Todas as amostras foram previstas corretamente como pertencentes à classe 1 (amostras adulteradas com óleo residual).

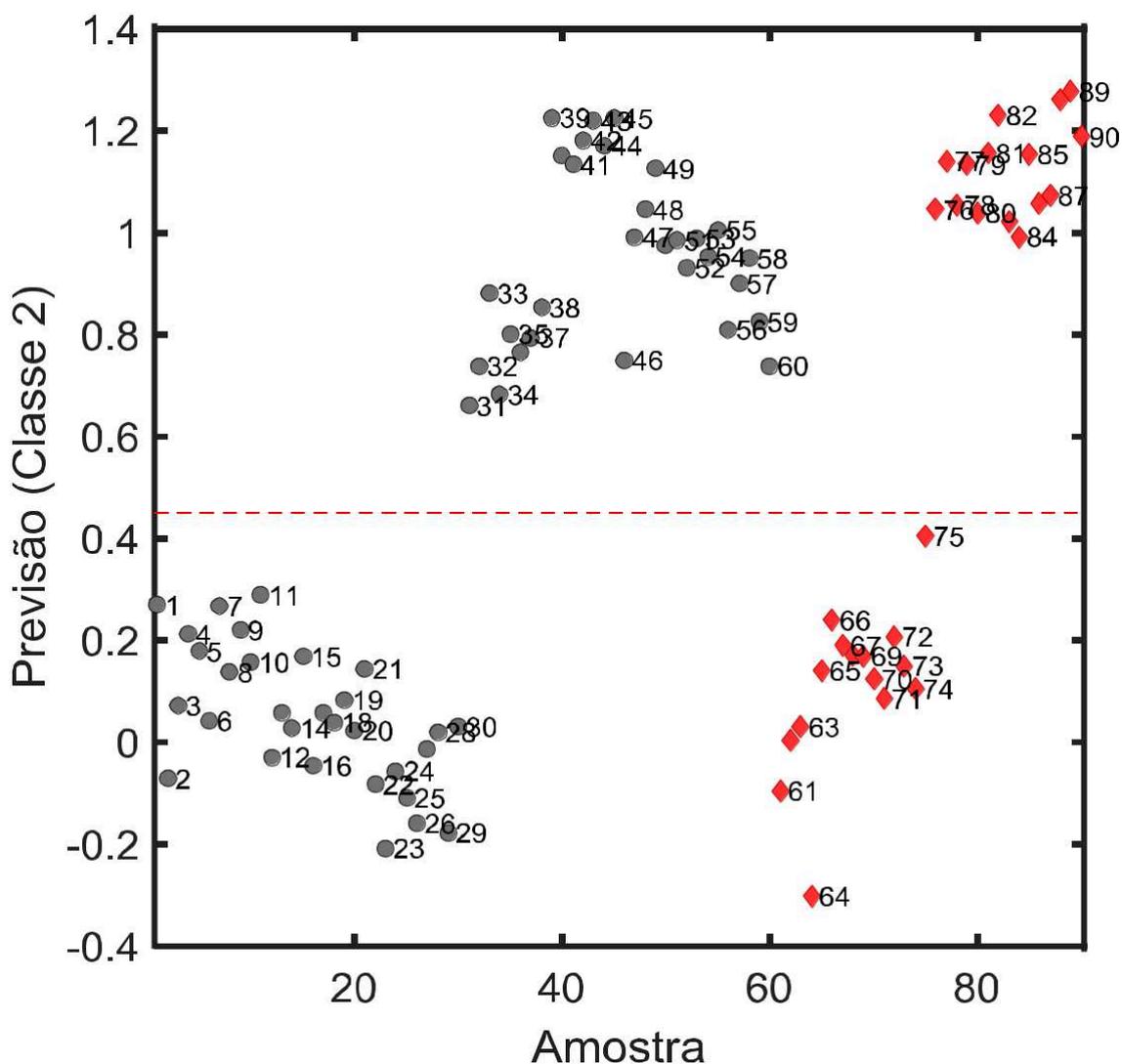
**Figura 35.** Estimativas das amostras autênticas de biodiesel/diesel de Crambe para o conjunto de adulteração com óleo residual.



**Fonte:** O autor, 2018.

No modelo de biodiesel metílico de Crambe adulterado com óleo automotivo usado foi construído com 4 variáveis latentes (VL) e capturou 96,52 e 89,70% de variância nos blocos **X** e **Y**, respectivamente. Nenhuma amostra *outlier* foi detectada nesse modelo. As amostras de 1 a 60 são as amostras de treinamento e as amostras de 61 a 90 (totalizando 30 amostras) pertencem ao conjunto de teste. A Figura 36 mostra as estimativas para a classe das autênticas.

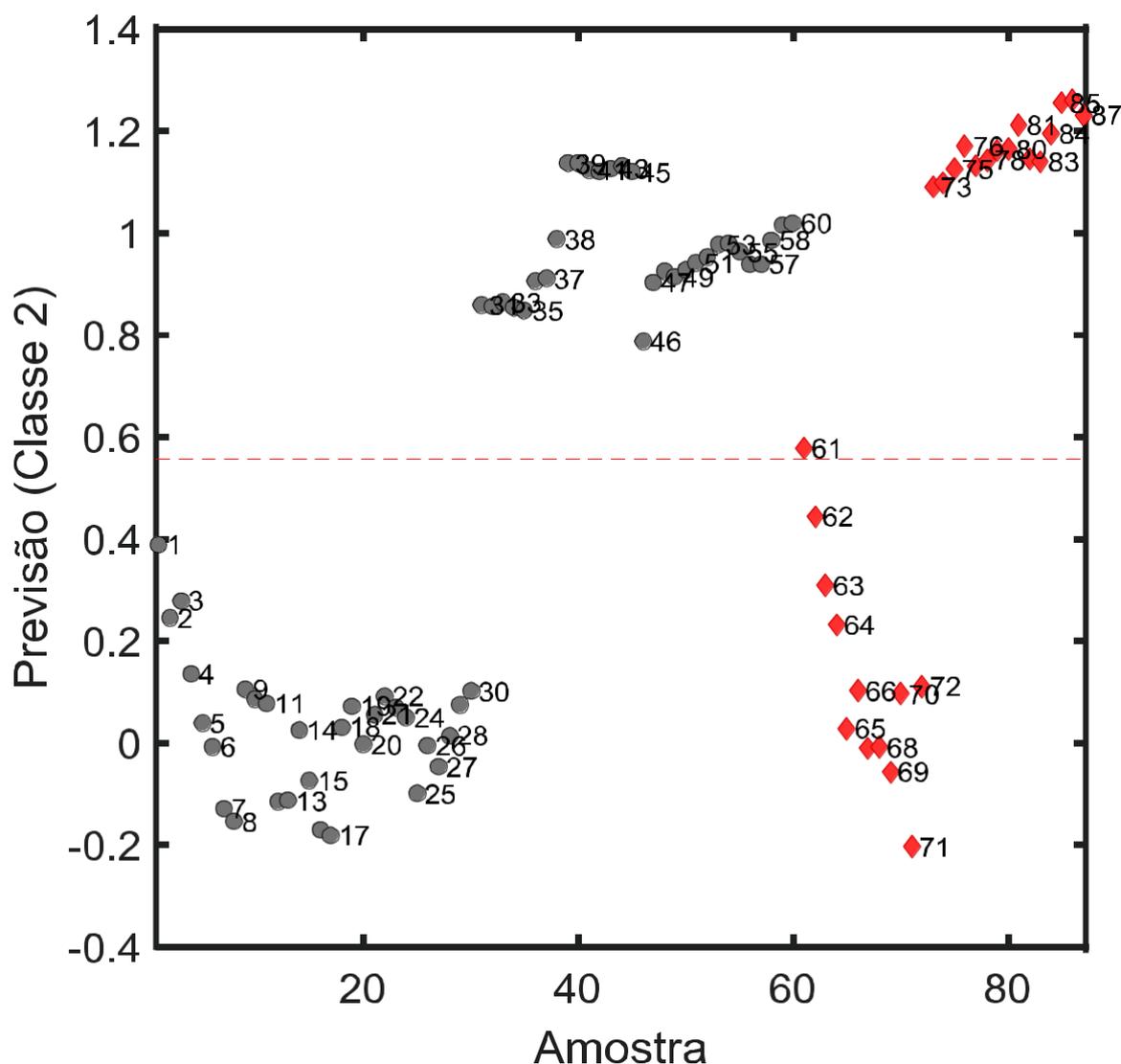
**Figura 36.** Estimativas das amostras autênticas de biodiesel/diesel de Crambe para o conjunto de adulteração com óleo automotivo usado.



Fonte: O autor, 2018.

A construção do modelo de biodiesel metílico de Crambe adulterado com gasolina foi realizada com 6 variáveis latentes (VL) e capturou 99,89 e 94,36% de variância nos blocos **X** e **Y**, respectivamente. Nenhuma amostra outlier foi detectada nesse modelo. As amostras de 1 a 60 são as amostras de treinamento e as amostras de 61 a 90 (totalizando 30 amostras) pertencem ao conjunto de teste. Uma das amostras adulteradas (amostra 61, adulterada com gasolina) foi prevista erroneamente como pertencente à classe das amostras autênticas pois ficou acima do threshold. A Figura 37 mostra as estimativas para a classe das autênticas.

**Figura 37.** Estimativas das amostras autênticas de biodiesel de Crambe/diesel para o conjunto de adulteração com gasolina.



Fonte: O autor, 2018.

De acordo com a tabela de confusão obtida para todos os modelos de biodiesel metílico de Crambe, observa-se que para os modelos de biodiesel metílico de Crambe adulterado com óleo de Soja (Tabela 7) e gasolina (Tabela 8) tiveram apenas uma das amostras adulteradas foi classificada erroneamente como amostra autêntica para cada um dos dois modelos no conjunto teste. Já os modelos de biodiesel metílico de Crambe adulterado com óleo residual (Tabela 7) e óleo automotivo usado (Tabela 8) não houve amostras autênticas classificadas como adulteradas e nem amostras adulteradas classificadas como autênticas tanto para o conjunto de treinamento como para o conjunto teste.

**Tabela 7.** Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Crambe com (a) óleo de Soja e (b) óleo residual.

adulterante	Óleo de Soja				Óleo residual			
	Treinamento		Teste		Treinamento		Teste	
	Real Classe		Real Classe		Real Classe		Real Classe	
	Classe Adulterada	Classe Autêntica						
<b>Previsto como Classe Adulterada</b>	30	0	14	1	30	0	15	0
<b>Previsto como Classe Autêntica</b>	0	30	0	15	0	30	0	15

Fonte: O autor, 2018.

**Tabela 8.** Tabela de confusão para os resultados do modelo PLS-DA para adulteração de biodiesel metílico de Crambe com (a) óleo automotivo usado e (b) gasolina.

adulterante	Óleo automotivo usado				gasolina			
	Treinamento		Teste		Treinamento		Teste	
	Real Classe		Real Classe		Real Classe		Real Classe	
	Classe Adulterada	Classe Autêntica	Classe Adulterada	Classe Autêntica	Classe Adulterada	Classe Autêntica	Classe Adulterada	Classe Autêntica
<b>Previsto como Classe Adulterada</b>	30	0	15	0	30	0	14	1
<b>Previsto como Classe Autêntica</b>	0	30	0	15	0	30	0	15

Fonte: O autor, 2018.

Os resultados para as duas classes, tanto no conjunto de treinamento quanto no de teste são mostrados na Tabela 9. Esses resultados com baixos valores de RMSEC ( $< 0,16\%$ ), RMSECV ( $< 0,20\%$ ), RMSEP ( $< 0,23\%$ ), sensibilidade e especificidade acima de 0,93 mostram que para todos os modelos de biodiesel metílico de Crambe adulterado com óleo de Soja, óleo residual, óleo automotivo usado e gasolina construídos para a classificação das amostras B10 foram eficientes para a distinção de amostras autênticas e adulteradas.

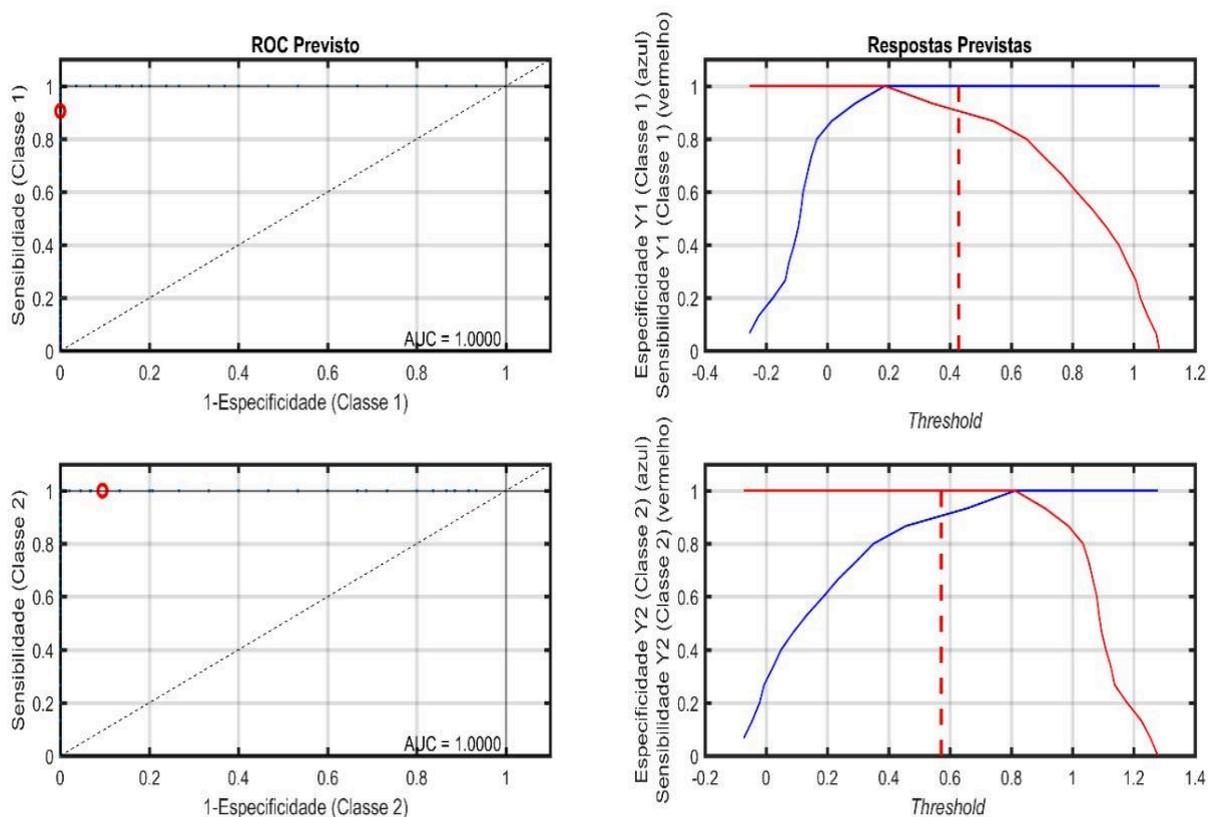
**Tabela 9** – Parâmetros de classificação obtidos pelo modelo PLS-DA para adulteração de biodiesel metílico de Crambe adulterado com: óleo de Soja, óleo residual, óleo automotivo usado e gasolina

Oleo Adulterante	Óleo de Soja		Óleo residual		óleo automotivo usado		gasolina	
	Classe 1	Classe 2	Classe 1	Classe 2	Classe 1	Classe 2	Classe 1	Classe 2
Parâmetros								
RMSECV (% v/v)	0,17	0,17	0,15	0,15	0,20	0,20	0,18	0,18
RMSEC (% v/v)	0,15	0,15	0,11	0,11	0,16	0,16	0,12	0,12
RMSEP (% v/v)	0,23	0,23	0,14	0,14	0,17	0,17	0,21	0,21
Especificidade <sub>treinamento</sub>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Especificidade <sub>teste</sub>	1,00	0,93	1,00	1,00	1,00	1,00	1,00	0,97
Sensibilidade <sub>treinamento</sub>	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Sensibilidade <sub>teste</sub>	0,93	1,00	1,00	1,00	1,00	1,00	1,00	0,97

**Fonte:** O autor, 2018.

As curvas ROC e a relação de especificidade, sensibilidade e *threshold* dos modelos são mostradas nas Figuras 38. Devido ao erro na previsão, o ponto de sensibilidade e especificidade do modelo de biodiesel metílico de Crambe adulterado com óleo de Soja sofreu um pequeno deslocamento. Pela Figura 38, observa-se que o limite para as classes 1 e 2 foram de 0,43 e 0,57, respectivamente. Portanto, as amostras cujo Y1 é previsto acima de 0,43 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,57 são consideradas como autênticas.

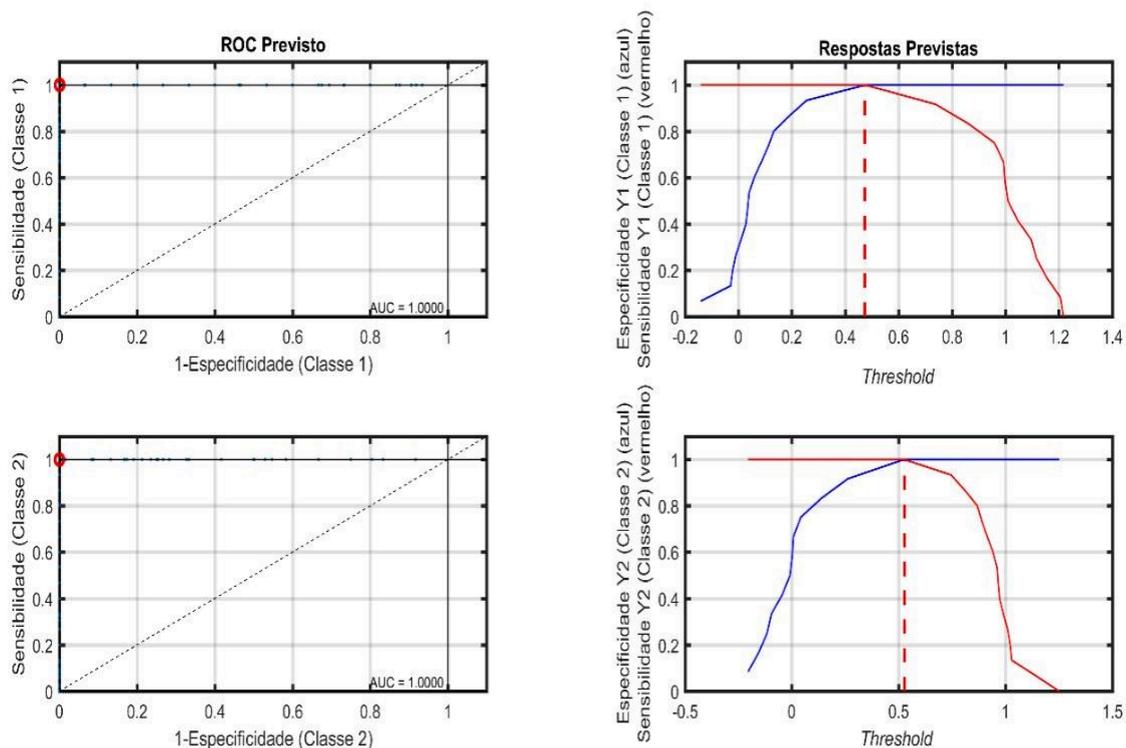
**Figura 38.** ROC das amostras biodiesel/diesel de Crambe para as classes adulteradas (classe1) e autênticas (classe2) para o conjunto de adulteração com óleo Soja



Fonte: O autor, 2018.

As curvas ROC e a relação de especificidade, sensibilidade e *threshold* para o modelo de biodiesel metílico de Crambe adulterado com óleo residual é mostrado no gráfico na Figura 39, o limite para as classes 1 e 2 foram de 0,48 e 0,51, respectivamente. Portanto, as amostras cujo Y1 é previsto acima de 0,48 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,51 são consideradas como autênticas.

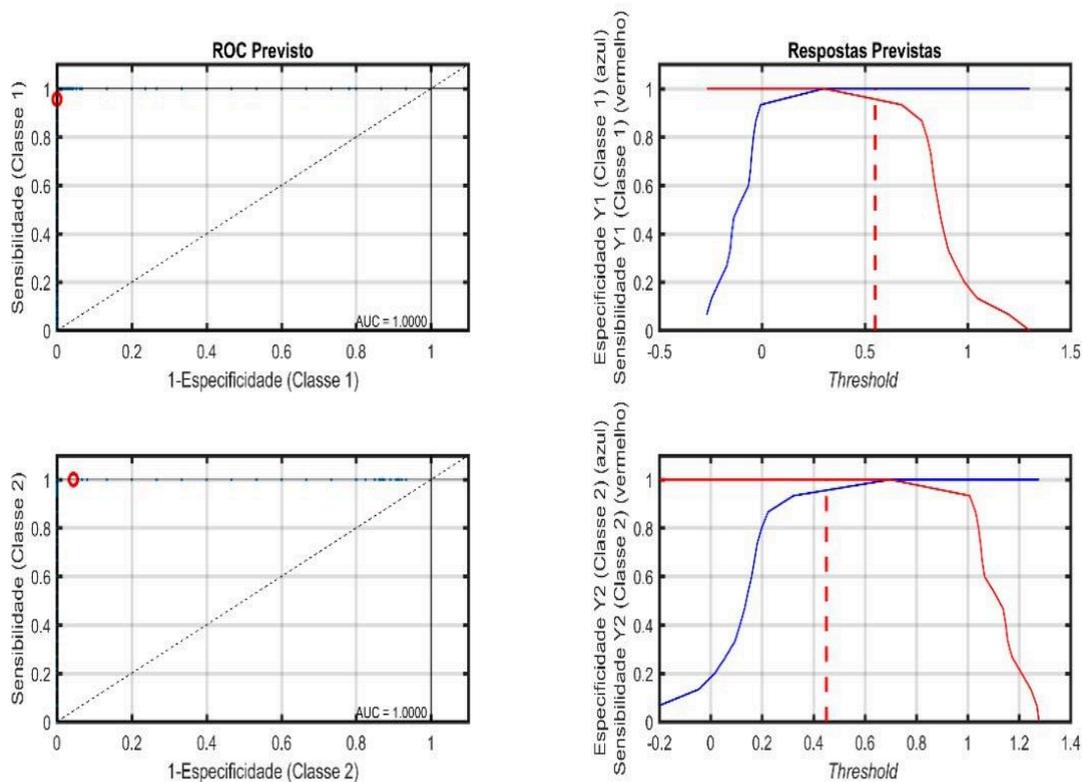
**Figura 39.** ROC das amostras biodiesel de Crambe/diesel para as classes adulteradas (classe1) e autênticas (classe2) para o conjunto de adulteração com óleo residual



**Fonte:** O autor, 2018.

As curvas ROC e a relação de especificidade, sensibilidade e *threshold* para o modelo de biodiesel metílico de Crambe adulterado com óleo automotivo usado é mostrado o gráfico de *threshold*, sensibilidade e especificidade (Figura 40), o limite para as classes 1 e 2 foram de 0,52 e 0,45, respectivamente. Portanto, as amostras cujo Y1 é previsto acima de 0,52 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,45 são consideradas como autênticas.

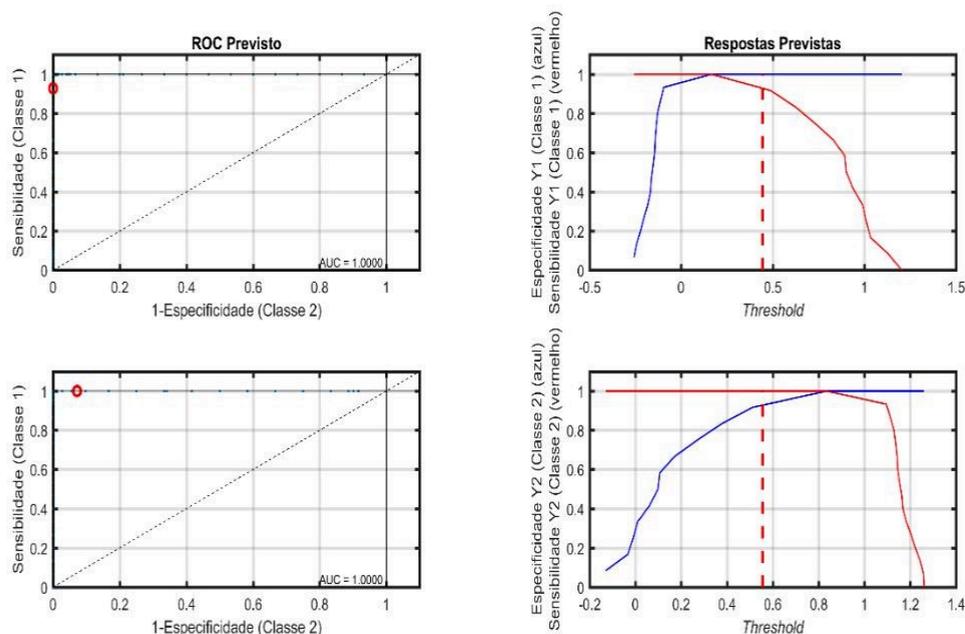
**Figura 40.** ROC das amostras biodiesel/diesel de Crambe para as classes adulteradas (class1) e autênticas (classe2) para o conjunto de adulteração com óleo automotivo usado.



**Fonte:** O autor, 2018.

Devido ao erro na previsão, o ponto de sensibilidade e especificidade do modelo sofreu um pequeno deslocamento. A partir do gráfico de *threshold*, sensibilidade e especificidade (Figura 41), o limite para as classes 1 e 2 foram de 0,48 e 0,58, respectivamente. Portanto, as amostras cujo Y1 é previsto acima de 0,48 são consideradas como adulteradas, e as amostras cujo Y2 é previsto acima de 0,58 são consideradas como autênticas.

**Figura 41.** ROC das amostras biodiesel de Crambe/diesel para as classes adulteradas (classe 1) e autênticas (classe 2) para o conjunto de adulteração com gasolina

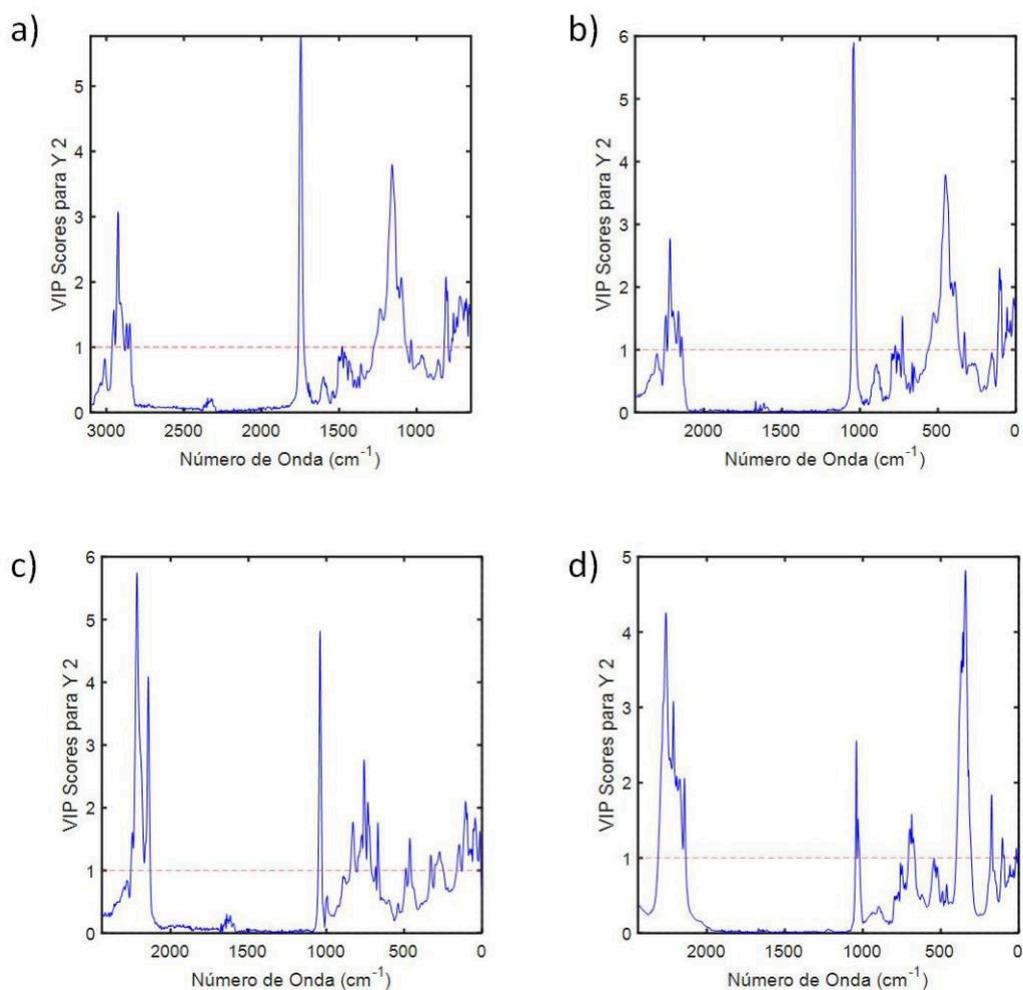


**Fonte:** O autor, 2018.

A análise dos escores VIP para os modelos de biodiesel metílico de Crambe adulterados com: óleo de Soja (Figura 42a), óleo residual (Figura 42b) são realizadas da mesma maneira que nos modelos de biodiesel metílico de Pinhão Manso uma vez que as regiões em destaque estão relacionadas à região da carbonila na faixa de  $1750\text{cm}^{-1}$  e  $1700\text{cm}^{-1}$  e também à região em torno de  $1300$  a  $1000\text{cm}^{-1}$  referentes às ligações C-O devido ao aumento da quantidade de óleo por causa da adulteração e conseqüente diminuição de diesel.

Nas Figuras 42c e 42d dá-se ênfase à região entre  $3000$  a  $2840\text{cm}^{-1}$  referentes aos derivados do petróleo.

**Figura 42.** VIP para PLS-DA das amostras biodiesel/diesel de Crambe para os conjuntos de adulteração com (a) óleo Soja; (b) óleo residual; (c) óleo automotivo usado e (d) gasolina.



**Fonte:** O autor, 2018.

Portanto, os modelos PLS-DA construídos foram eficientes tanto para o monitoramento de qualidade do teor de biodiesel (produzidos via rota metílica e usando óleos de Pinhão Manso ou Crambe) nas misturas com diesel, quanto a presença de adulterantes como o óleo de Soja, residual, óleo automotivo usado e gasolina. Os valores de sensibilidade e especificidade foram altos (acima de 0,93), com baixos erros de classificação. Além disso, foi possível identificar as regiões que apresentaram as maiores contribuições para a discriminação das amostras conformes ou não conformes, bem como associar com as variações químicas destas.

## 5. CONCLUSÕES

O uso combinado da espectroscopia MIR e método quimiométrico de classificação supervisionada PLS-DA permitiu o desenvolvimento de modelos capazes de classificar amostras segundo padrões exigidos pela ANP.

Os modelos foram validados usando parâmetros de qualidades como falso positivo, falso negativo, verdadeiro positivo, verdadeiro negativo, sensibilidade e especificidade, os quais foram concordantes com o exigido.

Para o caso dos modelos PLS-DA tanto para teor como para as adulterações os valores encontrados para os parâmetros de sensibilidade e especificidade estavam dentro do permitido.

Ferramentas gráficas como as curvas ROC obtiveram resultados de especificidade igual a 0 ou próximo de 0 e sensibilidade igual a 1 ou próximo de 1 para todos os modelos.

A partir dos gráficos de pesos e escores VIP, foi possível identificar as regiões dos espectros MIR responsáveis pela distinção das amostras, o que permite, em futuros trabalhos, o uso de métodos de seleção de variáveis visando melhorar o desempenho dos modelos.

Portanto, os métodos desenvolvidos neste trabalho se apresentam como uma alternativa viável para o controle de qualidade destes combustíveis, podendo vir a ser empregados nos órgãos de fiscalização e controle de qualidade e também em indústrias que visam qualidade deste produto.

Assim, podem identificar fraudes usando equipamentos infravermelhos portáteis fazendo análises rápidas, utilizando pequena quantidade de amostras e sendo estas análises feitas no local de produção industrial ou de comercialização do combustível.

## REFERÊNCIAS

- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS – ANP (Brasil). **Resolução ANP no. 454, de 11 maio 214**. Disponível em: [http://nxt.anp.gov.br/nxt/gateway.dll/leg/resolucoes\\_anp/2012/maio/ranp%2014%20-%202012.xml](http://nxt.anp.gov.br/nxt/gateway.dll/leg/resolucoes_anp/2012/maio/ranp%2014%20-%202012.xml) Acesso em: 07/01/2019.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS – ANP (Brasil). **Óleo Diesel**, 2017. Disponível em: <<http://www.anp.gov.br/petroleo-derivados/155-combustiveis/1857-oleo-diesel>>. Acesso em: 07/01/2019.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS – ANP (Brasil). **Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis**, Rio de Janeiro, 2018a. Disponível em: [http://www.anp.gov.br/images/publicacoes/anuario-estatistico/2018/anuario\\_2018.pdf](http://www.anp.gov.br/images/publicacoes/anuario-estatistico/2018/anuario_2018.pdf). Acesso em: 07/01/2019
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS – ANP (Brasil). **Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis**, Rio de Janeiro, 2019. Disponível em <http://www.anp.gov.br/arquivos/central-conteudos/anuario-estatistico/2019/anuario2019> Acesso em: 06/01/2019.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS – ANP (Brasil). **Informações de Mercado**. 2018b. Disponível em: <http://www.anp.gov.br/producao-de-biocombustiveis/biodiesel/informacoes-de-mercado>. Acesso em: 07/01/2019.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS – ANP (Brasil). **Seminário de Avaliação do Mercado de Combustíveis 2018**, 2018c. Disponível em: [http://www.anp.gov.br/images/Palestras/seminario\\_avaliao\\_2018.pdf](http://www.anp.gov.br/images/Palestras/seminario_avaliao_2018.pdf) . Acesso em: 07/01/2019.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS, ANP. Disponível em: <http://www.gov.br/producao-de-biocombustiveis/biodiesel/informacoes-de-mercado>. Acesso em: 02/11/2019.
- AHMAD, A. L. et al. Microalgae as a sustainable energy source for biodiesel production: A review. **Renewable and Sustainable Energy Reviews**, v. 15, p. 584, 2011. <https://doi.org/10.1016/j.rser.2010.09.018>
- 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. 305-311, 2013. <https://doi.org/10.1016/j.talanta.2013.09.025>
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS- ABNT. NBR-15568: **Biodiesel- Determinação do teor de biodiesel em óleo diesel por espectroscopia na região do infravermelho médio**. Rio de Janeiro, 2008. Disponível em: <https://www.target.com.br/produtos/normas-tecnicas/40641/nbr15568-biodiesel-determinacao-do-teor-de-biodiesel-em-oleo-diesel-por-espectroscopia-na-regiao-do-infravermelho-medio> Acesso em: 03/11/2019

ATABANI, A. E. et al. A comparative evaluation of physical and chemical properties of biodiesel synthesized from edible and non-edible oils and study on the effect biodiesel blending. **Energy**, v. 58, p. 296-304, 2013. <https://doi.org/10.1016/j.energy.2013.05.040>

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, p. 190-197, 2011. <https://doi.org/10.1016/j.aca.2011.01.041>

BALAT, M.; BALAT, H. Recent trends in global production and utilization of bio-ethanol fuel. **Applied Energy**, Londres, v. 86, p. 2273-2282, 2009. <https://doi.org/10.1016/j.apenergy.2009.03.015>

BALLABIO, D.; CONSONNI, V. Classification tools in chemistry. Part 1: linear models. PLS-DA. **Analytical Methods**, v. 5, p. 3790–3798, 2013. <https://doi: 10.1039/c3ay40582f>

BARBOSA, L. C. A. **Espectroscopia no infravermelho na caracterização de compostos orgânicos**. ISBN-13: 978-8572692809. Viçosa: UFV, 2007.

BARKER, M.; RAYENS, W. Partial least squares for discrimination. **Journal of Chemometrics**, Chichester, v. 17, p. 166-173, 2003. <https://doi.org/10.1002/cem.785>

BRANDÃO, L. F. P.; BRAGA, J. W. B.; SUAREZ, P. A. Z. Determination of vegetable oils fats adulterants in diesel oil by high performance liquid chromatography and multivariate methods. **Journal of cromatography A**, v.1225, p. 150-157, 2012. <https://doi: 10.1016/j.chroma.2011.12.076>.

BUIATTE, J. E. et al. Qualitative and Quantitative Monitoring of Methyl Cotton Biodiesel Content in Biodiesel/Diesel Blends Using MIR Spectroscopy and Chemometrics Tools. **Journal Brazilian Chemometrics Society**, v. 27, p. 84-90, 2016. <http://dx.doi.org/10.5935/0103-5053.20150251>

CHATARINO, R. R. et al. Biodiesel typification and quality control by direct infusion electrospray ionization mass spectrometry fingerprinting. **Energy & Fuels**, v.21, p. 3698-3701, 2007. <https://doi.org/10.1021/ef7003078>

CÉSAR, A. S. et al. The prospects of using *Acrocomia aculeata* (macaúba) a non-edible biodiesel feedstock in Brazil. **Renewable and Sustainable Energy Reviews**, v. 49, p. 1213-1220, 2015. <https://doi.org/10.1016/j.rser.2015.04.125>

Conselho Nacional de Política Energética – CNPE- publicado em 09-11-2017. Disponível em: <http://agenciabrasil.ebc.com.br/economia/noticia/2017-11/conselho-antecipa-para-marco-aumento-do-percentual-de-biodiesel-no-diesel> Acesso em : 06/09/2018

COATES, J. Interpretation of Infrared Spectra, A Pratical Approach, Coates Consulting, Newtown, USA, 2006.

COMPANHIA DE SANEAMENTO BÁSICO DO ESTADO DE SÃO PAULO-SABESP. Programa de Reciclagem de óleo de Fritura da Sabesp -São Paulo. Disponível em: [http://site.sabesp.com.br/uploads/file/asabesp\\_doctos/programa\\_reciclagem\\_oleo\\_completo.pdf](http://site.sabesp.com.br/uploads/file/asabesp_doctos/programa_reciclagem_oleo_completo.pdf) . Acesso em: 02/11/2019.

COSTA NETO, P. R.. et al. Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de Soja usado em frituras. **Química Nova**, v.23, p. 531-537, 2000. <http://dx.doi.org/10.1590/S0100-40422000000400017>

Ministério de Minas e Energia. Boletim mensal dos combustíveis renováveis, E. n 96 Fevereiro/2016.

FALASCA, S. L. et al. *Crambe abyssinica*: An almost unknown crop with a promissory future to produce biodiesel in Argentina. **International Journal of Hydrogen Energy**, Oxford, v. 35, p. 5808-5812, 2010. <https://doi.org/10.1016/j.ijhydene.2010.02.095>

FATEMI, S. H.; HAMMOND, E. G. Analysis of oleate, linoleate and linolenate hydroperoxides in oxidized ester mixtures. **Lipids**, Champaign, v. 15, p. 379-385, 1980. <https://doi.org/10.31274/rtd-180813-6390>

FENG, F.; WU, W.; ZENG, L. Rapid analysis of diesel fuel properties by near infrared reflectance spectra. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, Amsterdã, v. 149, p. 271-278, 2015. <https://doi.org/10.1016/j.saa.2015.04.095>

FERRARI, R. A.; OLIVEIRA, V. D. S.; A. SCABIO. Biodiesel de Soja- taxa de conversão em ésteres etílicos, caracterização físico-química e consumo em gerador de energia, **Química Nova**, v.28, p.19-23, 2005. <http://dx.doi.org/10.1590/S0100-40422005000100004>

FERREIRA, M. M. C. **Quimiometria. Conceitos, Métodos e Aplicações**. 1ª ed. Campinas. Unicamp, 2015.

GARMARUDI, B. A. et al. Origin based classification of crude oils by infrared spectrometry and chemometrics. **Fuel**, Londres, v. 236, p. 1093-1099, 2019. <https://doi.org/10.1016/j.fuel.2018.09.013>

GAYDOU, V.; KISTER, J.; DUPUY, 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, n. 2, p. 190-197, abr. 2011. <https://doi.org/10.1016/j.chemolab.2010.05.002>

GONTIJO, L. C.. **Uso de espectrometria no infravermelho médio, calibração multivariada e seleção de variáveis por intervalos na quantificação de biodieseis em misturas com diesel**. 2016. 106 f. Tese (Doutorado em Química) - Instituto de Química - Universidade Federal de Uberlândia, Uberlândia, 2016. Disponível em : < <https://repositorio.ufu.br/handle/123456789/17536>>  
Acesso em: 07/09/2018

GUI, M. M.; LEE, K. T.; BHATIA, S. Feasibility of edible oil vs. non-edible oil vs. waste edible oil as biodiesel feedstock. **Energy**, v.33, p.1646, 2008. <https://doi.org/10.1016/j.energy.2008.06.002>

GUIMARÃES, E. **Aplicação de Espectroscopia no Infravermelho Médio e Calibração Multivariada para a quantificação de adulterantes em biodieséis de óleo de Soja e residual e suas misturas com diesel.** 2014. 84 f. Dissertação (Mestrado em Química) – Instituto de Química – Universidade Federal de Uberlândia, Uberlândia, 2014. Disponível em: <https://repositorio.ufu.br/handle/123456789/17400>

GUIMARÃES, E. et al. Aplicação das Figuras de Mérito Multivariadas na Validação de Metodologias em Análises de Biocombustíveis empregando Espectroscopia no Infravermelho Médio e PLS, **Revista Virtual de Química**, v.7, p. 2242-2254, 2015. <https://10.5935/1984-6835.20150133>

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

KARPAGAM, R. et al. Characterization and fatty acid profiling in two fresh water microalgae for biodiesel production: Lipid enhancement methods and media optimization using response surface methodology. **Bioresource Technology**. v.188, p.177, 2015. <https://doi.org/10.1016/j.biortech.2015.01.053>

KNOTHE, G. et al. Determining the blend level of mixtures of biodiesel with conventional diesel fuel by fiber-optic near-infrared spectroscopy and <sup>1</sup>H nuclear magnetic resonance spectroscopy. **Journal of American Oil Chemical Society**, Chicago, v. 78, p. 1025-1028, 2001. <https://doi/10.1007/s11746-001-0382-0>

KNOTHE, G. et al. **Manual de biodiesel. Matérias primas alternativas e tecnologias para a produção de biodiesel.** 1 ed. São Paulo: Editora Egdgard Blücher LTDA, p. 46-61, 2006. <https://doi/10.21577/1984-6835.20170020>

KNOTHE, G.; KRAHL, J.; GERPEN, J.V. **The Biodiesel Handbook**, 2a. ed., AOCS: Urbana, 2009.

KUMAR, K.; MISHRA, A. K. Quantification of ethanol in ethanol-petrol and biodiesel in biodiesel-diesel blends using fluorescence spectroscopy and multivariate methods, **Journal of Fluorescence**, v.22, p. 339-347, 2012. <https://doi: 10.1007/s10895-011-0964-6>.

LAVIOLA, B. G. et al. Biometric and biotechnology strategies in Jatropha genetic breeding for biodiesel production. **Renewable and Sustainable Energy Reviews**. v.76, p. 894-904, 2017. <https://doi.org/10.1016/j.rser.2017.03.116>

LEUNG, D. Y.; KOO, B. C.; GUO, Y.; Degradation of biodiesel under different storage conditions. **Bioresource Technology**, Essex, v. 97, p. 250-256, jan. 2006. <https://doi.org/10.1016/j.biortech.2005.02.006>

LI, X. et al. Characterizing the moisture content of tea with diffuse reflectance spectroscopy using wavelet transform and multivariate analysis. **Sensors, Basel**, v. 12, p. 9847-9861, 2012. <https://doi.10.3390/s120709847>

LIM, B. Y. et al. review of processing and machinery for Jatropha curcas L. fruits and seeds in biodiesel production: Harvesting, shelling, pretreatment and storage. **Renewable and Sustainable Energy Reviews**.v.52, p. 991-1002, 2015. <https://doi.org/10.1016/j.rser.2015.07.077>

LÔBO, I. P.; FERREIRA, S. L. C.; CRUZ, R. S. Biodiesel: parâmetros de qualidade e métodos analíticos. **Química Nova**, São Paulo, v. 32, n. 6, p. 1596-1608, 2009. <http://dx.doi.org/10.1590/S0100-40422009000600044>

MÁQUINA, A. D. V. **Espectrometria no Infravermelho Médio e Métodos Quimiométricos PLS-DA e PLS: classificação e previsão do teor de biodiesel na mistura de biodiesel/diesel de mafurra, moringa e algodão**. 2017. 80 f. Dissertação (Mestrado em Química) - Instituto de Química - Universidade Federal de Uberlândia, Uberlândia, 2017. <https://repositorio.ufu.br/handle/123456789/1809>

MÁQUINA, A. D. V. ; SITEO, B. V. ; BUIATTE, J. E. ; SANTOS, D. Q ; NETO, W.B. . Quantification and classification of cotton biodiesel content in diesel blends, using mid-infrared spectroscopy and chemometric methods. **Fuel**, v. 237, p. 373-379, 2019. <https://doi.org/10.1016/j.fuel.2018.10.011>

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 Brazilian Chemical Society**, v.26, p. 642-648, 2015a. <http://dx.doi.org/10.5935/0103-5053.20150020>

MAZIVILA, S. J.; MITSUTAKE, H.; SANTANA, F. B.; 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**, v 142, p. 222-226, 2015b. <https://doi.org/10.1016/j.fuel.2014.11.014>

MASSART, D. L. et al. **Data Handling in Science and Technology: Part A**, v. 20. Amsterdã: Ed. Elsevier, 1998.

MEHMOOD, T.; LILAND, K. H.; SNIPEN, L. SÆBØ, S. A review of variable selection methods in Partial least squares regression. **Chemometrics and Intelligent Laboratory Systems**, Amsterdã, v. 118, p. 62-69, 2012. <https://doi.org/10.1016/j.chemolab.2012.07.010>

MENTEN, J. F. M.; MIYADA, V. S.; BERENCHTEIN, B. **Glicerol na alimentação animal. In: Simpósio sobre Manejo e Nutrição de Aves e Suínos**, ISSN 1806-92902008, Campinas, SP. Colégio Brasileiro de Nutrição Animal, p. 101-114, 2008.

Ministério de Minas e Energia (MME) Disponível em: <http://www.mme.gov.br/web/guest/secretarias/petroleo-gas-natural-e-combustiveis-renovaveis/publicacoes/boletim-mensal-de-biocombustiveis>. Acesso em: 06/11/2019

MITTELBAACH, M.; REMSCHMIDT, C. **Biodiesel the Comprehensive Handbook**, 2a. ed., Wiley: Viena, 2005.

MOFIJUR, M. et al. Effect of biodiesel from various feedstocks on combustion characteristics, engine durability and materials compatibility: A review. **Renewable and Sustainable Energy Reviews**, v. 28, p. 441-455, 2013. <https://doi.org/10.1016/j.rser.2013.07.051>

MOTA, J. A. C.; SILVA, X. A. C.; GONÇALVES, L. C. V. Gliceroquímica: novos produtos e processos a partir da glicerina de produção de biodiesel. **Química Nova**, v. 32, n. 3, p. 639-648, 2009. <http://dx.doi.org/10.1590/S0100-40422009000300008>

- MONTEIRO, M. R.. Evaluation of biodiesel-diesel blends quality using  $^1\text{H}$ NMR and chemometrics. **Talanta**, v. 78, p. 660–664, 2009. <http://dx.doi.org/10.1016/j.talanta.2008.12.026>.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT – OECD. OECD-FAO Agricultural Outlook 2018-2027, 2018. Disponível em: <https://www.oecd-ilibrary.org/oecd-fao-agricultural-outlook-> Acesso em: 08/10/2018.
- ONOREVOLI, B. et al. A one-dimensional and comprehensive two-dimensional gas chromatography study of the oil and the bio-oil of the residual cakes from the seeds of *Crambe abyssinica*. **Industrial Crops and Products**, v. 52, p. 8-16, 2014. <https://doi.org/10.1016/j.indcrop.2013.09.034>
- PADULA, A. D. et al. The emergence of the biodiesel industry in Brazil: Current figures and future prospects. **Energy Policy**, Surrey, v. 35, p. 395-405, 2012. <https://doi.org/10.1016/j.enpol.2012.02.003>
- PAGLIARO, M. et al. From Glycerol to Value-Added Products. *Angewandte Chemie International Edition*, v. 46, n. 24, p. 4434-4440, 2007. <https://doi.org/10.1002/anie.200604694>
- PEITER, G. C. et al. Alternativas para o uso do glicerol produzido a partir de biodiesel. **Revista Brasileira de Energia Renováveis**. v.5, n .4, 2016. <http://dx.doi.org/10.5380/rber.v5i4.46501>
- PINHO, D. M. M. et al. Evaluating use of EM 14078 for determination of biodiesel in diesel blends sold in the Brazilian Market. **Fuel**, Londres, v.136, p.136-142, 2014. <https://doi.org/10.1016/j.fuel.2014.07.043>
- POUSA, G. P. A. G.; SANTOS, A. L. F.; SUAREZ, P. A. Z. History and policy of biodiesel in Brazil. **Energy Policy**, Surrey, v. 35, n. 11, p. 5393-5398, 2007. <https://doi.org/10.1016/j.enpol.2007.05.010>
- QUINTELLA, C. M. et al. Cadeia do biodiesel da bancada à indústria: uma visão geral com prospecção de tarefas e oportunidades para P&D&I, **Química Nova**, São Paulo, v. 32, p. 793-808, 2009. <https://doi/10.1590/S0100-40422009000300022>
- RAMOS, L. P. et al. Biodiesel: Matérias Primas, Tecnologias de Produção e Propriedades Combustíveis. **Revista Virtual de Química**, v. 9, p.317-369, 2017. <https://doi/10.21577/1984-6835.20170020>
- SACRÉ, P. -Y. et al. Data processing of vibrational chemical imaging for pharmaceutical applications. **Journal of Pharmaceutical and Biomedical Analysis**, Amsterdam, v. 101, p. 123– 140, 2014. <https://doi.org/10.1016/j.jpba.2014.04.012>
- SANTOS, A. F. **Novas perspectivas da glicerina-síntese de novos nitratos com propriedades farmacológicas e melhoradores de cetano**. 2009. 116 f. Dissertação (Mestrado em Química) Instituto de Química - Universidade Federal da Paraíba, 2009. Disponível em: <https://repositorio.ufpb.br/jspui/handle/tede/7169> Acesso em: 06/07/2018.
- SANTOS, V. H.J.M. et al. Discriminant Analysis and Cluster Analysis of Biodiesel Fuel Blends Based on Fourier Transform Infrared Spectroscopy (FTIR), **Energy & Fuels**, Washington, v.30, p. 4905-4915, 2016. <https://doi.org/10.1021/acs.energyfuels.6b00447>

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v.36, p.1627, 1964.

<https://doi.org/10.1021/ac60214a047> Disponível em: 04/10/2018.

SKOOG, D.A.; HOLLER, f. J.; NIEMAN, T. A. **Princípios de análise instrumental**. 5. ed. São Paulo: Bookman, 2002.

SINGHABHANDHU, A.; TEZUKA, T. A perspective on incorporation of glycerin purification process in biodiesel plants using waste cooking oil as feedstock. Energy Economics Laboratory. Department of Socio-Environmental Energy Science. Kyoto University. Japan. **Energy**, p. 35, p. 2493-2504, 2010.

<https://doi.org/10.1016/j.energy.2010.02.047>

SMITH, B.C. **Fundamentals of Fourier transform infrared spectroscopy**. 2 ed. Boca Raton: CRC Press, 2011.

SOARES, I. P.; REZENDE, T. F.; FORTES, I. C. P. Study of the behavior changes in physical-chemistry properties of diesel/biodiesel (B2) mixtures with residual oil and its quantification by Partial Least-Squares Attenuated Total Reflection-Fourier Transformed Infrared Spectroscopy (PLS/ATR-FTIR). **Energy & Fuels**, Washington, v. 23, n. 8, p. 4143-4148, 2009. <https://doi.org/10.1021/ef900302q>

SORRELL, S. et al. Shaping the global oil peak: A review of the evidence on field sizes, reserve growth, decline rates and depletion rates. **Energy**, Oxford, v. 37, n. 1, p. 709-724, 2012. <https://doi.org/10.1016/j.energy.2011.10.010>

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. <https://doi.org/10.1016/j.fuel.2014.03.051>

TOLMASQUIM, M. T.; GUERREIRO, A.; GORINI, R. Matriz Energética Brasileira, : uma prospectiva. **Novos Estudos -CEBRAP**, nº 79, São Paulo, nov. 2007.

VANDEGINSTE, B. G. M. et al. **Handbook of Chemometrics and Qualimetrics: Part B**, Amsterdã: Ed. Elsevier, 1998.

VASCONCELOS, F. V. et al. Using near-infrared overtone regions to determine biodiesel content and adulteration of diesel/biodiesel blends with vegetable oils. **Analytica Chimica Acta**, Amsterdã, v. 716, p. 101-107, 2012. <https://doi.org/10.1016/j.aca.2011.12.027>.

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, p. 506-510, 2012. <https://doi.org/10.1016/j.meatsci.2012.05.019>

WAZILEWSKI, W. T. et al. Study of the methyl Crambe (*Crambe abyssinica* Hochst) and soybean biodiesel oxidative stability. **Industrial Crops and Products**, v. 43, p. 207-212, 2013. <https://doi.org/10.1016/j.indcrop.2012.07.046>