

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**TESE DE DOUTORADO**

**METODOLOGIAS ANALÍTICAS PARA AVALIAR A  
BIODEGRADABILIDADE DO DIESEL, BIODIESEL E BLENDS B10**

**Doutoranda: Camila Correa**

**Orientador: Prof. Dr. Marco Flôres Ferrão**

**Co-orientadora: Profa. Dra. Fátima Menezes Bento**

**Porto Alegre, maio de 2021.**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**  
**INSTITUTO DE QUÍMICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

CAMILA CORREA

**METODOLOGIAS ANALÍTICAS PARA AVALIAR A**  
**BIODEGRADABILIDADE DO DIESEL, BIODIESEL E BLENDA B10**

Tese apresentada como requisito  
parcial para a obtenção do grau  
de Doutora em Química

Prof. Dr. Marco Flôres Ferrão

Orientador

Profa. Dra. Fátima Menezes Bento

Co-orientadora

Porto Alegre, maio de 2021.

A presente tese foi realizada inteiramente pelo autor, exceto as colaborações as quais serão devidamente citadas nos agradecimentos, no período entre (mês/ano de ingresso) e (mês/ano da defesa), no Instituto de Química da Universidade Federal do Rio Grande do Sul sob Orientação do Professor Doutor Marco Flôres Ferrão e Co-orientação da Professora Doutora Fátima Menezes Bento. A tese foi julgada adequada para a obtenção do título de Doutora em Química pela seguinte banca examinadora.

**Comissão Examinadora:**

Dr(a). Adriane Ramos Zimmer

Prof. Dr. Luiz Antônio Mazzini Fontoura

Prof. Dr. Adriano de Araújo Gomes

Profa. Dra. Marcia Messias da Silva

Prof. Dr. Marco Flôres Ferrão

Profa. Dra. Fátima Menezes Bento

Camila Correa

## **Agradecimentos**

Ao professor Marco Flôres Ferrão pela orientação, ensinamentos, amizade, confiança e paciência durante a realização deste trabalho.

À professora Fátima Menezes Bento pela co-orientação.

Ao professor Pedro Henrique A. Konzen pela sua significativa contribuição no desenvolvimento do método de contagem de unidades formadoras de colônias.

Aos meus amigos e colegas de laboratório Fábio, Fernanda e Juliana

Aos meus amigos e colegas de trabalho Cristina, Adriana, Vagner e Litiele.

As minhas amigas Jucelaine e Camila Cerveira.

Ao meu pai Delci e meus irmãos Mateus e Lucas.

As minhas tias Carolina e Joessane.

À memória de minha mãe, pelo amor, dedicação e valorizar a importância da educação na vida dos seus filhos.

A Universidade Federal do Rio Grande do Sul pelos recursos estruturais e financeiros para realização deste trabalho.

# SUMÁRIO

Lista de Figuras .....	vi
Lista de Tabelas.....	viii
Lista de Abreviaturas.....	ix
Resumo.....	x
Abstract.....	xi
<b>1 Introdução.....</b>	<b>1</b>
<b>2 Objetivo.....</b>	<b>3</b>
2.1 Geral.....	3
2.2 Específicos .....	3
<b>3 Revisão Bibliográfica.....</b>	<b>4</b>
3.1 Diesel.....	4
3.2 Biodiesel.....	5
3.3 Biodegradação dos Combustíveis.....	6
3.3.1 <i>Fatores que Contribuem para a Biodegradação</i> .....	7
3.3.1.1 Água.....	7
3.3.1.2 Tempo de Armazenamento.....	7
3.3.2 <i>Biodegradação do Diesel e Biodiesel</i> .....	8
3.4 <i>Análise da Biodegradação em Combustíveis</i> .....	9
3.5 <i>Espectroscopia no Infravermelho com Transformada Fourier (FTIR)</i> .....	11
3.6 <i>Quimiometria</i> .....	12
3.6.1 <i>Pré-processamento dos dados</i> .....	13
3.6.1.1 Normalização.....	15
3.6.1.2 Centrar na Média .....	15
3.6.1.3 Primeira e Segunda derivada .....	16
3.6.2 <i>Análise por Componentes Principais (PCA)</i> .....	17
3.6.3 <i>Calibração Multivariada</i> .....	19
3.6.3.1 Regressão por Mínimos Quadrados Parciais (PLS).....	19
3.7 <i>Análise de Imagens</i> .....	21
<b>4 Resultados .....</b>	<b>22</b>
<b>4.1 Pré-Processamento Espectral e Análise por Componentes Principais (PCA) de Diesel com Diferentes Teores de Enxofre Submetidos a Biodegradação .....</b>	<b>23</b>
4.1.1 <i>Objetivo</i> .....	23

4.1.2	<i>Materiais e Métodos</i> .....	23
4.1.2.1	Descrição das amostras de diesel.....	23
4.1.2.2	Descrição do experimento de estocagem simulada .....	23
4.1.2.3	Espectroscopia no Infravermelho .....	24
4.1.2.4	Análise Multivariada dos dados.....	24
<b>4.2</b>	<b>Resultados e Discussão</b> .....	<b>26</b>
4.2.1	<i>Caracterização do Diesel por FTIR</i> .....	26
4.2.2	<i>Otimização do Pré-processamento e Análise Exploratória dos dados de FTIR – ATR de Diesel com diferentes teores de enxofre submetidos a biodegradação</i> .....	27
4.2.3	<i>Análise Multivariada dos Espectros de Diesel S10</i> .....	32
4.2.4	<i>Análise Multivariada dos Espectros de Diesel S50, S500 e S1800</i> .....	39
<b>4.3</b>	<b>Conclusão</b> .....	<b>41</b>
<b>4.4</b>	<b>Quantificação da Biomassa Produzida na Biodegradação de Blendas B10 durante estocagem simulada usando Infravermelho com Transformada de Fourier (FTIR) e Regressão por Mínimos Quadrados Parciais (PLS)</b> .....	<b>42</b>
4.4.1	<i>Objetivo</i> .....	42
4.4.2	<i>Materiais e Métodos</i> .....	42
4.4.2.1	Amostras .....	42
4.4.2.2	Quantificação da Biomassa.....	42
4.4.2.3	Espectroscopia no Infravermelho .....	43
4.4.2.4	Análise Multivariada dos Dados.....	43
<b>4.5</b>	<b>Resultados e Discussão</b> .....	<b>45</b>
4.5.1	<i>Caracterização da Blenda B10 por FTIR</i> .....	45
4.5.2	<i>Quantificação da Biomassa por Calibração Multivariada</i> .....	46
<b>4.6</b>	<b>Conclusão</b> .....	<b>50</b>
<b>4.7</b>	<b>Utilização de Imagens Digitais para a Contagem de Colônias de Microrganismos Deteriogênicos de Biodiesel</b> .....	<b>51</b>
4.7.1	<i>Objetivo</i> .....	51
4.7.2	<i>Materiais e Métodos</i> .....	51
4.7.2.1	Cultura dos Microrganismos.....	51
4.7.2.2	Aquisição das Imagens .....	52
4.7.3	<i>Contador Semi-Automático de Colônias</i> .....	53
<b>4.8</b>	<b>Resultados e Discussão</b> .....	<b>55</b>
4.8.1	<i>Bacillus Pumilus</i> .....	56

4.8.2 <i>Meyerozyma Guilliermondii</i> .....	56
<b>4.9 Conclusão .....</b>	<b>58</b>
<b>5 Considerações Finais .....</b>	<b>59</b>
<b>6 Referências.....</b>	<b>60</b>
ANEXOS.....	75

## Lista de Figuras

<b>Figura 1:</b> Reação de transesterificação de um triglicerídeo para produção de biodiesel. ....	6
<b>Figura 2:</b> Fluxograma das etapas envolvidas na aplicação da análise multivariada de dados..	13
<b>Figura 3:</b> Dados centrados na média: a) espectros brutos e b) espectros centrados na média	16
<b>Figura 4:</b> Curva com uma inclinação e um deslocamento de linha de base (a) , (b) aplicação da primeira derivada e (c) aplicação da segunda derivada. ....	17
<b>Figura 5:</b> Estrutura matemática do modelo PCA. ....	18
<b>Figura 6:</b> Espectros de FTIR- ATR das amostras de diesel com diferentes teores de enxofre. ....	26
<b>Figura 7:</b> Espectros de FTIR- ATR das amostras de diesel com diferentes teores de enxofre, azul (10 ppm), verde (50 ppm), preto (500 ppm) e vermelho (1800 ppm). ....	27
<b>Figura 8:</b> Gráfico de escores da PCA dos espectros de FTIR- ATR centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800(vermelho). ....	28
<b>Figura 9:</b> Gráfico de escores da PCA dos espectros de FTIR- ATR normalizados e centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800(vermelho). ....	29
<b>Figura 10:</b> Gráfico de escores PC1xPC2 dos espectros de FTIR- ATR normalizados, com primeira derivada e centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800(vermelho). ....	30
<b>Figura 11:</b> Gráfico de escores da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800(vermelho). ....	30
<b>Figura 12:</b> Gráfico de pesos da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel com diferentes teores de enxofre: a) PC1 e b) PC2. ....	31
<b>Figura 13:</b> Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-ATR normalizados e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho). ....	33
<b>Figura 14:</b> Gráfico de escores PC1xPC3 da PCA dos espectros de FTIR-ATR normalizados e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho). ....	33
<b>Figura 15:</b> Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR- ATR normalizados, com primeira derivada e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho). ....	34
<b>Figura 16:</b> Gráfico de escores PC1xPC3 da PCA dos espectros de FTIR- ATR normalizados, com primeira derivada e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho). ....	35
<b>Figura 17:</b> Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho). ....	36

<b>Figura 18:</b> Gráfico de escores PC1xPC3 da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).....	36
<b>Figura 19:</b> Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318 cm <sup>-1</sup> para o diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).....	38
<b>Figura 20:</b> Gráfico de pesos da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318 cm <sup>-1</sup> para o diesel S10. ....	38
<b>Figura 21:</b> Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318 cm <sup>-1</sup> para o diesel S50: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).....	39
<b>Figura 22:</b> Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318 cm <sup>-1</sup> para o diesel S500: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).....	40
<b>Figura 23:</b> Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318 cm <sup>-1</sup> para o diesel S1800: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).....	40
<b>Figura 24:</b> Espectros de FTIR-ATR das amostras de B10.....	45
<b>Figura 25:</b> Espectros de B10 na região de 1800 a 1650 cm <sup>-1</sup> .....	46
<b>Figura 26:</b> Espectro com destaque das regiões 22 e 24 selecionadas no modelo de regressão s2iPLS32. ....	48
<b>Figura 27:</b> Gráfico dos valores de RMSECV por número de variáveis latentes utilizadas pelo modelo s2iPLS32. ....	49
<b>Figura 28:</b> Gráfico dos valores de referência versus valores previstos por siPLS para determinação de biomassa.....	49
<b>Figura 29:</b> Imagens da placa de petri com colônias de bactérias (esquerda) e leveduras (direita). ....	52
<b>Figura 30:</b> Amostras de imagens redimensionadas e achatadas das culturas selecionadas de <i>Bacillus pumilus</i> (esquerda) e <i>Meyerozyma guilliermondii</i> (direita). ....	54
<b>Figura 31:</b> Amostras de imagens filtradas das culturas selecionadas de <i>Bacillus pumilus</i> (esquerda) e <i>Meyerozyma guilliermondii</i> (direita). ....	54
<b>Figura 32:</b> Unidades formadoras de colônias identificadas em imagens de culturas selecionadas de <i>Bacillus pumilus</i> (esquerda) e <i>Meyerozyma guilliermondii</i> (direita). Os círculos vermelhos estão centrados nos pontos máximos locais. ....	55
<b>Figura 33:</b> Gráficos de calibração (esquerda) e validação (direita) para o caso de teste <i>Bacillus pumilus</i> . ....	56
<b>Figura 34:</b> Gráficos de calibração (esquerda) e validação (direita) para o caso de teste <i>Meyerozyma guilliermondii</i> .....	57

## **Lista de Tabelas**

<b>Tabela 1:</b> Regiões do espectro de infravermelho. ....	11
<b>Tabela 2:</b> Descrição da nomenclatura das amostras .....	24
<b>Tabela 3:</b> Descrição dos conjuntos de calibração e previsão. ....	44
<b>Tabela 4:</b> Resultados das figuras de mérito dos modelos de regressão por PLS, iPLS e siPLS para determinação da biomassa. ....	47

## **Lista de Abreviaturas**

**ANP** – Agência Nacional do Petróleo, Gás Natural e Biocombustíveis

**ATR** - Reflexão Total Atenuada (do inglês, *Attenuated Total Reflectance*)

**ASTM** - American Society of Testing and Materials

**CCM** - Carta de Controle Multivariada

**CEN** - Comité Europeu de Normalização

**HCA** – Análise Hierárquica por Agrupamento (do inglês, *hierarchical cluster analysis*)

**IR** - Infravermelho (do inglês *Infrared*)

**FIR** - Infravermelho distante (do inglês *Far Infrared*)

**FTIR** - Espectroscopia no Infravermelho Médio com Transformada de Fourier (do inglês *Fourier Transform Infrared*)

**HATR** - Reflexão Total Atenuada Horizontal (do inglês, *Horizontal Attenuated Total Reflectance*)

**i-PCA** - Análise por Componentes Principais por Intervalo

**MIR** – Infravermelho médio (do inglês *Mid Infrared*)

**NIR** - Infravermelho próximo (do inglês *Near Infrared*)

**PCA** - Análise por Componentes Principais (do inglês *Principal Component Analysis*)

**PC** - Componente Principal

**PLS** – Regressão por quadrados mínimos parciais (do inglês *Partial Least Squares*)

**iPLS** – Regressão por quadrados mínimos parciais em intervalos

**siPLS** - Regressão por quadrados mínimos parciais por sinergismo de intervalos

**RMSEC** - Raiz Quadrada do Erro Médio de Calibração (do inglês, *Root Mean Square Error of Calibration*)

**RMSECV** - Raiz Quadrada do Erro Médio de Validação Cruzada (do inglês, *Root Mean Square Error of Cross Validation*)

**RMSEP** - Raiz Quadrada do Erro Médio de Predição (do inglês, *Root Mean Square Error of Prediction*)

## Resumo

Neste trabalho foram desenvolvidas metodologias de análise para avaliar a biodegradabilidade do diesel, biodiesel e blendas B10, mediante as técnicas de espectroscopia no infravermelho (FTIR) associada com quimiometria e o emprego de análise de imagem digitais. O objetivo do trabalho foi desenvolver métodos analíticos capazes de avaliar biodegradação em diesel, quantificar biomassa formada na biodegradação de misturas biodiesel/diesel e por fim quantificar o número de colônias de microrganismos deteriorogênicos de biodiesel. Para isso, foram utilizados métodos de análise multivariada como análise por componentes principais (PCA) e regressão multivariada por mínimos quadrados (PLS) e a análise de imagens digitais. Para o estudo de biodegradação de diesel foi investigado diferentes combinações de métodos de pré-processamentos aplicados em espectros de infravermelho médio para analisar amostras de diesel com diferentes teores de enxofre que foram submetidas a estocagem simulada com contaminação microbiana por um período de 40 dias. Para quantificação de biomassa em blendas B10 (10% de biodiesel puro e 90% de óleo diesel) foram utilizadas amostras submetidas a biodegradação por microrganismos e analisadas através de espectroscopia no infravermelho médio e regressão multivariada. Para a quantificação do número de colônias, foram utilizadas imagens digitais de placas de petri contendo colônias da bactéria *Bacillus pumilus* e da levedura *Meyerozyma guilliermondii* e um código Python foi desenvolvido para contar automaticamente as unidades formadoras de colônias a partir das imagens digitais. Os principais resultados obtidos demonstraram que o FTIR-ATR pode ser utilizado para avaliação da degradação de diesel e que a região de 1528 - 1318  $\text{cm}^{-1}$  é a mais indicada e uma possível combinação de pré-processamentos com normalização/segunda derivada/centrado na média. Para a quantificação utilizando espectros FTIR-ATR foram avaliadas as figuras de mérito da metodologia, sendo o método considerado promissor para o seu objetivo. Na contagem de unidades formadoras de colônias, por um método semi-automático, foram avaliadas a calibração e validação em relação à contagem manual clássica, os resultados mostraram curvas de calibração e validação com coeficiente de determinação ( $R^2$ ) de 0,99 e 0,98, respectivamente.

**Palavras-chave:** biodiesel, diesel, microrganismos, análise multivariada, análise de imagem.

## Abstract

In this work, analysis methodologies were developed to evaluate the biodegradability of diesel, biodiesel and B10 blends, using infrared spectroscopy (FTIR) technique associated with chemometry and the use of the digital image analysis. The aim of this work was to develop analytical methods capable of evaluating biodegradation in diesel, quantifying biomass formed in biodegradation of biodiesel/diesel mixtures and finally quantifying the number of colonies of deteriorogenic microorganisms in biodiesel. For this, multivariate analysis methods were used such as principal component analysis (PCA) and multivariate least squares regression (PLS) and the digital image analysis. For the study of diesel biodegradation was investigated different combinations of pre-processing methods applied in mid-infrared spectra to analyze diesel samples with different sulfur contents that were subjected to simulated storage with microbial contamination for a period of 40 days. For quantification of biomass in B10 blends (10% pure biodiesel and 90% diesel oil), samples subjected to biodegradation by microorganisms were used and analyzed using medium infrared spectroscopy and multivariate regression. To quantify the number of colonies, digital images of petri dishes containing colonies of the genus *Bacillus pumilus* and the yeast *Meyerozyma guilliermondii* were used and a Python code was developed to automatically count the colony forming units from the digital images. The main results obtained demonstrated that the FTIR-ATR can be used to assess diesel degradation and that the region of 1528 - 1318  $\text{cm}^{-1}$  is the most suitable and a possible combination of pre-processing with normalization/second derivative/centered on average. For the quantification using FTIR-ATR spectra, the figures of merit of the proposed method were evaluated, being considered appropriate for its objective. The counting of colony forming units, by a semi-automatic method, the calibration and validation were compared with the classic manual counting. The results showed calibration and validation curves with coefficient of determination ( $R^2$ ) of 0.99 and 0.98, respectively.

**Keywords:** biodiesel, diesel, microorganisms, multivariate analysis, image analysis.

# 1 Introdução

O óleo diesel é o principal combustível utilizado no transporte rodoviário de passageiros, produtos industriais e na produção agrícola no Brasil. Trata-se de um derivado do petróleo, sendo assim uma fonte finita e não renovável.<sup>1</sup> A busca por alternativas ao seu uso levou a produção de biodiesel, um combustível derivado de fontes renováveis, como óleo de soja, girassol, canola e gordura animal (sebo bovino).<sup>2,3</sup>

No ano de 2005, com base na lei nº 11.097, o biodiesel passou a fazer parte da matriz energética, podendo ser adicionado facultativamente ao diesel e a partir de 2008 a adição passou a ser obrigatória em um percentual mínimo de 2%, essa mistura foi denominada B2 (2% de biodiesel e 98% de diesel) e o percentual foi aumento sucessivamente, atualmente porcentagem de biodiesel mínima adicionado ao diesel é de 10%.<sup>4,5</sup>

O biodiesel é tecnicamente competitivo com o diesel convencional (petroquímico), tendo algumas vantagens sobre este, como: origem nacional, favorecendo a indústria agrícola local, a biodegradabilidade, redução na emissões de material particulado, monóxido de carbono, baixas ou nenhuma quantidade de enxofre e compostos aromáticos.<sup>6,7</sup> No entanto, a adição do biodiesel ao diesel, agrega novas características intrínsecas à mistura, como: aumento da higroscopicidade, poder de solvência e tendência natural à oxidação, o que pode gerar particulados de natureza orgânica e inorgânica.<sup>8</sup>

A biodegradação do biodiesel é favorecida pela natureza da composição química (mistura de ésteres de ácidos graxos) que podem ser mais facilmente reconhecidos pelos microrganismos e a higroscopicidade.<sup>9</sup> Segundo relatado na literatura, a adição de biodiesel ao diesel promove um maior incremento de populações microbianas, principalmente quando existe a presença de água.<sup>9-12</sup> Levando-se em consideração o ponto de vista ambiental, em casos de derramamento do biodiesel, sua característica de biodegradabilidade, é uma vantagem frente ao diesel. No entanto em casos de armazenagem essa característica passa a ser um problema, uma vez que compromete o tempo de estabilidade do produto.<sup>13</sup>

Uma grande variedade de microrganismos já foram isolado, identificados e avaliados o crescimento em óleo diesel, biodiesel e misturas BX simulando a estocagem. A composição química dos combustíveis é um dos fatores responsável por esta característica pois a presença de hidrocarbonetos alifáticos e aromáticos no óleo diesel, bem como os ésteres presentes no biodiesel, podem ser fontes de carbono e energia para o crescimento e proliferação de bactérias e fungos.<sup>14-16</sup>

Outros fatores que favorecem o desenvolvimento de microrganismos deteriorogênicos dos combustíveis são o tempo de armazenamento e a densidade de microrganismo nativos.<sup>17</sup> As principais consequências causadas pela atividade microbiana é a perda da qualidade do produto final, uma vez que a biomassa formada durante o processo de biodeterioração pode ocasionar o entupimento de filtros de combustíveis, reduzindo assim a sua vida útil, deterioração de tubulações e tanques de armazenamento nas refinarias e usuário final.<sup>18-20</sup>

Diversos métodos analíticos são utilizados para avaliação de contaminação microbiana em combustíveis. Dentre eles, a contagem em placa é o método mais comum utilizado para detecção e quantificação de colônias de bactérias e fungos em amostras contaminadas.<sup>21</sup> Técnicas moleculares são empregadas para a identificação de microrganismos<sup>22-24</sup> e a quantificação de biomassa formada em amostras contaminadas é feita por gravimetria (peso seco).<sup>25,26</sup>

Devido ao avanço tecnológico, houve um desenvolvimento de técnicas analíticas que possibilitaram a união entre a química analítica e a microbiologia. Atualmente a caracterização de produtos de biodegradação é possível ser avaliada através da cromatografia a gás com espectrometria de massa acoplada (GC-MS)<sup>27,28</sup> e da cromatografia a gás com detector de ionização em chama (GC-FID)<sup>29</sup> e da ressonância nuclear de hidrogênio ( $H^1$  NMR).<sup>30</sup>

Recentemente, a aplicação de espectroscopia no infravermelho médio associada com quimiometria vem sendo utilizada para a avaliação de processos de biodegradação em combustíveis.<sup>15,31</sup> Outra técnica que vem sendo proposta é a análise por imagens digitais, cuja metodologia foi utilizada para diferenciar tipos de bactérias<sup>32</sup> e, quantificar colônias de microrganismos.<sup>33</sup> Ambas são técnicas de análise rápida, de baixo custo e que exigem pouco ou nenhum preparo de amostra quando comparadas à outros métodos instrumentais.

Os métodos baseados em espectroscopia no infravermelho médio e análise de imagens podem incrementar ou substituir técnicas convencionais no estudo da biodegradação de combustíveis, auxiliando na aquisição de informações sobre a biodegradabilidade de diesel, biodiesel e na mistura diesel/biodiesel de forma rápida e sem altos custos para garantia da qualidade destes combustíveis.

## **2 Objetivo**

### **2.1 Geral**

O objetivo deste estudo foi desenvolver metodologias analíticas utilizando espectroscopia no infravermelho médio associada com quimiometria e o emprego da análise de imagens digitais para avaliar a biodegradabilidade do diesel, biodiesel e blendas B10 (10% biodiesel e 90% diesel).

### **2.2 Específicos**

I – Aplicar pré-processamentos espectrais e Análise por Componentes Principais (PCA) de diesel com diferentes teores de enxofre submetidos a biodegradação;

II - Quantificar a biomassa produzida na biodegradação de blendas B10 usando Infravermelho com Transformada de Fourier por Reflexão Total Atenuada (FTIR-ATR) e Regressão por Mínimos Quadrados Parciais (PLS);

III – Empregar imagens digitais obtidas por um escâner de mesa para a contagem de colônias de microrganismos deteriogênicos de biodiesel.

## 3 Revisão Bibliográfica

### 3.1 Diesel

O óleo diesel é um dos principais combustíveis líquidos derivado do petróleo, composto principalmente por uma mistura complexa de hidrocarbonetos parafínicos, naftênicos, olefínicos e aromáticos, variando de 8 a 28 átomos de carbono e, em menor concentração substâncias contendo oxigênio, nitrogênio, enxofre e metais.<sup>34,35</sup> É utilizado nos motores de combustão interna e ignição por compressão (motores ciclo Diesel) em veículos rodoviários, ferroviários e marítimos e em geradores de energia elétrica. Atualmente é o combustível mais utilizado no país, impulsionado pela produção agrícola e também devido a predominância da rede rodoviária no transporte de produtos e passageiros. O volume comercializado em 2019 foi de aproximadamente 57,3 milhões de m<sup>3</sup> de diesel.<sup>36</sup>

A formulação do óleo diesel é feita pela mistura de várias frações provenientes dos estágios de processamento do petróleo bruto. A proporção de cada componente é ajustada para garantir as especificações de qualidade estabelecidas pela legislação vigente. A composição química específica do óleo diesel depende do processo de produção na refinaria, origem do petróleo bruto, adições de frações do processo de craqueamento e adição de aditivos para aumentar a eficiência do motor. No Brasil, o óleo diesel é um derivado do tipo parafínico-naftênico do petróleo.<sup>37</sup>

O teor de enxofre no diesel é considerado um dos principais contribuintes para a poluição do ar atmosférico pela emissão de gases de óxidos de enxofre, sendo que a sua redução é uma exigência mundial.<sup>38</sup> Uma das especificações de qualidade do diesel regulamentada pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), é o teor de enxofre. O diesel brasileiro apresenta três níveis de enxofre na sua composição: o diesel S10 (10 ppm de enxofre) e o diesel S500 (500 ppm de enxofre) utilizados em veículos automotivos e em máquinas agrícolas, industriais e construção civil e também existe o diesel S1800 (1800 ppm de enxofre) utilizado no transporte ferroviário, mineração e geração de energia elétrica.<sup>39,40</sup>

Com a utilização do biodiesel, o óleo diesel brasileiro foi dividido em duas classes: o tipo A, que é sem adição de biodiesel e o tipo B que é o óleo diesel acrescido de biodiesel.<sup>39</sup>

### 3.2 Biodiesel

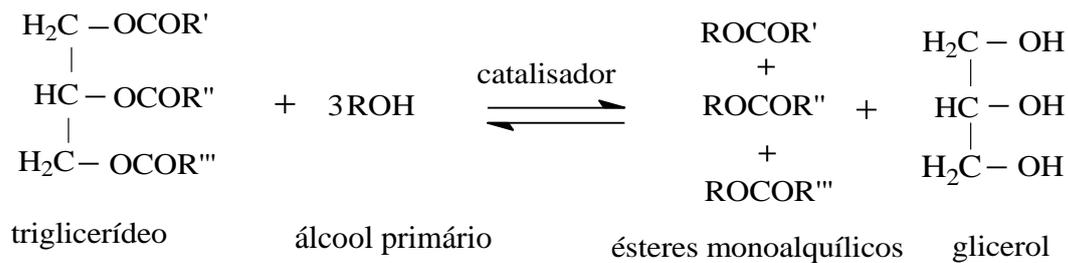
A crescente demanda de energia no mundo industrializado proveniente do aumento na produção agrícola, indústria e setor de transporte, combinado com a escassez dos recursos de reservas de petróleo e impactos ambientais causados pela queima dos combustíveis fósseis, motivou diversas pesquisas em fontes alternativas de energia, estima-se que o consumo mundial de energia comercializada no mundo deverá aumentar em 48% de 2012 a 2040.<sup>41</sup> Nesse cenário, o biodiesel ganhou ênfase por apresentar propriedades semelhantes ao óleo diesel, podendo ser utilizados nos motores de ignição por compressão sem modificações. Pode ser usado puro ou misturado com diesel de petróleo devido à sua completa miscibilidade em todas as proporções de mistura e compatibilidade com a distribuição de combustível existente na infraestrutura.<sup>3,42</sup>

O biodiesel é comercializado misturado ao diesel formando o óleo diesel tipo B. As misturas de diesel e biodiesel são conhecidas como BXX, em que XX indica a quantidade de biodiesel na mistura. Assim, o B100 corresponde ao biodiesel puro, e um B10 seria 10% de biodiesel e 90% de diesel em volume.

A partir de 2008, o óleo diesel tipo B passou a ser comercializado de forma que o percentual mínimo de biodiesel era de 2%, essa mistura foi denominada B2 (2% biodiesel e 98% diesel). Este percentual mínimo está em crescimento: em 2010 passou para 5% (B5), passando para 6% ou B6 em julho de 2014, 7% ou B7 em novembro de 2014, 11% ou B11 em 2019 e em 2020 passou a ser de 12% ou B12. Estima-se que o teor de biodiesel no diesel deve chegar a 15% em 2023.<sup>4,43</sup>

O biodiesel é uma mistura de ésteres alquílicos de ácidos carboxílicos de cadeia longa produzido através de uma reação de transesterificação de ácidos graxos de gorduras de origem vegetal ou animal.<sup>2,44</sup> Uma das maiores vantagens do biodiesel frente a outros combustíveis é a variedade de matérias primas disponível para sua produção que inclui óleo de soja, algodão, milho e gordura animal. As mais utilizadas no país são o óleo soja com 71,21% de participação na produção devido à sua grande disponibilidade, seguido pela gordura bovina 11,51%.<sup>45</sup>

A produção em escala industrial é realizada através de uma reação de transesterificação, como ilustrado na Figura 1, onde um triglicerídeo reage com um álcool de cadeia curta, o metanol é o mais utilizado em função da sua maior reatividade.



**Figura 1:** Reação de transesterificação de um triglicerídeo para produção de biodiesel.

Para uma reação de transesterificação estequiometricamente completa, uma proporção molar 3:1 de álcool por triglicerídeo é necessária. Pode-se utilizar um catalisador alcalino ou ácido, porém a catálise alcalina homogênea é a mais empregada, obtendo assim uma mistura de ésteres metílicos de ácidos graxos, que constitui o biodiesel, e o coproduto glicerol.<sup>46</sup>

### 3.3 Biodegradação dos Combustíveis

A presença de bactérias e fungos em combustíveis como diesel, biodiesel, gasolina e querosene é amplamente relatada em diversos estudos.<sup>23,47,48</sup> Uma grande variedade de microrganismos já foi isolada e identificada, sendo que muitos apresentam a capacidade de crescer e degradar diferentes fontes de combustíveis.<sup>14-16</sup>

Os microrganismos são ubíquos (estão em todos os lugares) na natureza e alguns são capazes de consumir hidrocarbonetos como fonte de energia e carbono. Podem degradar vários tipos de hidrocarbonetos - de cadeia curta, de cadeia longa e vários compostos aromáticos, incluindo os hidrocarbonetos aromáticos policíclicos.<sup>24,49</sup> Portanto os combustíveis derivados do petróleo e o biodiesel constituem uma importante e potencial fonte de carbono.<sup>50,51</sup>

A degradação dos combustíveis principalmente durante o armazenamento é um problema econômico reconhecido com consequências causadas pela atividade microbiana como o comprometimento da qualidade final do produto, entupimento e redução da vida útil de filtros, injetores de combustíveis ocasionados pela biomassa formada durante o processo de biodeterioração,<sup>19,20</sup> bem como a corrosão de tubulações e tanques de armazenamento nas refinarias, distribuidoras e para o usuário final. Neste contexto, o monitoramento e controle de qualidade é de suma importância para garantir que os combustíveis estejam adequados para comercialização no mercado.

### *3.3.1 Fatores que Contribuem para a Biodegradação*

#### *3.3.1.1 Água*

A disponibilidade de água é um fator determinante para o crescimento de microrganismos, pois estes são incapazes de sobreviver na ausência de água.<sup>19,52</sup> A presença de água em tanques de armazenamento propicia condições ideais para o crescimento de populações microbianas, níveis de cerca de 1% é suficiente para o desenvolvimento microbiano.<sup>53</sup> A água pode estar na forma dissolvida; emulsificada ou como fase livre, no fundo dos tanques.

A entrada de água nos tanques de armazenamento pode ocorrer de diferentes maneiras tais como carreada pelo combustível, pela condensação do ar nas paredes, pelo metabolismo microbiano, entrada pelos respiros, durante a lavagem dos tanques ou ainda pode ser colocada intencionalmente, como lastro.

A densidade da água é maior que a do combustível, conseqüentemente fica depositada no fundo dos tanques, formando um sistema bifásico “combustível-água”. Os microrganismos se desenvolvem, então, no fundo dos tanques, na interface entre a água e o combustível onde todos os requisitos para o crescimento são fornecidos: uma substância orgânica como fonte de energia e nutrientes inorgânicos.<sup>52,53</sup>

#### *3.3.1.2 Tempo de Armazenamento*

A deterioração dos combustíveis é mais problemática em tanques de armazenamento em que a taxa de rotatividade é baixa. A permanência do combustível por um período superior a trinta dias, pode ocasionar a proliferação de bactérias e fungos que ocorre na interface entre água e o óleo devido ao maior tempo de contato entre o combustível e as populações microbianas.<sup>18</sup>

Não há como prevenir a entrada de microrganismos em tanques de armazenamento de combustíveis, a contaminação pode ocorrer pelo solo, via ar, água de lavagem utilizadas nos tanques de armazenamento, tubulações contaminadas, biofilmes presentes nas paredes dos tanques (decorrente de limpeza insuficiente) ou estão presentes no próprio combustível.<sup>14</sup>

### 3.3.2 *Biodegradação do Diesel e Biodiesel*

A maioria dos microrganismos é capaz de oxidar hidrocarbonetos com comprimento de cadeia de C<sub>10</sub> a C<sub>20</sub>, portanto o diesel que contém vários hidrocarbonetos alifáticos com comprimentos de cadeia variando de C<sub>8</sub> a C<sub>26</sub> é propenso a ataques microbianos.<sup>16,54,55</sup> O diesel é o combustível que apresenta os mais variados problemas de contaminação microbiana, sendo que várias espécies de bactérias são capazes de usar o diesel como fonte de carbono.<sup>56-59</sup>

A biodegradabilidade do biodiesel está associada a sua composição química que é composta por ésteres graxos que podem ser facilmente oxidados e também a ausência de hidrocarbonetos aromáticos e enxofre.<sup>60,61</sup> Diversos estudos investigaram a degradabilidade da mistura diesel/biodiesel e apontaram que a adição do biodiesel aumentou a biodegradação da mistura e que o aumento na fração de biodiesel aumenta a degradação.<sup>10,62</sup>

Um dos motivos do aumento da biodegradabilidade com o aumento da fração de biodiesel está relacionado com o aumento da higroscopicidade da mistura. O biodiesel apresenta uma polaridade maior que o diesel, devido à presença de ligações de éster, assim tem maior afinidade com a água do que o diesel. Alguns estudos indicam que a presença de ésteres metílicos de ácidos graxos pode aumentar a solubilidade e a dispersabilidade de água nos combustíveis.<sup>8,63</sup>

### 3.4 Análise da Biodegradação em Combustíveis

Os parâmetros de qualidade do óleo diesel devem atender a especificações da resolução ANP N° 50 DE 23/12/2013 e segundo a legislação atual, é obrigatória a adição de 10% de biodiesel ao diesel, com variação máxima de  $\pm 0,5\%$  em volume. A mistura de diesel/biodiesel comercializadas no Brasil, bem como os parâmetros de qualidade do biodiesel devem atender às especificações estabelecidas na Resolução ANP N° 47/2012, que tem como base os padrões de qualidade elaborados pela American Society of Testing and Materials (ASTM) através da norma ASTM D6751 e pelo Comité Europeu de Normalização (Comité Européen de Normalisation – CEN) através da norma EN1421428.<sup>39,64</sup>

A Resolução ANP N° 47/2012, estabelece que o biodiesel produzido tenha um prazo máximo de um mês, a contar da data de certificação, para ser comercializado. Para a liberação do produto diversos parâmetros de qualidade são avaliados, entre eles a análises de massa específica a 20 °C, o teor de água, o índice de acidez e a estabilidade à oxidação a 110 °C.

A bioprospecção da contaminação microbiana em tanques armazenadores de combustível pode ser conduzida tanto na fase óleo, fase água ou da interface óleo/água conforme sugerido na Norma da ASTM 6469-99- Guia para Contaminação Microbiana em Combustíveis e Sistemas de combustível.<sup>65</sup> Os microrganismos podem ser isolados a partir da filtragem do combustível e ou da fase aquosa, conforme preconizado em Normas ASTM D 6974-09 e ASTM E 1326.<sup>66,67</sup>

O método preconizado nas normas para avaliar contaminação por microrganismos envolve a quantificação de colônias de bactérias e fungos pela contagem em placa, que consiste na contagem direta do número de unidades formadoras de colônias em placas com um meio de cultura sólido. Uma amostra de volume conhecido do combustível passa por um processo de filtração utilizando uma membrana com porosidade de 0,45 $\mu$ m e esta membrana é inoculada em placas com meios de cultura sólidos e o número de colônias é contado após o crescimento. E a técnica utilizada na quantificação de biomassa formada durante o processo de biodegradação é a gravimetria.<sup>25,26</sup>

Atualmente, técnicas instrumentais estão sendo aplicadas para a análise da degradação em combustíveis, a caracterização de produtos de biodegradação é possível ser avaliada através da cromatografia a gás com espectrometria de massa acoplada (GC-MS)<sup>27,28</sup> e da cromatografia a gás com detector de ionização em chama (GC-FID)<sup>29</sup> e da ressonância magnética nuclear de hidrogênio ( $H^1$  NMR).<sup>30</sup>

E o emprego da Espectroscopia no Infravermelho com Transformada de Fourier (FTIR) em conjunto com ferramentas quimiométricas vem se destacando no tratamento de dados espectrais para estudar a degradação de petróleo e seus derivados.<sup>27,68</sup> E também na análise dos produtos de degradação em biodiesel e suas misturas com o diesel. Em seu trabalho, Zimmer e colaboradores (2013), monitoraram a degradação de combustíveis ao longo do tempo e avaliaram a fase óleo através de espectroscopia no infravermelho médio em conjunto com análise multivariada utilizando Análise Hierárquica de Cluster (HCA). O dendrograma evidenciou a degradação do biodiesel ao longo do tempo.<sup>69</sup>

Utilizando uma abordagem semelhante, Bücken e colaboradores (2014) avaliaram a ação de diferentes biocidas no controle da biodegradação da blenda diesel/biodiesel (B10), onde a fase óleo foi analisada através de FTIR e HCA, o dendrograma evidenciou a ação diferenciada do biocida de acordo com a concentração utilizada no experimento.<sup>23</sup> Da mesma forma Gassen e colaboradores (2015), utilizaram a análise de FTIR e PCA para avaliar o efeito da adição de biodiesel no diesel no crescimento do microrganismo *P. variotii*, as análises mostraram a degradação do combustível.<sup>15</sup>

Embora já existam metodologias para avaliar a biodegradação de combustíveis o desenvolvimento de novos métodos analíticos pode incrementar ou substituir técnicas convencionais de forma mais eficiente no estudo da biodegradação de diesel, biodiesel e blends diesel/biodiesel.

Uma descrição de Espectroscopia no Infravermelho com Transformada de Fourier, quimiometria e análise de imagens digitais será agora apresentado.

### 3.5 Espectroscopia no Infravermelho com Transformada Fourier (FTIR)

A Espectroscopia no Infravermelho com Transformada de Fourier (FTIR) é uma técnica instrumental utilizada para analisar e identificar a estrutura química de compostos orgânicos e inorgânicos que apresentam ligações covalentes, através da medida da absorção da radiação na região do infravermelho.<sup>70</sup>

A maioria dos compostos orgânicos e inorgânicos apresentam ligações covalentes que absorvem várias frequências de radiação eletromagnética na região do infravermelho. A região do espectro eletromagnético que corresponde ao infravermelho vai de 12.000 a 10  $\text{cm}^{-1}$  e é dividido três regiões: infravermelho próximo (NIR), infravermelho médio (MIR) e infravermelho distante (FIR). A Tabela 1, apresenta os intervalos correspondente a cada região.<sup>71</sup>

**Tabela 1:** Regiões do espectro de infravermelho.

Região	Número de onda ( $\text{cm}^{-1}$ )	Comprimento de onda (nm)
NIR	12.800 – 4.000	780 – 2.500
MIR	4.000 - 400	2.500 – 5.000
FIR	200 - 10	5.000 – 100.000

O fundamental para que haja absorção de energia no infravermelho é que a molécula apresente variação do momento de dipolo elétrico que é em decorrência do seu movimento vibracional ou rotacional, assim o espectro de infravermelho é consequência da absorção de energia em frequências específicas que causam um aumento na amplitude dos movimentos vibracionais ou rotacionais.

Em um processo de absorção de energia, quando uma molécula absorve radiação no infravermelho são excitadas para um estado de maior energia. Na temperatura ambiente a maioria das moléculas estão em seus níveis de energia vibracional fundamental, quando ocorre absorção de radiação no infravermelho são excitadas para um estado de maior energia.<sup>72</sup>

A espectroscopia de infravermelho é empregada na análise de compostos orgânicos, pois a energia de transição correspondente a mudanças no estado de energia vibracional para muitos grupos funcionais está localizada na região de infravermelho médio, assim, o aparecimento de uma banda de absorção nesta região pode ser usado para determinar se grupos funcionais específicos que estão presentes em uma amostras.<sup>73,74</sup> É amplamente utilizada na análise de combustíveis.<sup>75-78</sup>

### 3.6 Quimiometria

A análise multivariada de dados deu origem a quimiometria, uma área altamente interdisciplinar que abrange métodos matemáticos e estatísticos aplicados para a avaliação e interpretação de dados químicos visando extrair o máximo de informações relevantes através da análise dos dados e também empregada no planejamento experimental.<sup>79,80</sup>

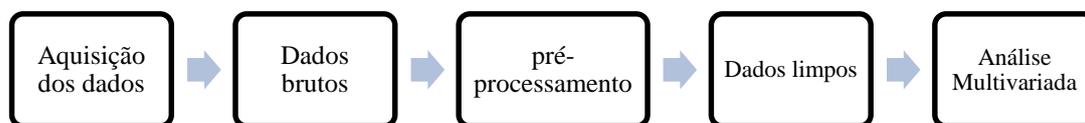
O termo quimiometria foi proposto pelo químico orgânico Svante Wold em 1971 e em colaboração com Bruce R. Kowalski criaram em 1974 a *International Chemometrics Society*, juntos são considerados os pioneiros na publicação de trabalhos na área de quimiometria. Em 1974, Wold definiu formalmente a quimiometria como “ a arte de extrair informações quimicamente relevantes a partir de dados produzidos em experimentos químicos”<sup>80,81</sup> e no ano de 1975 em uma publicação de Kowalski a quimiometria foi definida como uma disciplina da química.<sup>82</sup>

Desde então, duas razões contribuíram para o desenvolvimento da quimiometria, o avanço na instrumentação analítica que passou a gerar resposta multivariada (ex: espectro de infravermelho que contém em média 3.000 comprimentos de onda) e o progresso da informática associada à disponibilidade de computadores em laboratório de química.<sup>83</sup>

A quimiometria compreende a utilização de ferramentas estáticas para a análise multivariada de dados visando extrair informações, identificar padrões e características a partir de um grande volume de dados, é significativamente empregada dentro da química analítica,<sup>84</sup> pois a instrumentação analítica envolve a obtenção de um largo número de variáveis, como cromatografia em fase gasosa (CG),<sup>85-87</sup> cromatografia em fase líquida de alto desempenho (HPLC),<sup>88,89</sup> ressonância magnética nuclear (RMN)<sup>90-93</sup> e espectroscopia vibracional<sup>94-96</sup> que produzem uma grande quantidade de dados multivariados,  $n$  amostras podem gerar milhares de sinais medidos (variáveis), exemplo valores de absorbância.

O trabalho desenvolvido dentro da quimiometria pode ser dividido em etapas como pode ser observado na Figura 2, que apresenta um fluxograma com as etapas usuais envolvida na análise de dados, que inicia pela aquisição dos dados, que podem ser espectros ou cromatogramas, e após há necessidade de um pré-tratamento ou também conhecido como pré-processamentos dos dados para remover variações ocasionadas na

aquisição dos dados (ex: ruído instrumental) e após ferramentas quimiométricas são aplicadas na análise multivariada.<sup>80</sup>



**Figura 2:** Fluxograma das etapas envolvidas na aplicação da análise multivariada de dados.

Para realizar a análise multivariada, os dados são organizados em uma matriz, as linhas representam as amostras e as colunas as variáveis (valores instrumentais medidos), a dimensão da matriz será definida pelo número de amostras analisadas e o número de variáveis medidas.

Diversos métodos de análise multivariada são utilizados dentro da quimiometria, que podem ser divididos em dois grupos de acordo com o objetivo da análise, propósitos qualitativos (análise exploratória e reconhecimento de padrões) e quantitativos (regressão multivariada ou calibração multivariada).<sup>97</sup>

Os métodos de análise exploratória de dados mais empregado são HCA (*Hierarchical Cluster Analysis*) e PCA (*Principal Component Analysis*). Para calibração multivariada, PCR (*Principal Component Regression*) e PLS (*Partial Least Squares*), e métodos de reconhecimento de padrão aplicados em modelos de classificação, KNN (*k-Nearest Neighbor*), SIMCA (*Soft Independent Modelling of Class Analogy*) e PLS-DA (*PLS for Discriminant Analysis*).<sup>98-101</sup>

### 3.6.1 Pré-processamento dos dados

O pré-processamento dos dados é uma etapa fundamental dentro da quimiometria, que visa remover ou minimizar as fontes de variabilidade espectral associado ao ruído aleatório e a variações sistemáticas no conjunto de dados e também pode destacar diferenças sutis entre as amostras.<sup>102</sup>

Nas etapas de análise dos dados, Figura 2, o pré-processamento é a etapa prévia aplicado aos dados brutos para reduzir as variações quimicamente irrelevantes com o objetivo de melhorar a exatidão e precisão da análise multivariada (regressão multivariada, modelos de classificação ou análise exploratória).<sup>103,104</sup>

Os espectros são apresentados como uma série de picos, que são sobrepostos ao ruído. Este é uma perturbação indesejada em um sinal e que pode ser ocasionado, por exemplo,

devido a sensibilidade dos detectores dos instrumentos, ou a fonte de excitação (fonte de luz eletromagnética). A elevação do ruído espectral pode mascarar características espectrais sutis, assim a qualidade espectral é geralmente avaliada pela relação sinal-ruído (SNR).<sup>105</sup>

As variações sistemáticas incluem:<sup>106</sup>

- Desvios da linha de base devido aos efeitos de dispersão da luz relacionados à variação no tamanho das partículas;
- Deslocamento da linha de base devido a variações instrumentais;
- Diferenças na intensidade do sinal global devido a variações na quantidade de amostra.

Os espectros de FTIR- ATR geralmente são afetados por variações sistemáticas, que causam desvios na linha de base e tamanho ou quantidade de amostra vai afetar o valor de absorbância. Estas variações prejudicam a análise dos dados, interferindo na extração de informações relevantes, bem como suas interpretações.<sup>107</sup>

Existem vários métodos de pré-processamento que podem ser aplicadas com o objetivo de reduzir tanto as variações aleatórias (ruído) quanto as informações sistemáticas indesejáveis. Não existe uma regra que determine a utilização de uma técnica ou a combinação de técnicas. A escolha de um método ótimo de pré-processamento ou combinações de métodos vai depender do conjunto de dados e no objetivo da análise e tendo o cuidado para não distorcer os resultados e comprometer a análise dos resultados. Uma estratégia de tentativa e erros é o método mais comumente usado, onde a aplicação de diferentes métodos de pré-processamento e suas combinações é selecionado de acordo com o desempenho através de modelos de regressão multivariada ou análise exploratória de dados.<sup>103,105,108</sup>

Alguns métodos matemáticos são amplamente empregados no pré-processamento de dados de espectroscopia no infravermelho. A normalização é aplicada para reduzir variações relacionadas ao tamanho ou quantidade de amostras.<sup>106</sup> A aplicação de derivadas reduz diferenças de linha de base e também podem resolver sobreposição espectral de bandas.<sup>107</sup> E centrar os dados na média, um dos métodos de pré-processamento mais comuns aplicado para remover deslocamento espectral, permitindo focar a análise nas diferenças entre as amostras, ao invés da direção geral da variância.<sup>109</sup>

### 3.6.1.1 Normalização

A normalização é um procedimento muito importante, que visa remover variações sistemáticas, que podem estar associadas a quantidade de amostra, efeitos de flutuações de energia (fonte de radiação MIR) e espalhamento. É um procedimento que corrige disparidade nos níveis de intensidade do sinal analítico, por exemplo, a intensidade de uma dada banda de FTIR da mesma amostra será tão semelhante quanto possível ao longo dos espectros registrados sob os mesmos parâmetros experimentais, mas que pode ser afetado pelas variações descritas acima.<sup>102</sup>

Nesse pré-processamento, os valores de cada uma das variáveis de uma dada amostra  $i$  são divididos por um fator de normalização, que pode ser a norma dessa amostra, representada por  $\|\mathbf{X}_i\|$ , equação 1. Assim, a normalização iguala a magnitude de cada amostra dando o mesmo peso para todas as absorvâncias.<sup>110,111</sup>

$$x_{ij(norm)} = \frac{x_{ij}}{\|\mathbf{X}_i\|} \quad (1)$$

As normas mais utilizadas são:

$\|\mathbf{X}_i\|_\infty = \max|x_{ij}|$  Denominada norma *sup*, ou norma infinita,  $l_\infty$ ;

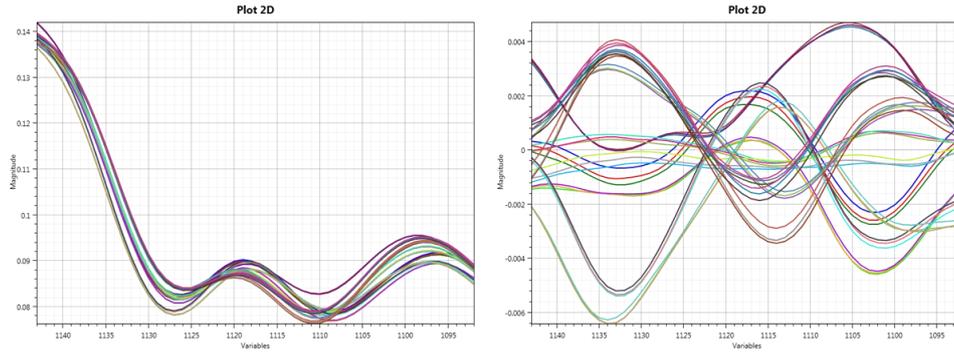
$\|\mathbf{X}_i\|_1 = \sum_{j=1}^J |x_{ij}|$  Denominada de norma um,  $l_1$ ;

$\|\mathbf{X}_i\|_2 = \sqrt{\sum_{j=1}^J x_{ij}^2}$  Norma Euclidiana ou norma dois,  $l_2$ .

Cada uma das normas utilizadas causa um efeito, a norma *sup*, a resposta máxima de cada uma das amostras se torna igual a 1. A norma um é dada pela soma dos valores absolutos das variáveis da amostra  $i$ , que corresponde à área sob a curva definida pelos dados da  $i$ -ésima amostra. Na norma dois cada espectro terá comprimento igual a 1.<sup>110</sup>

### 3.6.1.2 Centrar na Média

A centralização na média é um dos métodos de pré-processamento mais comum e utilizado em análise de dados multivariados.<sup>80</sup> Este procedimento remove o deslocamento espectral de cada espectro medido e permite avaliar as diferenças presentes nas amostras, que é a variação relevante. Ao centrar os dados na média, a intensidade do sinal flutua em torno de zero do valor médio dos sinais,<sup>109</sup> como mostrado na Figura 3.



**Figura 3:** Dados centrados na média: a) espectros brutos e b) espectros centrados na média

Centrar os dados na média consiste em calcular o espectro médio e subtrair este de todos os espectros das amostras, ou seja, o valor médio de cada coluna da matriz de dados é determinado de acordo com a equação 1, e subsequentemente subtraindo esse valor médio de todos os elementos nessa coluna, equação 2. Resultando em uma matriz de dados que possui novas colunas com médias zero.<sup>110</sup>

$$\bar{x}_j = \frac{1}{n} \sum_{i=0}^n x_{ij} \quad (1)$$

$$x_{ij(cm)} = x_{ij} - \bar{x}_j \quad (2)$$

onde:

$\bar{x}_j$  é a média dos dados

$n$  é o número de amostras que fazem parte do conjunto de dados

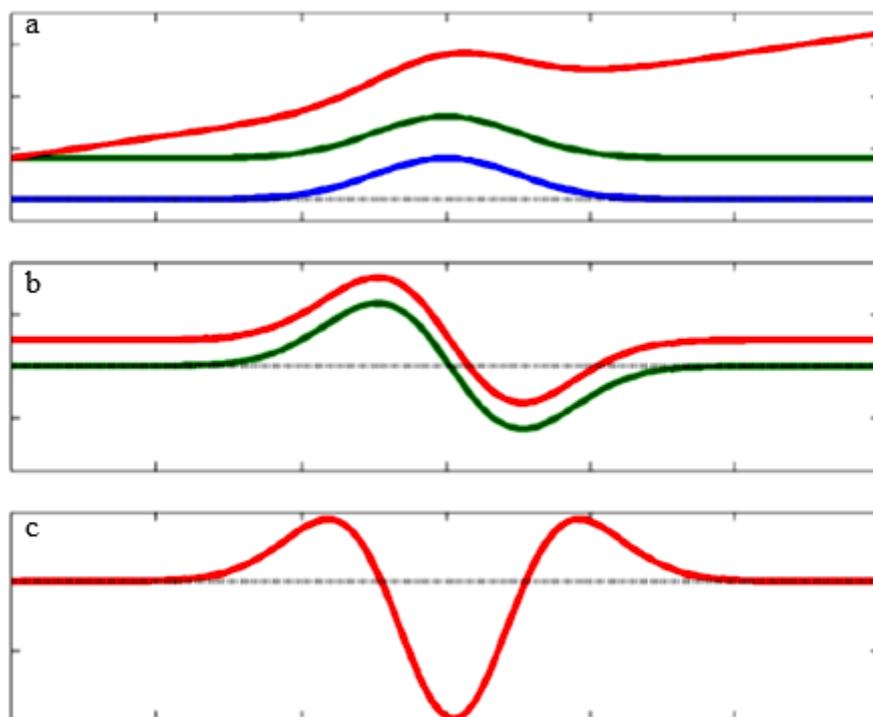
O efeito desse pré-processamento é apenas uma translação de eixos para um valor médio de cada um deles, a estrutura dos dados não é modificada.

### 3.6.1.3 Primeira e Segunda derivada

O pré-processamento com derivadas é amplamente empregado em espectroscopia vibracional para correção de linha de base, a primeira derivada reduz o deslocamento espectral constante na linha de base (*offset*) e a segunda derivada remove inclinações de linha de base (tendência) e também podem resolver sobreposição espectral com base na diferenciação dos espectros.<sup>111-114</sup>

A Figura 4a mostra a representação gráfica de uma curva com inclinação (vermelho) e deslocamento de linha de base (verde) e uma curva sem qualquer variação (azul), e nas

Figuras 4b e 4c o resultado da aplicação da primeira e segunda derivada que eliminaram as variações na linha de base.



**Figura 4:** Curva com uma inclinação e um deslocamento de linha de base (a) , (b) aplicação da primeira derivada e (c) aplicação da segunda derivada.<sup>104</sup>

A aplicação de derivadas não altera a relação da área abaixo de um pico do espectro e a concentração, a correlação quantitativa entre a concentração e as medidas espectrais permanece igual. Uma limitação dos métodos derivativos é a diminuição da relação sinal/ruído, devido à alta sensibilidade do método. O método mais empregado para o cálculo de derivadas de um espectro é o método de Savitzky-Golay.<sup>115</sup>

### 3.6.2 Análise por Componentes Principais (PCA)

Análise por componentes principais (PCA) é um método de análise de dados multivariados não supervisionado aplicado para avaliar as inter relações entre um grande número de variáveis com o objetivo de reconhecer padrões na distribuição das amostras, avaliar a relação entre amostras e as variáveis e ainda detectar a presença de amostras que apresentam um comportamento distinto (*outliers*).<sup>116</sup>

A PCA é uma técnica que consiste na redução da dimensionalidade do conjunto de dados originais, ou seja, transforma um grande número de variáveis em um número menor de variáveis. Essa redução é obtida através de combinações lineares das variáveis originais, que são agrupadas em função das correlações existentes, com o objetivo de

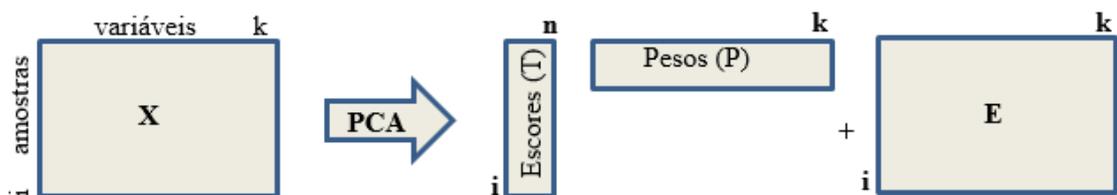
representar as variações presentes em muitas variáveis, através de um número menor de novas variáveis, chamadas de componentes principais (PC's).

A base do método é a correlação entre as variáveis, agrupando aquelas que estão altamente correlacionadas em componentes principais, preservando ao máximo as informações originais. Um pressuposto básico para a PCA é que a primeira componente principal deve conter a máxima variância possível, a segunda componente representa a variância não explicada na primeira componente e assim por diante. E as PC's são ortogonais entre si.<sup>117,118</sup>

Usando a PCA, por exemplo, é possível reduzir um grande conjunto de dados de espectroscopia com centenas de variáveis em algumas componentes principais que representam a maior parte das informações originais das amostras, tornando a interpretabilidade dos dados mais administrável.

O modelo matemático da PCA pode ser descrito através da Figura 5, a matriz X de dados originais é decomposta em um produto de duas matrizes, onde T é matriz dos escores e contém as novas coordenadas do sistema de eixo (PC's); P é matriz dos pesos em que os elementos de cada coluna correspondem aos coeficientes das combinações lineares das variáveis originais e mais uma matriz E de resíduo, conforme a Equação 4:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (4)$$



**Figura 5:** Estrutura matemática do modelo PCA.

O resultado fornecido pela PCA é visualizado através de um gráfico de escores que mostra a relação entre as amostras, com diferentes agrupamentos de amostras através da identificação de semelhanças e diferenças, tendências e *outliers* (amostras anômalas). E um gráfico de pesos que mostra quais variáveis (colunas) da matriz X definem as componentes principais, permitindo avaliar quais variáveis de medição contribuíram para o padrão apresentado no gráfico de escores.<sup>116</sup>

### 3.6.3 Calibração Multivariada

Em química analítica a quantificação de um constituinte em uma amostra não pode ser medida diretamente, requer o desenvolvimento de um processo de calibração, que relaciona a intensidade de um sinal analítico (ex: absorvância) com a concentração, que é a quantidade do constituinte de interesse presente em uma amostra. O procedimento de calibração envolve uma série de etapas na construção de um modelo matemático que relaciona uma grandeza medida com um conjunto de amostras padrões que apresentam uma concentração conhecida da espécie de interesse.<sup>119,120</sup>

Os métodos de calibração podem ser classificados de acordo com a complexidade ou a dimensionalidade dos dados medidos por amostra em calibração univariada ou calibração de ordem zero e calibração multivariada que pode ser de primeira, segunda, terceira ou quarta dependendo da quantidade de informação medida por amostra, sendo a ordem uma propriedade dos dados instrumentais para uma única amostra que caracteriza o objeto matemático que pode ser construído com os dados.<sup>121,122</sup>

A calibração de ordem zero é o método tradicionalmente empregado na química analítica clássica, é baseada na relação entre a medição de um único valor numérico (escalar), por exemplo a absorvância em um único comprimento de onda, e a propriedade de interesse na amostra e análise de um conjunto de amostras gera um vetor.<sup>123</sup>

Na calibração de primeira ordem um conjunto de dados numéricos, por exemplo um espectro de infravermelho (absorvância em vários comprimentos de onda) é relacionado com a propriedade de interesse presente na amostra, cada espectro pode ser representado por um vetor e o conjunto de amostras pode ser organizado em uma matriz.<sup>124</sup>

Os métodos mais empregado em calibração multivariada para dados de primeira ordem são Regressão Linear Múltipla (RLM), Regressão por Componentes Principais (PCR) e Regressão por Mínimos Quadrados Parciais (PLS), que associados à espectroscopia no infravermelho médio são amplamente empregados para determinações quantitativas de parâmetros físico-químicos em amostras de combustíveis.<sup>125-127</sup>

#### 3.6.3.1 Regressão por Mínimos Quadrados Parciais (PLS)

A regressão por mínimos quadrados parciais é o modelo de calibração mais popular para dados de primeira ordem. A construção do modelo de calibração é baseada na relação linear entre as variáveis dependente (concentração ou propriedade de interesse) e as variáveis independentes (sinal analítico).

A construção do modelo de calibração por PLS é baseado em uma decomposição bilinear, tanto a matriz  $\mathbf{X}$  (variáveis independentes) e a matriz  $\mathbf{Y}$  (variável dependente) são decompostas. O modelo PLS estabelece uma relação linear entre a matriz  $\mathbf{X}$  e  $\mathbf{Y}$ , através de variáveis latentes que são determinadas por combinações lineares, ou seja, as variáveis da matriz  $\mathbf{X}$  e  $\mathbf{Y}$  estão relacionadas entre si por meio dessas variáveis latentes.<sup>128</sup>

O modelo matemático é apresentado nas equações 5 e 6, as matrizes  $\mathbf{X}$  e  $\mathbf{Y}$  são decompostas em uma soma do produto entre dois vetores, escores e pesos, Equações 5 e 6.<sup>129</sup>

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} = \sum \mathbf{t}_h \mathbf{p}_h^T + \mathbf{E} \quad (5)$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{F} = \sum \mathbf{u}_h \mathbf{q}_h^T + \mathbf{F} \quad (6)$$

onde  $\mathbf{T}$  e  $\mathbf{U}$  são as matrizes de escores das matrizes  $\mathbf{X}$  e  $\mathbf{Y}$ ;  $\mathbf{P}$  e  $\mathbf{Q}$  são as matrizes de pesos das matrizes  $\mathbf{X}$  e  $\mathbf{Y}$ ,  $h$  corresponde ao número de variáveis latentes,  $\mathbf{E}$  e  $\mathbf{F}$  correspondem às matrizes de resíduos. Cada variável latente, que relaciona  $\mathbf{X}$  e a propriedade de interesse  $\mathbf{Y}$ , é obtida maximizando a covariância entre os escores das matrizes  $\mathbf{X}$  e  $\mathbf{Y}$ .

Uma relação linear entre as matrizes  $\mathbf{X}$  e  $\mathbf{Y}$  é obtida correlacionando-se as matrizes escores  $\mathbf{u}$  e  $\mathbf{t}$ , através do coeficiente de regressão linear de cada componente principal ( $\mathbf{b}_h$ ), Equação 7:

$$\mathbf{u}_h = \mathbf{b}_h \mathbf{t}_h \quad (7)$$

O coeficiente de regressão linear ( $\mathbf{b}_h$ ) é determinado pela equação 8:

$$\mathbf{b}_h = \frac{\mathbf{u}_h^T \mathbf{t}_h}{\mathbf{t}_h^T \mathbf{t}_h} \quad (8)$$

onde,  $\mathbf{u}$  e  $\mathbf{t}$  são os escores das matrizes  $\mathbf{U}$  e  $\mathbf{T}$ , respectivamente.

A avaliação da eficiência do modelo é baseada na análise da Raiz Quadrada dos Erros Médios (RMSE, do inglês, *Root Mean Square Error*), que compreende os erros: Raiz Quadrada do Erro Médio de Calibração (RMSEC, do inglês, *Root Mean Square Error of Calibration*); Raiz Quadrada do Erro Médio de Predição (RMSEP, do inglês *Root Mean Square Error of Prediction*) e Raiz Quadrada do Erro Médio de Validação Cruzada (RMSECV, do inglês *Root Mean Square Error of Cross Validation*).<sup>124</sup>

### 3.7 Análise de Imagens

Hoje em dia, a análise de imagem, de produtos e processos, é cada vez mais importante devido à sua capacidade de ser rápida, não invasiva e de baixo custo. O processamento digital de imagens é uma disciplina que parte de uma ou de uma sequência de imagens e utiliza técnicas apropriadas para extrair/interpretar informações relevantes para cada aplicação.<sup>130</sup>

O interesse em processamento de imagens surgiu de duas áreas: melhoramento da qualidade para interpretação humana e percepção de máquina extraindo informações das imagens, de forma tratável pelo computador. Atualmente os processos computacionais de melhoramento de imagens para percepção humana têm sido utilizados em diversas áreas, nas quais se destacam a química, biologia e geografia.

Na resolução de problemas de percepção por máquina, busca-se extrair informações adequadas ao processamento computacional. Estas informações geralmente diferem das que são utilizadas pelo homem na interpretação de uma imagem (momentos estáticos, invariantes, coeficientes da transformada de Fourier). Dentre os problemas típicos de percepção por máquina, podemos citar o reconhecimento de padrões, muito aplicada na química analítica moderna.<sup>131</sup>

Uma imagem digital é formada por uma estrutura quadriculada e cada quadrado recebe o nome de pixel e cada pixel possui um valor numérico de intensidade que representa a mistura de três cores primária: vermelho, verde e azul (RGB). Os valores numéricos que representam as cores são empregados para construção de uma matriz de dados com as características de cor de um conjunto de imagens. A informação numérica dos espaços de cores é processada com ferramentas quimiométricas.<sup>132</sup>

Muitos trabalhos estão sendo publicados utilizando a análise de imagens como técnica analítica associada a ferramentas quimiométricas para determinação de metais,<sup>133</sup> corantes em alimentos,<sup>134</sup> detecção e quantificação de adulterantes em vinho, autenticação de vinhos<sup>135</sup>, classificação de biodiesel,<sup>136</sup> na classificação de bactérias<sup>32</sup> e análise de drogas.<sup>137</sup>

E uma metodológica que se tornou um grande tópico na literatura é a contagem de unidades formadoras de colônias de microrganismo por meio de imagens obtidas de placas de Petri. Atualmente vem sendo empregado diferentes meios de aquisição de imagens: câmeras digitais,<sup>138,139</sup> imagens obtidas através de um microscópio acoplado a uma câmera CCD,<sup>140</sup> imagens microscópicas,<sup>141</sup> imagens hiperespectrais<sup>142</sup> e um sistema

de aquisição de imagem de placas de petri em comprimentos de onda próximos ao infravermelho.<sup>143</sup>

Ainda que existem diversas técnicas para a contagem de colônias de microrganismo, nenhuma foi aplicada no estudo de colônias de microrganismos deteriorogênicos de combustíveis.

## 4 Resultados

A apresentação dos resultados deste trabalho foi dividida em capítulos: O capítulo 4.1 descreve a investigação de diferentes métodos de pré-processamento espectral aplicados em espectros de infravermelho médio de amostras de diesel com diferentes teores de enxofre que foram submetidas a estocagem simulada por um período de 40 dias. O capítulo 4.4 apresenta um método de quantificação da biomassa produzida no processo biodegradação durante estocagem simulada de Blendas B10 usando infravermelho com Transformada de Fourier (FTIR) e regressão por mínimos quadrados parciais (PLS) e no capítulo 4.7 é descrito o desenvolvimento de um método de contagem semi-automático para a quantificação do número de colônias da bactéria *Bacillus pumilus* e da levedura *Meyerozyma guilliermondii*, ambas isoladas a partir do óleo diesel, empregando imagens digitais.

## **4.1 Pré-Processamento Espectral e Análise por Componentes Principais (PCA) de Diesel com Diferentes Teores de Enxofre Submetidos a Biodegradação**

### *4.1.1 Objetivo*

Investigar diferentes combinações de métodos de pré-processamento aplicados em espectros de infravermelho médio para analisar amostras de diesel com diferentes teores de enxofre que foram contaminadas com microrganismo e avaliadas por um período de 40 dias. Este trabalho foi desenvolvido em parceria com o Laboratório de Biodeterioração de Combustíveis e Biocombustíveis (LABBIO), UFRGS /RS.

### *4.1.2 Materiais e Métodos*

As amostras da fase óleo foram cedidas pelo LABBIO e fizeram parte do projeto de pesquisa de Azambuja e Colaboradores. O diesel foi recebido da empresa Ipiranga Distribuidores de Combustíveis, Canoas, RS.<sup>12</sup>

#### *4.1.2.1 Descrição das amostras de diesel*

Diesel com teor de enxofre ultra alto - UHSD  $\leq 1800$  mg L<sup>-1</sup> de enxofre (nível de enxofre na amostra empregada: 1366 mg L<sup>-1</sup> S); diesel com alto teor de enxofre - HSD  $\leq 500$  mg L<sup>-1</sup> de enxofre (nível de enxofre na amostra empregada: 408 mg L<sup>-1</sup> S); diesel com baixo teor de enxofre - LSD  $\leq 50$  mg L<sup>-1</sup> de enxofre (nível de enxofre na amostra empregada: 7,1 mg L<sup>-1</sup> S) e diesel de ultra baixo teor de enxofre - ULSD  $\leq 10$  mg L<sup>-1</sup> de enxofre (nível de enxofre na amostra empregada: 4,6 mg L<sup>-1</sup> S).

#### *4.1.2.2 Descrição do experimento de estocagem simulada*

Foram simuladas condições de estocagem em que as amostras de diesel foram contaminadas com uma mistura de microrganismos não identificados. Foram utilizados frascos de vidro de 100 mL, com tampa de rosca de plástico. Um volume de 40 mL de diesel puro com diferentes teores de enxofre, foi adicionado a 10 mL de um meio mineral Bushnell-Haas estéril para simular a fase aquosa. Os frascos foram incubados a 30°C no escuro. Foram analisadas amostras da fase óleo nos tempos zero, 10 dias, 20 dias, 30 dias e 40 dias através espectroscopia no infravermelho.

Cada tempo amostral foi feito em triplicata, e para cada tempo havia amostras controle (amostras esterilizadas e sem a adição de microrganismos), e amostras em dois níveis de contaminação microbiana: baixo nível foram amostras de diesel recebidas do distribuidor

que não foram esterilizadas e amostras com nível médio de contaminação, que foram produzidas através de um processo de esterilização e foi adicionado uma mistura de microrganismos não caracterizados<sup>12</sup>, as amostras foram divididas em grupos como descrito na Tabela 2.

**Tabela 2:** Descrição da nomenclatura das amostras

Amostras	Descrição
<b>Tempo zero (T0)</b>	Amostras da fase óleo (diesel) no início do experimento.
<b>Tempo 10 - Controle (T10C)</b> <b>Tempo 20 - Controle (T20C)</b> <b>Tempo 30 - Controle (T30C)</b> <b>Tempo 40 - Controle (T40C)</b>	Amostras estéreis.
<b>Tempo 10 - Inóculo (T10I)</b> <b>Tempo 20 - Inóculo (T20I)</b> <b>Tempo 30 - Inóculo (T30I)</b> <b>Tempo 40 - Inóculo (T40I)</b>	Amostras com médio nível de contaminação.
<b>Tempo 10 - Recebido (T10R)</b> <b>Tempo 20 - Recebido (T20R)</b> <b>Tempo 30 - Recebido (T30R)</b> <b>Tempo 40 - Recebido (T40R)</b>	Amostras recebidas do distribuidor com baixo nível de contaminação.

#### 4.1.2.3 Espectroscopia no Infravermelho

Foram analisadas por Espectroscopia no Infravermelho com Transformada de Fourier acoplada com refletância total atenuada (FTIR- ATR) as amostras da fase óleo nos tempos zero, 10 dias, 20 dias, 30 dias e 40 dias. Os espectros foram obtidos em um equipamento FTIR Spectrum 400, Perkin Elmer - acoplado num acessório de refletância total atenuada horizontal (HATR) com cristal de seleneto de zinco (ZnSe)). Todas as amostras foram obtidas em duplicata e na faixa de 4.000 a 650  $\text{cm}^{-1}$ , a temperatura ambiente, com resolução de 4  $\text{cm}^{-1}$  e 16 varreduras.

#### 4.1.2.4 Análise Multivariada dos dados

Neste estudo foram avaliados os métodos de pré-processamentos geralmente aplicados para dados de FTIR-ATR com amostras de combustíveis<sup>76,125,144-147</sup> e o software utilizado foi o ChemoStat<sup>®148</sup>:

- 1) Normalização dos espectros para um máximo de 1;
- 2) Aplicação de primeira e segunda derivada usando o algoritmo de Savitzky-Golay (polinômio de segunda ordem e 15 pontos por janela);

### 3) Centrar na média.

A aplicação do pré-processamento aos espectros brutos foi realizada de acordo com seguinte ordem:

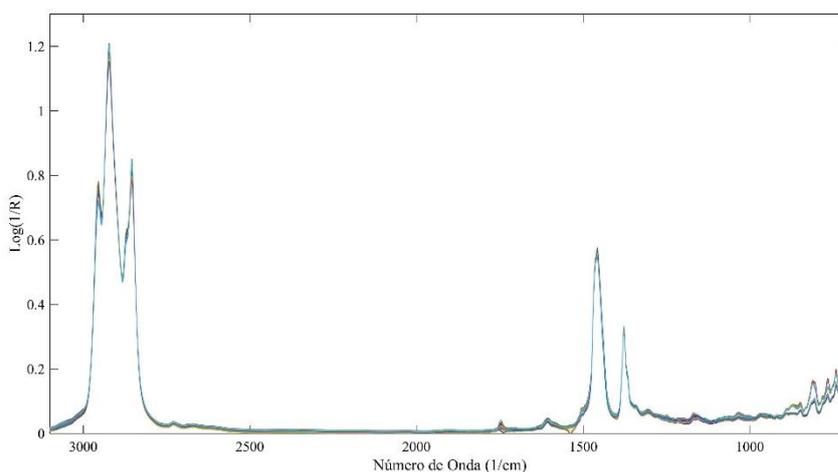
- centrar na média;
- normalização + centrar na média;
- normalização + 1º Derivada + centrar na média;
- normalização + 2º Derivada + centrar na média.

Os métodos de pré-processamento foram avaliados através de análise multivariada utilizando Análise por Componentes Principais (PCA) e Análise por Componentes Principais por intervalo (*i*PCA) para seleção de variáveis. O pré-processamentos e as combinações de métodos foram aplicados aos espectros com o intuito de avaliar quais apresentavam a melhor discriminação entre as amostras.

## 4.2 Resultados e Discussão

### 4.2.1 Caracterização do Diesel por FTIR

Os espectros de infravermelho médio nas regiões de  $3.100 - 700 \text{ cm}^{-1}$  e  $1.100 - 700 \text{ cm}^{-1}$  das amostras de diesel com diferentes teores de enxofre, estão apresentados nas Figuras 6 e 7.



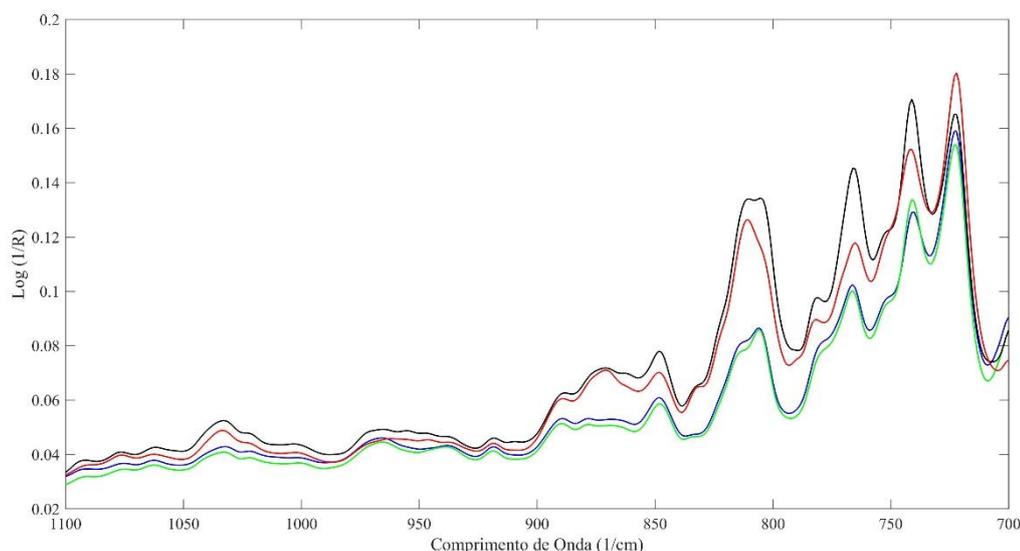
**Figura 6:** Espectros de FTIR- ATR das amostras de diesel com diferentes teores de enxofre.

Os espectros apresentam modos e combinações vibracionais referentes à composição química do diesel, as bandas de absorção observadas na região de  $3.050 - 2.800 \text{ cm}^{-1}$  podem ser atribuídas às vibrações de deformação axial das ligações C-H dos grupamentos metila ( $\text{CH}_3$ ), metileno ( $\text{CH}_2$ ) e das ligações duplas ( $\text{CH}=\text{CH}$ ), as bandas de absorção na região de  $1.500 - 1.350 \text{ cm}^{-1}$  correspondem a deformação angular de  $\text{CH}_2$  e  $\text{CH}_3$ , e na região de  $1.000 - 900 \text{ cm}^{-1}$  as bandas de absorção podem ser atribuída a deformação angular simétrica fora do plano de ligações C-H de olefinas e na região de  $900 - 700 \text{ cm}^{-1}$  a deformação angular assimétrica de  $\text{CH}_2$  e ligações C- H fora do plano de compostos aromáticos.<sup>149</sup>

Na Figura 7, observa-se na região de  $1.100 - 700 \text{ cm}^{-1}$  a diferença dos espectros do diesel de acordo com os teores de enxofre. O perfil dos espectros é semelhante, porém a intensidade do sinal é maior no diesel S500 e S1800.

A diferença dos espectros não se deve a absorção de enxofre, que está presente no diesel como mercaptanos e sulfetos porém em ambos os casos as bandas de absorção no espectro FTIR são muito fracas, e em consequência, praticamente não são percebidas visualmente.<sup>35</sup> Em geral, vibrações C-S, S-H e/ou S-S específicas não são muito

detectáveis por IR devido à pequena alteração do momento de dipolo durante as vibrações desses elementos estruturais.<sup>149</sup>



**Figura 7:** Espectros de FTIR- ATR das amostras de diesel com diferentes teores de enxofre, azul (10 ppm), verde (50 ppm), preto (500 ppm) e vermelho (1800 ppm).

O espectro de infravermelho é diretamente proporcional a quantidade dos componentes químicos presentes na amostra, os picos observados na região de 900 a 700  $\text{cm}^{-1}$  podem atribuídos a ligações C-H fora do plano de compostos aromáticos, ligações de  $\text{CH}_2$  de compostos com número de carbonos acima de 3 e ligações  $=\text{CH}$ <sup>149</sup>, logo diesel S500 e S1800 são mais ricos nesses compostos.

#### 4.2.2 *Otimização do Pré-processamento e Análise Exploratória dos dados de FTIR – ATR de Diesel com diferentes teores de enxofre submetidos a biodegradação*

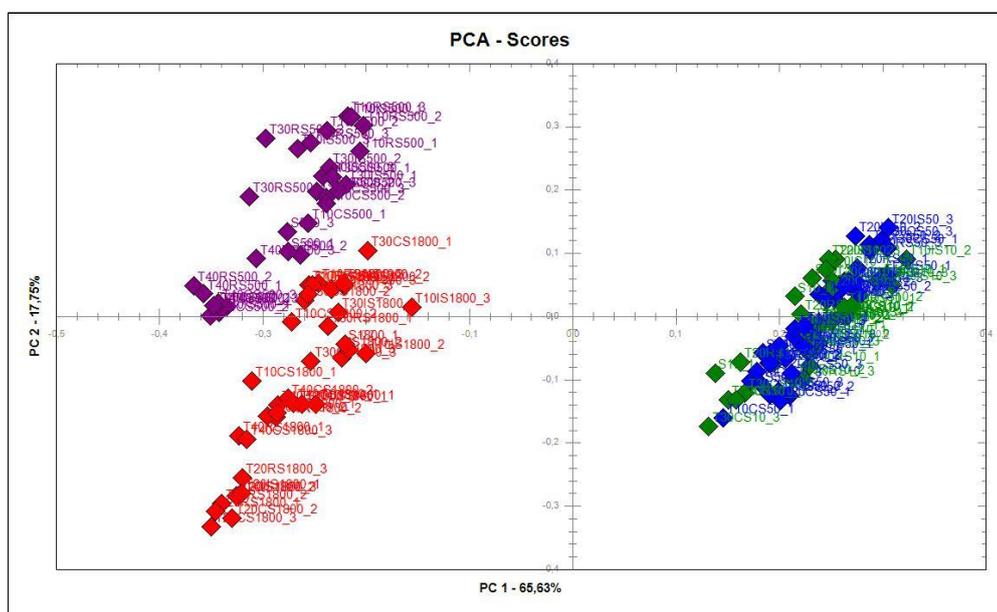
Neste estudo foram analisados quatro tipos de pré-processamentos e suas combinações: centrar na média, normalização, primeira e segunda derivadas e a análise exploratória dos dados foi feita através da Análise por Componentes Principais (PCA).

O objetivo da análise multivariada das amostras foi avaliar alterações que podem ter ocorrido nas amostras de diesel contaminadas com microrganismos após o período de estocagem simulada, verificar quais e se ocorreu alterações nos espectros entre amostras controle (T10C, T20C, T30C e T40C), amostras com nível médio de contaminação (T10I, T20I, T30I e T40I) e amostras com baixo nível de contaminação (T10R, T20R, T30R e T40R) no período de 10, 20, 30 e 40 dias.

A primeira etapa foi analisar todas as amostras sem separar pelo teor de enxofre para verificar como se comportavam. A primeira avaliação foi pelo tempo amostral, de acordo com o tempo de estocagem simulada, as diferentes combinações de pré-processamento foram aplicadas e analisadas através de uma análise exploratória com PCA. No entanto, o efeito do pré-processamento também foi estudado para diferenciar as amostras controle e contaminadas.

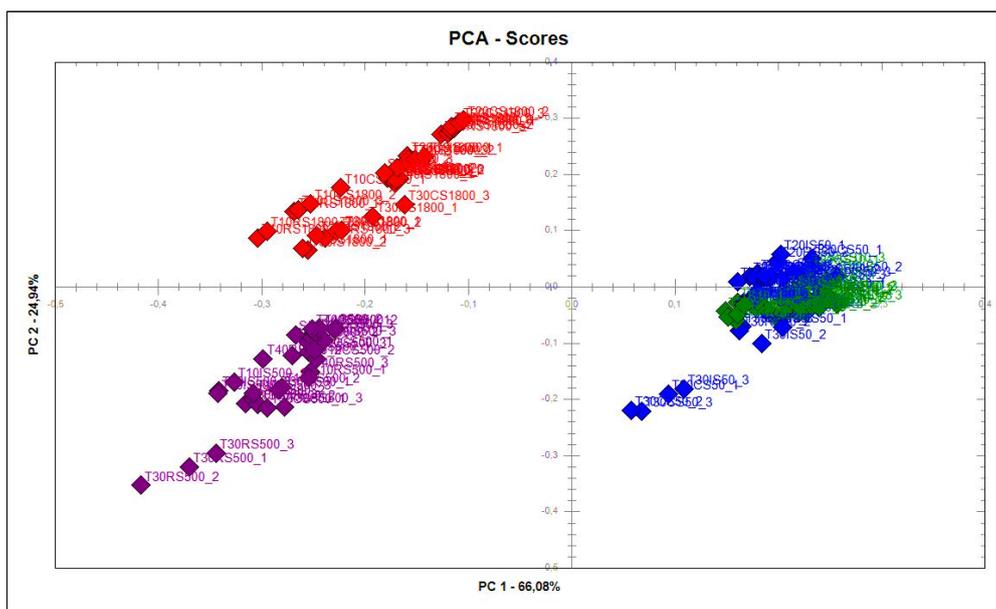
Nas Figuras 8, 9, 10 e 11 estão apresentados os gráficos de escores da PCA dos espectros de diesel submetidos as seguintes pré-processamento e combinações: centrados na média; normalização e centrados na média; normalização, primeira derivada e centrados na média; normalização, segunda derivada e centrados na média.

Na Figura 8, está apresentado o gráfico de escores da PC1xPC2 dos espectros do diesel que foram centrados na média, representando 83,38% da variância total dos dados e na Figura 9, é apresentado o gráfico de escore da PC1xPC2 dos espectros do diesel que foram normalizados (entre 0-1) e centrados na média, representando 91,02% da variância total dos dados. Há formação evidente de três grandes grupos, diesel S500 (roxo), diesel S1800 (vermelho) e uma mistura entre diesel S50 e S10 (azul e verde), nos dois casos, porém após a normalização dos espectros houve uma maior distinção das amostras de S500 e S1800 e o agrupamento entre as amostras ficou mais evidente.



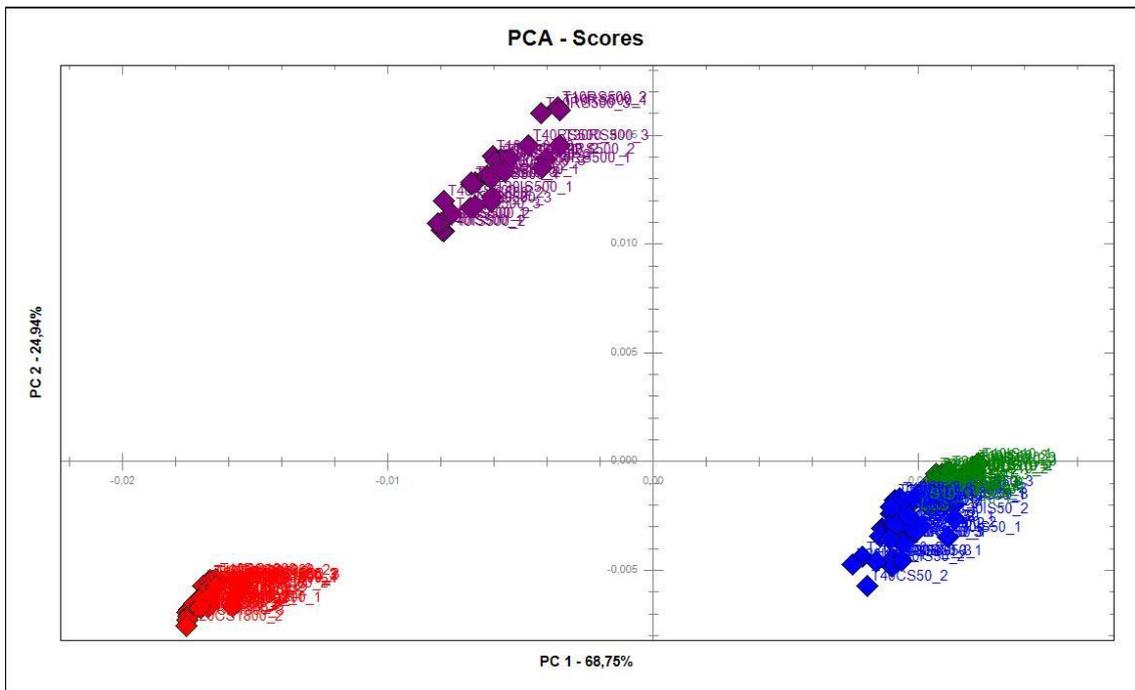
**Figura 8:** Gráfico de escores da PCA dos espectros de FTIR- ATR centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800 (vermelho).

Analisando o gráfico de escores da PC1xPC2, Figura 9, as amostras de S10 e S50 separaram-se das amostras de S500 e S1800. A separação é identificada no lado positivo da PC1 com 66,08% variância total. Já na PC2 houve a separação das amostras de S500 e S1800 com 24,94% da variância total, S500 com escores positivos e S1800 com escores negativos, e houve a separação de dois grupos de amostras do tempo 30 (T30RS500 e T30CS50), ambas com escores negativos na PC2, essa separação não ocorre nas análises seguintes.

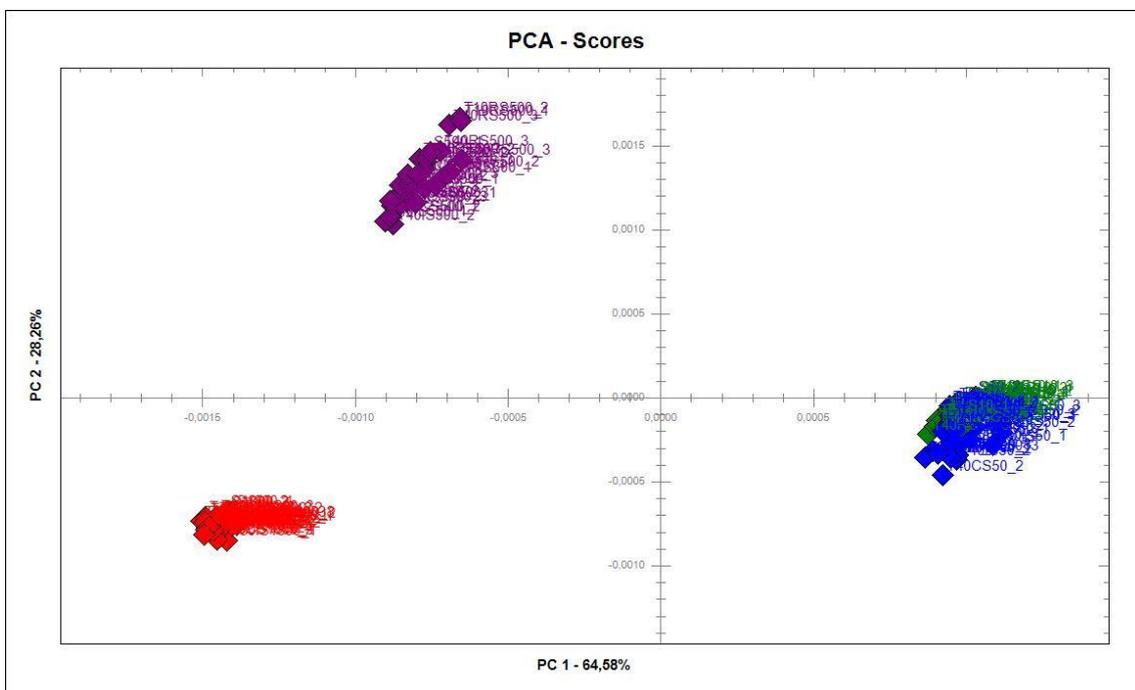


**Figura 9:** Gráfico de escores da PCA dos espectros de FTIR- ATR normalizados e centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800 (vermelho).

Na Figura 10, está apresentado o gráfico de escore da PC1xPC2 dos espectros do diesel com a seguinte combinação de pré-processamento: normalização, primeira derivada e centrados na média, representando 93,7% da variância total dos dados e na Figura 11, é apresentado o gráfico de escore da PC1xPC2 dos espectros do diesel que foram normalizados, segunda derivada e centrados na média, representando 92,84% da variância total dos dados. A aplicação da segunda derivada alterou pouco a variância total da PC1xPC2 em relação a aplicação da primeira derivada, ficando evidente que a PC2 é responsável pela separação das amostras de diesel S500 com escores positivos das amostras de diesel S1800 com escores negativos. Após a aplicação da primeira derivada houve uma separação das amostras S10 e S50 na PC2 como poder ser observado na Figura 10.

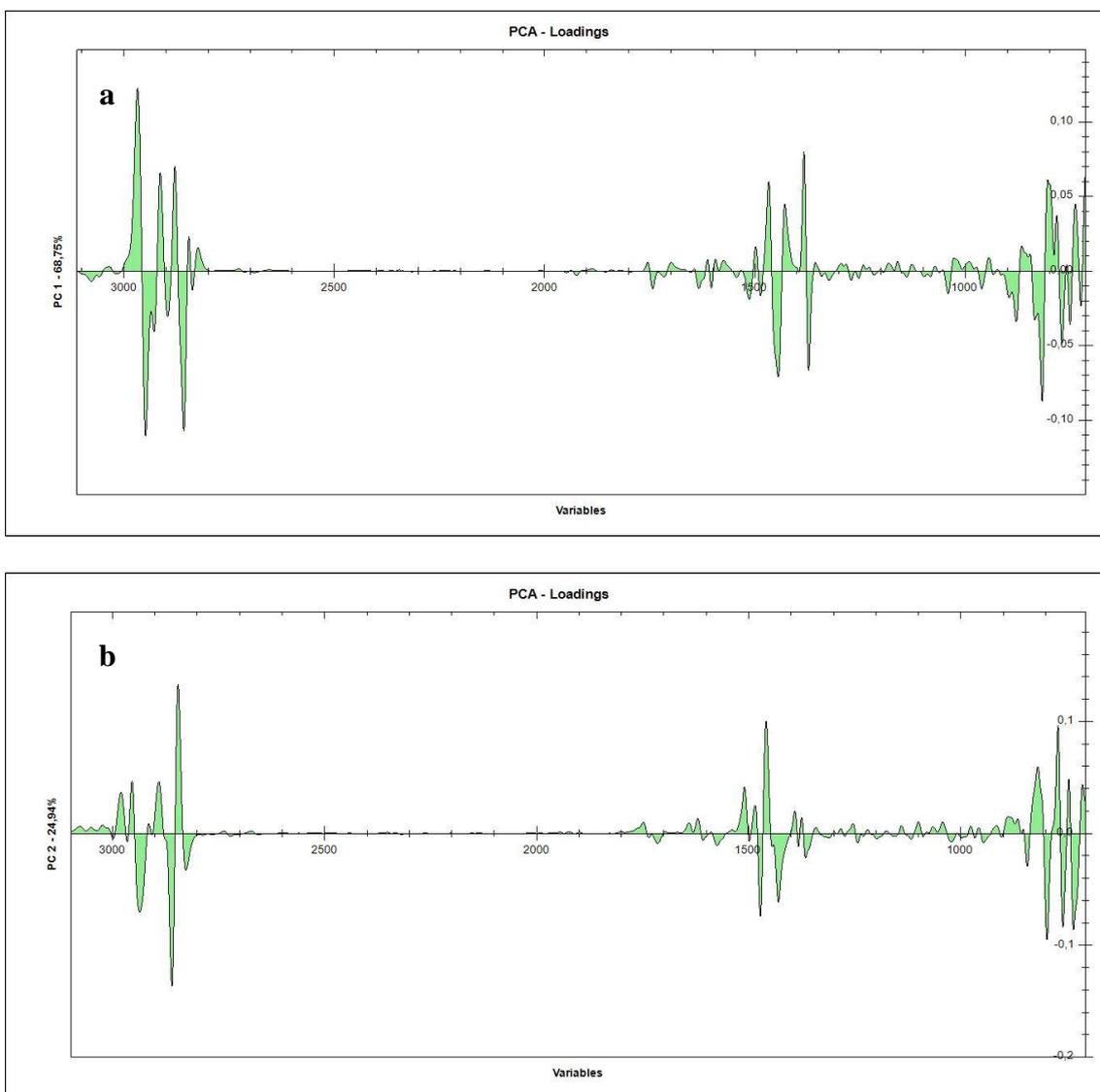


**Figura 10:** Gráfico de escores PC1xPC2 dos espectros de FTIR- ATR normalizados, com primeira derivada e centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800 (vermelho).



**Figura 11:** Gráfico de escores da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel com diferentes teores de enxofre. Diesel S10 (verde), Diesel S50 (azul), Diesel S500 (roxo), Diesel S1800 (vermelho).

Para verificar quais regiões do espectro contribuíram para a separação das amostras foi analisado o gráfico de pesos da PC1 e PC2, Figuras 12a e 12b, das amostras que foram submetidas a combinação de pré-processamento normalização, primeira derivada e centrados na média. Os pesos da PCA elucidam a correlação entre as variáveis e as amostras.



**Figura 12:** Gráfico de pesos da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel com diferentes teores de enxofre: a) PC1 e b) PC2

A maior carga da PC1 e PC2 está relacionada às regiões mais importantes do espectro do diesel (Figura 6). A região de  $1500 - 700 \text{ cm}^{-1}$ , têm grande influência na separação das amostras de S10 e S50 das amostras de S500 e S1800 que são mais ricas nos compostos

que absorvem nessa região e a separação das amostras S500 e S1800 é influenciada pela região de 3000 -2800  $\text{cm}^{-1}$ .

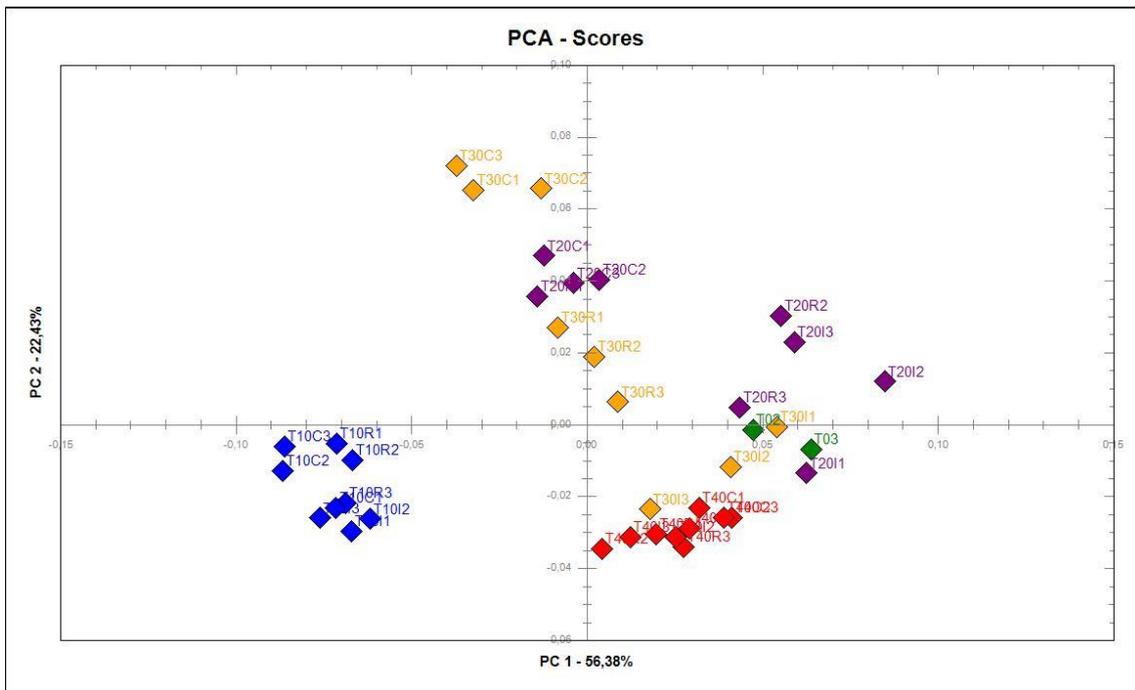
A separação das amostras pelo tempo de estocagem simulada não ocorreu, mas observa-se que houve a formação de grupos distintos relacionados com o teor de enxofre. A aplicação da primeira e segunda derivada nos dados removeu os sinais da linha de base e enfatizou as diferenças entre as amostras como pode ser observado nas Figuras 10 e 11, em comparação com as amostras que foram apenas centrados na média, Figura 8 e centrados na média + normalização, Figura 9, porém a informação sobre o teor de enxofre prevaleceu sobre possíveis alterações que tenham ocorrido nos espectros das amostras que foram submetidas a biodegradação e a estratégia foi avaliar as amostras separadamente pelo teor de enxofre.

#### 4.2.3 *Análise Multivariada dos Espectros de Diesel S10*

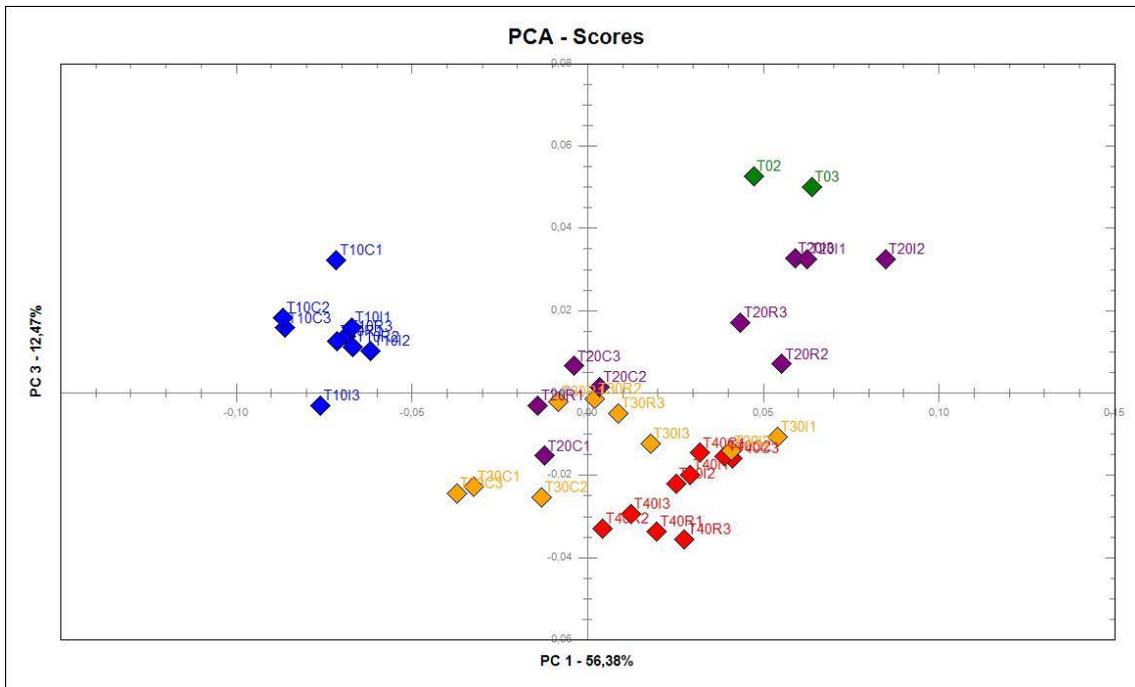
O mesmo procedimento com os pré-processamento e análise exploratória foi aplicado nos dados de diesel S10. O método PCA foi aplicado aos espectros previamente tratados (normalizados, derivados em primeira e segunda ordem e centrados na média).

As Figuras 13, 14, 15, 16, 17 e 18 mostram os escore para as três componentes principais (PCs) da PCA de todos os espectros de amostras de S10 com diferentes tempos amostrais. A correta separação das amostras de acordo com o tempo de estocagem simulada, não ficou clara, mas uma tendência de separação pode ser observada após a análise multivariada dos dados.

Nas Figuras 13 e 14 são apresentados os gráficos de escores da PC1xPC2 e PC1xPC3 para os dados normalizados e centrados na média, considerou-se a análise das três primeiras componentes principais, que resultaram em uma variância acumulada de 91,28%. Analisando a primeira componente (PC1), Figura 13, houve uma separação das amostras do tempo 10 das amostras dos tempos 0, 20, 30 e 40. A separação é identificada no lado negativo do PC1, com 56,38% da variância total. A segunda componente (PC2), com 22,43% da variância, distingue as amostras dos tempos 10 e 40 (T10 e T40) das amostras controle dos tempos 20 e 30 (T20C e T30C). E na terceira componente principal (PC3), com 12,47% da variância, distinguiu as amostras do tempo zero (T0), Figura 14.



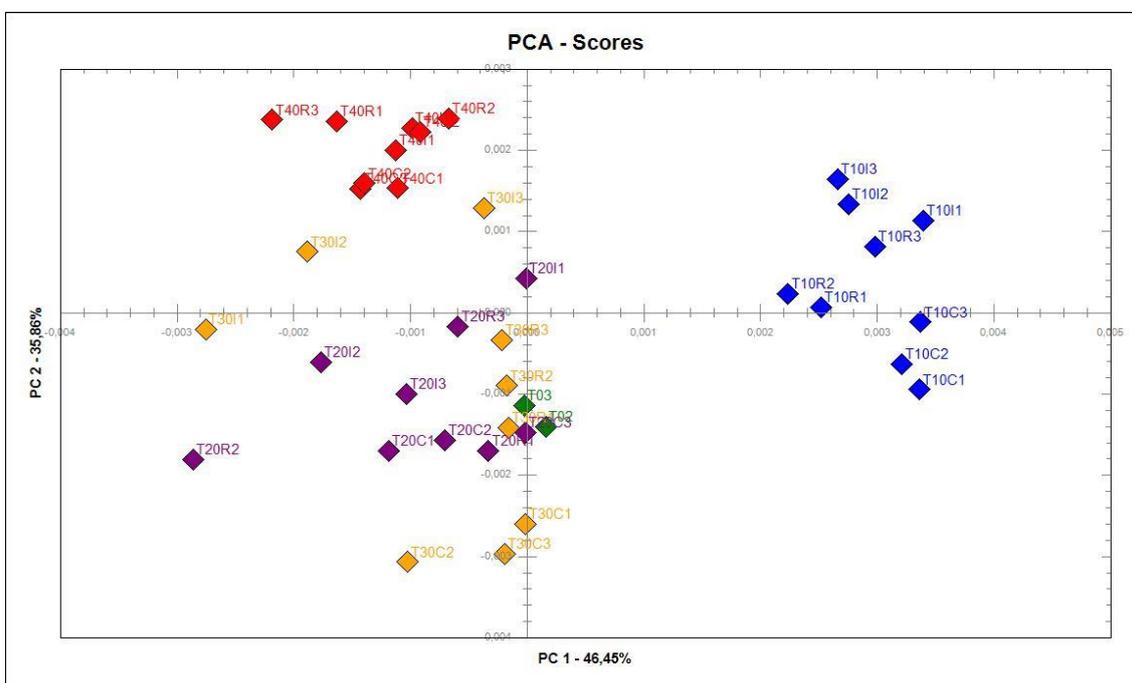
**Figura 13:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-ATR normalizados e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).



**Figura 14:** Gráfico de escores PC1xPC3 da PCA dos espectros de FTIR-ATR normalizados e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

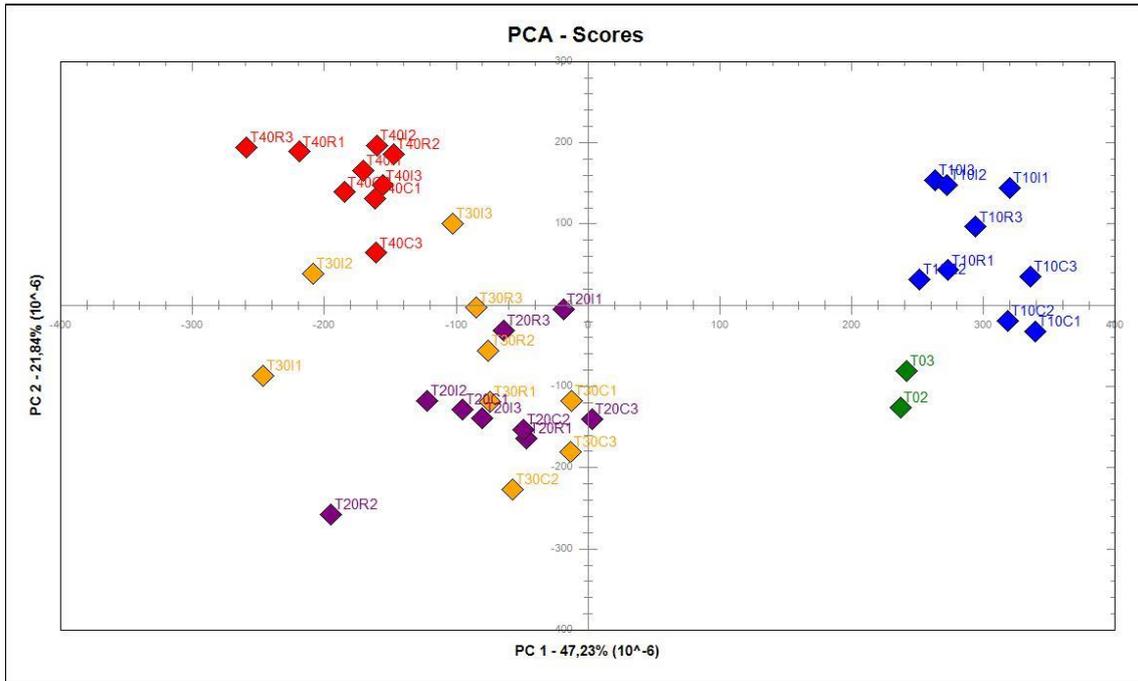
Nas Figuras 15 e 16 são apresentados os gráficos de escores da PC1xPC2 e PC1xPC3 das amostras após normalização, aplicação da primeira derivada e centrados na média, considerou-se a análise das três primeiras componentes principais, que resultaram em uma variância acumulada de 87,73% e nas Figuras 17 e 18 são apresentados os gráficos de escores da PC1xPC2 e PC1xPC3 das amostras após normalização, aplicação da segunda derivada e centrados na média, que resultaram em uma variância acumulada de 79,61%.

Analisando o gráfico de escores, Figura 15, a primeira componente (PC1) com 46,45% da variância, separa as amostras do tempo do tempo 10 (T10) das amostras dos tempos 20, 30 e 40 (T20, T30 e T40) e as amostras do tempo zero (T0) ficaram próximas das amostras dos tempos 20 e 30 (T20 e T30). A segunda componente (PC2), com 35,86% da variância, distingui as amostras dos tempos 10 e 40 (T10 e T40) e também pode ser observada a separação das amostras controle do tempo 30 (T30C), no lado negativo da PC2. E na terceira componente principal (PC3), com 5,42% da variância, houve a separação de duas amostras dos tempos 20 e 30 (T20 e T30), Figura 16.

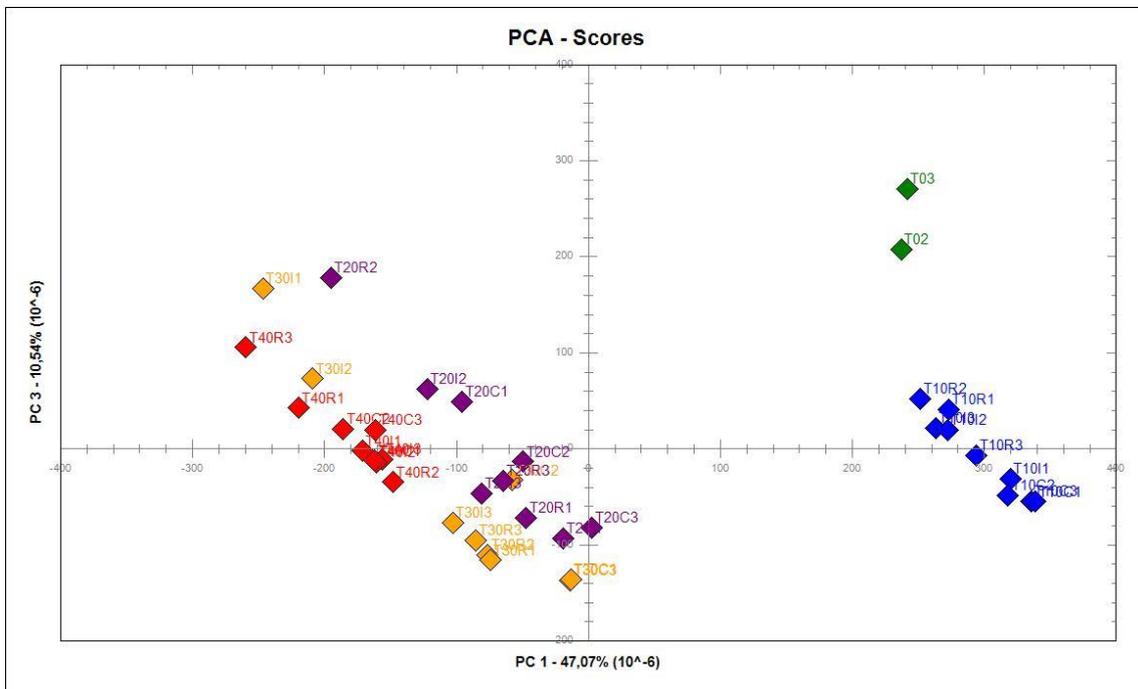


**Figura 15:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR- ATR normalizados, com primeira derivada e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).





**Figura 17:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).



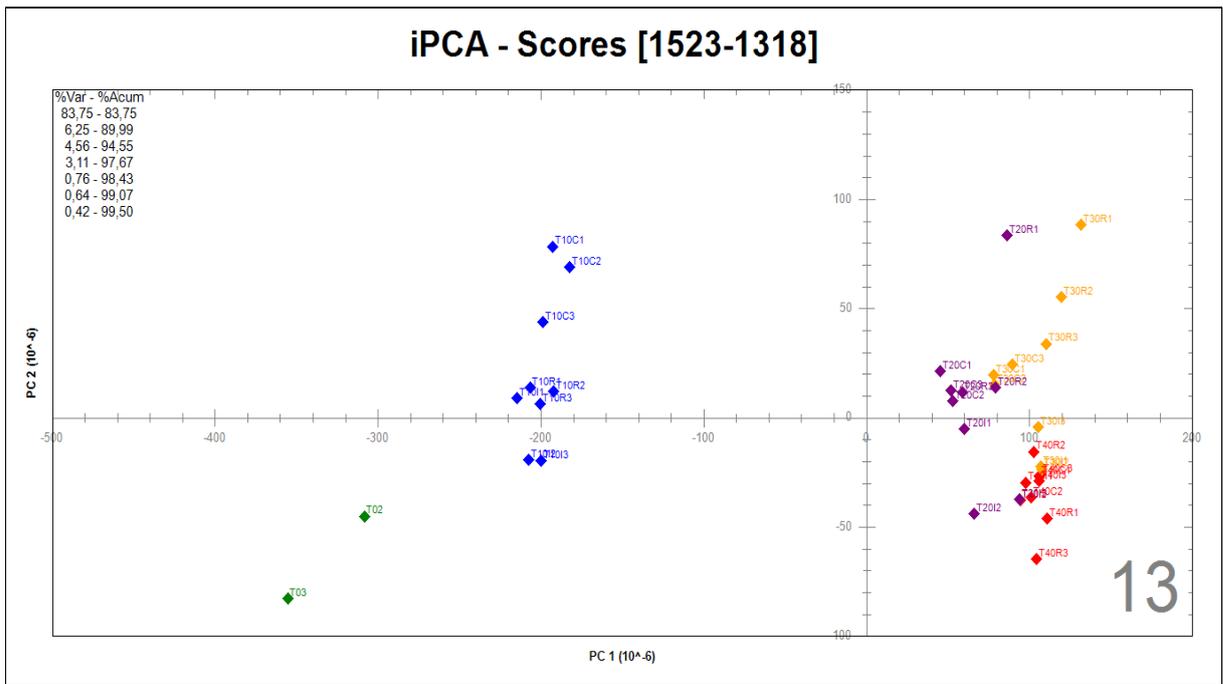
**Figura 18:** Gráfico de escores PC1xPC3 da PCA dos espectros de FTIR- ATR normalizados, com segunda derivada e centrados na média das amostras de diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

Em nenhuma das combinações de pré-processamento houve uma separação evidente dos tempos 20 e 30 dos demais tempos amostrais. E em nenhum dos casos houve separação das amostras que foram contaminadas com o microrganismo (T10I, T20I, T30I e T40I) das amostras controle (T10C, T20C, T30C e T40C) e amostras recebidas (T10R, T20R, T30R e T40R). Porém, ficou evidente que a aplicação de combinações de pré-processamento diferentes melhorou a interpretação dos resultados, e a utilização de derivadas resultou em uma melhor separação das amostras.

A próxima etapa foi investigar as regiões espectrais separadamente através da análise por componentes principais por intervalo (*i*PCA), com a melhor combinação de pré-processamento determinada anteriormente.

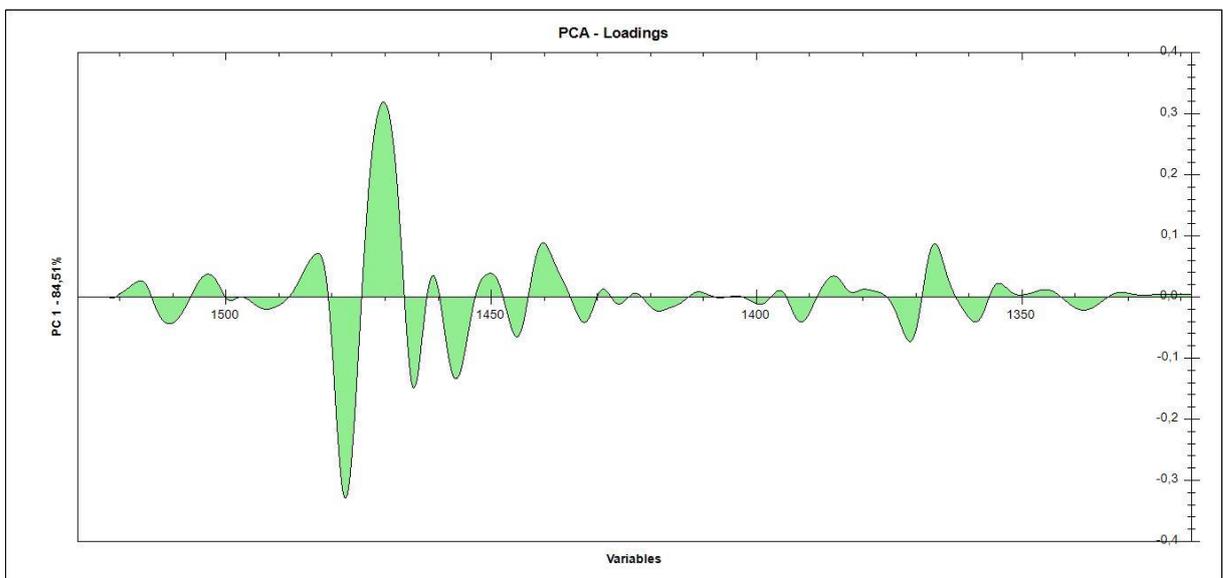
A *i*PCA é similar à PCA, porém antes de realizar as operações matriciais, a *i*PCA divide o espectro em *i* intervalos equidistantes. Ou seja, para  $i = 16$ , o espectro é dividido em dezesseis intervalos, e a PCA é, então, conduzida para cada intervalo separadamente. No presente trabalho, os valores de *i* foram 4, 8, 16 e 32, e cada PCA foi avaliada separadamente.

A melhor separação das amostras de acordo com o tempo amostral foi encontrada na divisão em 16 intervalos (*i*16PCA), na Figura 19, está apresentado o gráfico de escores da PC1 e PC2 referente ao intervalo 13, com variância total de 93,37%. É possível observar que houve uma tendência na separação das amostras por tempo amostral na PC1, as amostras do tempo 0 e tempo 10 com escores negativos e tempos 20, 30 e 40 escores positivos. A região de  $1.523 - 1318 \text{ cm}^{-1}$  corresponde às bandas de absorção de deformação angular de grupos metilas ( $\text{CH}_3$ ) e metilênicos ( $\text{CH}_2$ ).



**Figura 19:** Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318  $\text{cm}^{-1}$  para o diesel S10: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

De acordo com a análise do gráfico de pesos, Figura 20, amostras dos tempos zero e 10 são mais ricas em grupos metilas ( $\text{CH}_3$ ) e amostras dos tempos 20, 30 e 40 em grupos metilênicos  $\text{CH}_2$ .

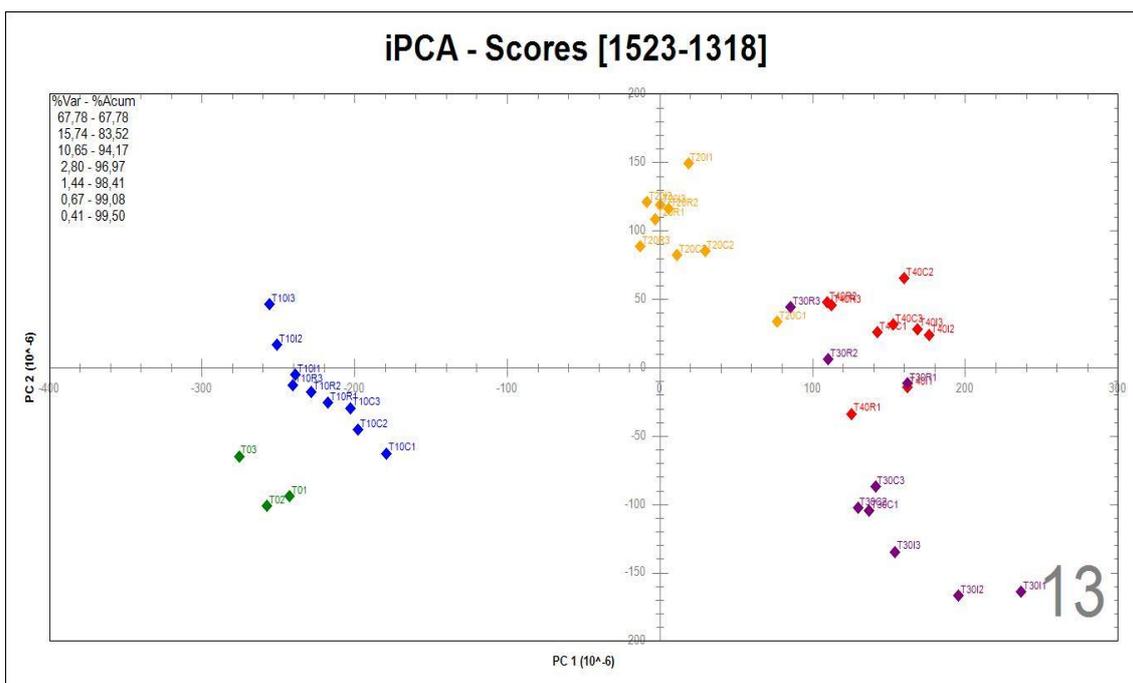


**Figura 20:** Gráfico de pesos da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318  $\text{cm}^{-1}$  para o diesel S10.

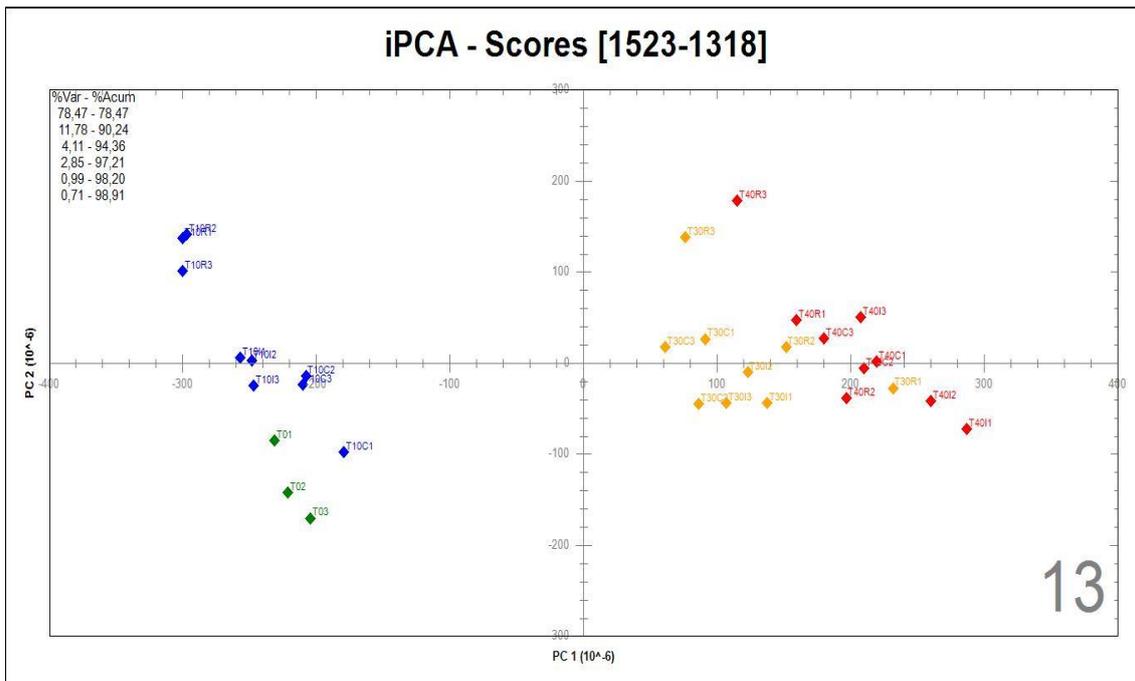
#### 4.2.4 Análise Multivariada dos Espectros de Diesel S50, S500 e S1800

O mesmo procedimento com tratamento prévio através de normalização, primeira e segunda derivada, centrados na média utilizado com os espectros de S10 foi aplicado nos espectros de S50, S500 e S1800, e os resultados obtidos foram semelhantes. Assim, os gráficos escores da PC1 e PC2 estão apresentados nos anexos, espectros de diesel S50, Figuras 1, 2 e 3, espectros de diesel S500, Figuras 4, 5 e 6, e espectros de diesel S1800, Figuras 7, 8, 9.

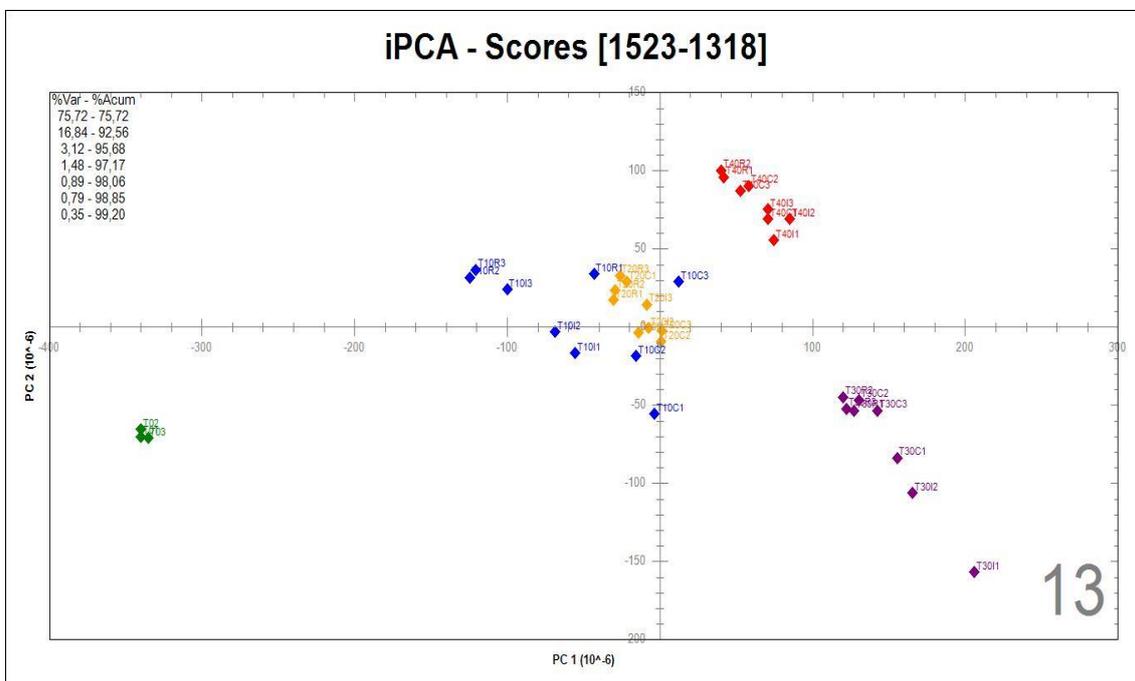
Nas Figuras 21, 22 e 23 são apresentados os gráficos de escore da PC1 e PC2 referente ao intervalo 13 da *i*PCA das amostras de S50, S500 e S1800, com a combinação de pré-processamento de normalização, segunda derivada e centrado na média. É possível observar em todos os casos, que a região de melhor discriminação das amostras foi 1523 - 1318  $\text{cm}^{-1}$ , que separou as amostras do tempo 0 e tempo 10 com escores negativos dos tempos 20, 30 e 40 com escores positivos na PC1.



**Figura 21:** Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318  $\text{cm}^{-1}$  para o diesel S50: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).



**Figura 22:** Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318 cm<sup>-1</sup> para o diesel S500: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).



**Figura 23:** Gráfico de escores da PCA dos espectros de FTIR- ATR do intervalo 13, região de 1523 - 1318 cm<sup>-1</sup> para o diesel S1800: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

### **4.3 Conclusão**

O presente estudo analisou diferentes combinações de pré-processamento para evidenciar possíveis alterações nos espectros de FTIR-ATR de amostras de diesel que foram submetidas à estocagem simulada com a contaminação por microrganismo e ficou evidente a necessidade de avaliação prévia dos tipos de pré-processamento a serem aplicados nos espectros para a melhor interpretação dos resultados.

Foi demonstrado que o FTIR-ATR pode ser empregado para avaliação a degradação de diesel, no entanto, a região de  $1528 - 1318 \text{ cm}^{-1}$  é a mais indicada e uma possível combinação de pré-processamento com normalização/segunda derivada/centrado na média foi apresentada.

#### **4.4 Quantificação da Biomassa Produzida na Biodegradação de Blendas B10 durante estocagem simulada usando Infravermelho com Transformada de Fourier (FTIR) e Regressão por Mínimos Quadrados Parciais (PLS)**

##### *4.4.1 Objetivo*

Quantificar a degradação das blendas B10 (10% de biodiesel puro e 90% de óleo diesel) durante estocagem simulada com microrganismos (Inóculo não caracterizado – Norma ASTM E 1259) através de Espectroscopia no Infravermelho Médio (FTIR) e Regressão por Mínimos Quadrados Parciais (PLS). Este trabalho foi desenvolvido em parceria com o Laboratório de Biodeterioração de Combustíveis e Biocombustíveis (LABBIO), UFRGS /RS.

##### *4.4.2 Materiais e Métodos*

###### *4.4.2.1 Amostras*

As amostras da fase óleo de um experimento que fizeram parte do trabalho de pesquisa de Azambuja e colaboradores foram cedidas pelo LABBIO.<sup>150</sup>

As blendas B10 foram preparadas em laboratório pela mistura de biodiesel e óleo diesel com diferentes teores de enxofre: Diesel com teor de enxofre ultra alto - UHSD  $\leq$  1800 mg/L de enxofre; diesel com alto teor de enxofre - HSD  $\leq$  500 mg/L de enxofre; diesel com baixo teor de enxofre - LSD  $\leq$  50 mg/L de enxofre e diesel de ultrabaixo teor de enxofre - ULSD  $\leq$  10 mg/L de enxofre. O combustível foi fornecido pela empresa Ipiranga Distribuidores de Combustíveis, Canoas, RS

Para avaliar a biodegradação foram simuladas condições de estocagem em que as amostras de B10 foram contaminadas com microrganismos não identificados. Foram utilizados frascos de vidro de 100 mL, com tampa de rosca de plástico. Um volume de 20 mL da blenda B10, foi adicionado a 10 mL de um meio mineral Bushnell-Haas estéril para simular a fase aquosa. Os frascos foram incubados a 30°C no escuro. Foram analisadas amostras da fase óleo nos tempos zero, 10 dias, 20 dias, 30 dias e 40 dias e para cada período havia amostras controle (que não foram contaminadas) e amostras que foram contaminadas com microrganismo.

###### *4.4.2.2 Quantificação da Biomassa*

A quantificação da biomassa produzida pela biodegradação foi determinada por gravimetria. Após cada tempo amostral, a biomassa formada na interface óleo/água foi

removida por filtração, utilizando membranas de papel filtro que foram previamente submetidas a secagem em estufa bacteriológica a 30°C por 48 horas e mantidas em dessecador por 24 horas e na sequência pesadas em balança de precisão.

A biomassa foi recuperada por filtração adicionando-se 4 mL de hexano para remoção dos resíduos oleosos e após foi realizada uma etapa de secagem e pesagem das membranas com a biomassa.<sup>150</sup>

A quantificação da biomassa foi expressa em mg mL<sup>-1</sup> de peso seco e obtida pela diferença entre o peso final e peso inicial da membrana. Os resultados para 98 amostras são apresentados nos anexos (Tabela 1), os valores de biomassa variam de 0 a 12,2 mg mL<sup>-1</sup>.

#### 4.4.2.3 Espectroscopia no Infravermelho

Os espectros FTIR – ATR da fase óleo foram obtidos empregando um espectrofotômetro (Perkin Elmer modelo Spectrum 400), acoplado a um acessório de refletância total atenuada (ATR) com cristal de seleneto de zinco (ZnSe). Os espectros foram obtidos em duplicata para cada amostra na faixa de 4000 – 650 cm<sup>-1</sup>, a temperatura ambiente, com resolução de 2 cm<sup>-1</sup> e 32 varreduras.

#### 4.4.2.4 Análise Multivariada dos Dados

A análise multivariada das amostras de B10 foi realizada a partir dos dados FTIR-ATR e biomassa (peso seco) e o programa utilizado para estes estudos foi o Matlab 7,11 (MathWorks Inc, Natick, MA), pacote iToolbox e computador com processador Intel Core i5 2,40 GHz.

Foram construídos modelos de regressão multivariada por PLS, *i*PLS e *si*PLS para quantificação de biomassa utilizando dados de FTIR-ATR e dados de referência obtidos por gravimetria de 98 amostras. Para a análise multivariada foram utilizados os espectros médios e dois pré-processamentos foram aplicados aos espectros: normalização (0,1) e centrar na média.<sup>125</sup>

O método de regressão por mínimos quadrados parciais por intervalo (*i*PLS) é uma extensão do PLS, que consiste na divisão da região espectral em intervalos equidistantes e para cada intervalo um novo modelo PLS é construído a partir das variáveis selecionadas. O principal objetivo é fornecer informações relevantes em diferentes subdivisões do espectro total, removendo as regiões espectrais cujas variáveis não apresentam informações relevantes e/ou interferentes.<sup>151</sup>

A regressão por quadrados mínimos por sinergismo de intervalos *si*PLS é um método adicional ao *i*PLS, que consiste na combinação dos intervalos resultantes da divisão de regiões do espectro. Um novo modelo é construído com cada combinação.

Para construção dos modelos de regressão multivariada as 98 amostras foram divididas em conjuntos de calibração e previsão. O número de amostras selecionadas para o conjunto de calibração foi 68 (cerca de 2/3 do total de amostras) e para o conjunto de previsão 30 amostras, a seleção das amostras foi realizada utilizando algoritmo de Kennard-Stone.<sup>152</sup> A Tabela 3 apresenta a descrição dos conjuntos de calibração e previsão, a matriz **X** são variáveis independentes (valores de absorbância medidos na faixa de 4000 – 650 cm<sup>-1</sup>) e a matriz **Y** as variáveis dependentes (valores de biomassa).

**Tabela 3:** Descrição dos conjuntos de calibração e previsão.

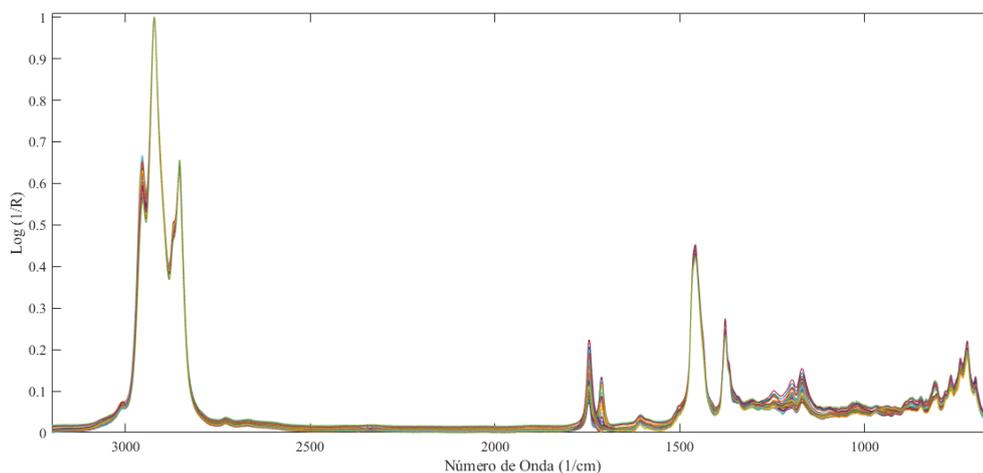
<b>Conjunto</b>	<b>N° de amostras</b>	<b>Matriz</b>	<b>Dimensão</b>	<b>Matriz</b>	<b>Dimensão</b>
<b>Calibração</b>	68	X	68x3500	Y	68x1
<b>Previsão</b>	30	X	30x3500	Y	30x1

Na construção do modelo PLS foi utilizado toda a faixa espectral (4000 – 650 cm<sup>-1</sup>). Para o *i*PLS o espectro foi dividido 2, 4, 8, 16 e 32 intervalos e para o *si*PLS a combinação de 2, 3 e 4 intervalos foi usado quando a região espectral foi dividida em 8, 16 e 32 intervalos. A seleção do melhor modelo foi baseada na análise do número de variáveis latentes, raiz quadrada do erro de validação cruzada (RMSECV), erro quadrático médio de predição (RMSEP) e coeficientes de determinação (Rc<sup>2</sup> e Rp<sup>2</sup>).

## 4.5 Resultados e Discussão

### 4.5.1 Caracterização da Blenda B10 por FTIR

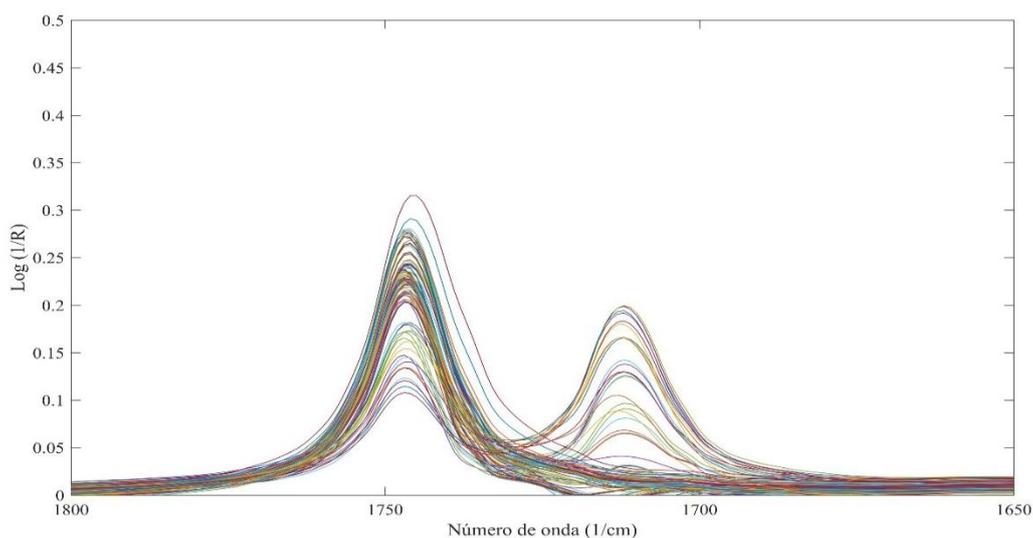
Os espectros obtidos na região de  $4000 - 650 \text{ cm}^{-1}$  a partir das amostras da fase óleo de B10 são apresentados na Figura 24.



**Figura 24:** Espectros de FTIR-ATR das amostras de B10.

As principais diferenças observadas no espectro com a adição de biodiesel ao diesel, é a banda que aparece aproximadamente na região de  $1.745 \text{ cm}^{-1}$  que é referente às vibrações de deformação axial do grupo carbonila ( $\text{-C=O}$ ) presente nos ésteres alifáticos constituintes do biodiesel. E na região de  $1.300$  a  $900 \text{ cm}^{-1}$ , que contém parte da região de “impressão digital” dos compostos, encontram-se as bandas de absorção referentes às vibrações de deformação axial da ligação  $\text{C-O}$  dos ésteres constituintes do biodiesel.<sup>153</sup>

Analisando os espectros de FTIR, umas das principais características das amostras que sofreram degradação é o surgimento de uma banda na região de  $1710 \text{ cm}^{-1}$ , a banda da carbonila diminui sua intensidade com o surgimento desta nova banda, como pode ser observado na Figura 25, o surgimento desta banda indica a formação de novos compostos que pode ser atribuído aos produtos formados durante o processo de degradação, grupo carbonila ( $\text{-C=O}$ ) de cetonas e ácidos carboxílicos absorvem nessa região.<sup>154</sup>



**Figura 25:** Espectros de B10 na região de 1800 a 1650  $\text{cm}^{-1}$ .

#### 4.5.2 *Quantificação da Biomassa por Calibração Multivariada*

Os resultados do modelo de regressão por PLS, *i*PLS e *si*PLS foram construídos empregando validação cruzada *leave-one-out* para a predição da biomassa. Os resultados das figuras de méritos são apresentados na Tabela 4, e estão organizados pelo número de divisões equidistantes feitas no espectro e no caso do *si*PLS de suas combinações de 2, 3 ou 4 intervalos.

A avaliação dos modelos de calibração foi baseada na análise do RMSECV, do  $R_c^2$ , no número de variáveis latentes e considerando as regiões que possuíam informação espectral mais relevante no caso de *i*PLS e *si*PLS. E o melhor modelo de predição foi definido com base naquele que apresentava os menores valores de RMSECV e RMSEP, sendo desejado que não haja grandes diferenças entre os erros para não ocorrer um subajuste ou sobre ajuste, no número de VLs (onde o ideal é a utilização do menor número possível), e nos coeficientes de determinação ( $R_c^2$  e  $R_p^2$ ).

**Tabela 4:** Resultados das figuras de mérito dos modelos de regressão por PLS, *i*PLS e *si*PLS para determinação da biomassa.

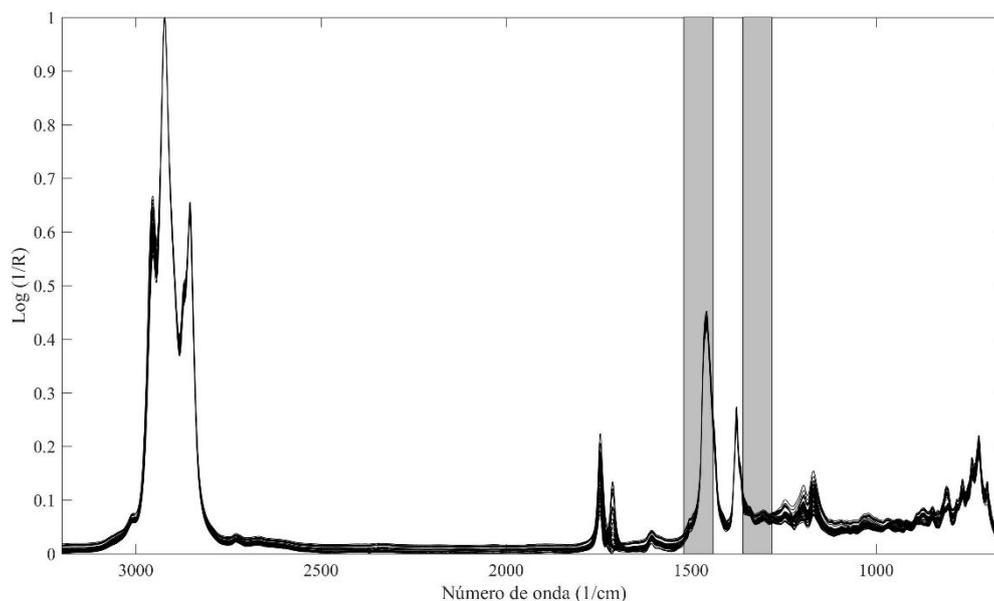
Modelo <sup>a</sup>	Intervalos selecionados	VLs <sup>b</sup>	Rc <sup>2</sup>	RMSECV (mg mL <sup>-1</sup> )	Rp <sup>2</sup>	RMSEP (mg mL <sup>-1</sup> )
PLS	1	11	0,802	1,617	1,09	0,79
<i>i</i> PLS2	2	9	0,770	1,745	0,765	0,951
<i>i</i> PLS4	3	9	0,758	1,806	0,783	1,022
<i>i</i> PLS8	6	9	0,765	1,766	0,748	1,073
<i>i</i> PLS16	10	7	0,727	1,903	0,680	1,205
<i>i</i> PLS32	23	8	0,710	1,971	0,710	1,065
s2 <i>i</i> PLS8	6, 7	10	0,786	1,692	0,732	1,401
s3 <i>i</i> PLS8	2, 3, 6	11	0,803	1,617	0,670	1,605
s4 <i>i</i> PLS8	2, 3, 5 e 6	7	0,743	1,850	0,762	0,953
s2 <i>i</i> PLS16	10, 14	9	0,773	1,740	0,746	1,163
s3 <i>i</i> PLS16	7, 9, 12	11	0,834	1,467	0,740	1,010
s4 <i>i</i> PLS16	4, 10, 13, 16	10	0,786	1,690	0,702	1,132
<b>s2<i>i</i>PLS32</b>	<b>22, 24</b>	<b>9</b>	<b>0,822</b>	<b>1,540</b>	<b>0,815</b>	<b>1,350</b>
s3 <i>i</i> PLS32	20, 22, 25	9	0,836	1,471	0,780	1,420
S4 <i>i</i> PLS32	8, 18, 22, 25	10	0,851	1,396	0,79	1,310

<sup>a</sup> *i*PLS – número de divisões do espectro; <sup>s</sup> – número de intervalos combinados; <sup>b</sup> número de variáveis latentes; Formação em negrito ao melhor modelo de regressão.

A partir dos resultados da análise do número de variáveis latentes, RMSCEV, RMSEP, Rc<sup>2</sup> e Rp<sup>2</sup> dos modelos de regressões usando o método de PLS, *i*PLS e *si*PLS definiu-se que o melhor modelo foi o s2*i*PLS32, onde ocorreu a divisão em 32 partes equidistantes, combinação de 2 intervalos, que apresentou os melhores coeficientes de determinação

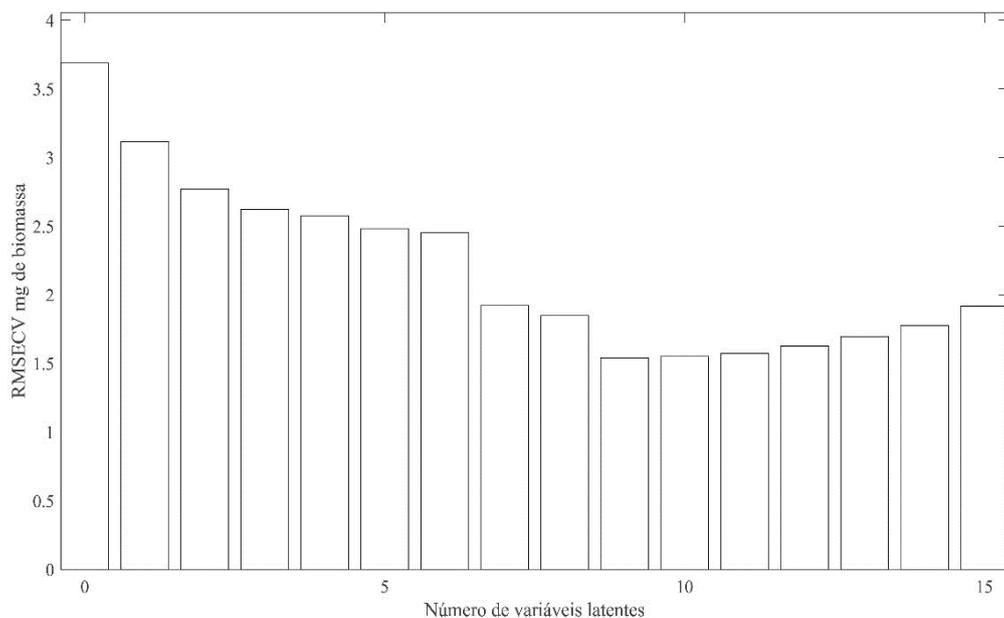
( $R_c^2=0,822$  e  $R_p^2= 0,815$ ), associados a baixos valores de erros:  $RMSECV =1,540$  mg  $mL^{-1}$  e  $RMSEP= 1,350$  mg  $mL^{-1}$ .

Os intervalos selecionados no modelo s2iPLS32 foram 22 e 24, que respectivamente correspondem as faixas espectrais  $1520 - 1441$   $cm^{-1}$  e  $1360 - 1282$   $cm^{-1}$ , Figura 26. Os grupos metilas presentes nos ésteres do biodiesel e no diesel absorvem na região de  $1520-1441$   $cm^{-1}$ .<sup>155</sup>



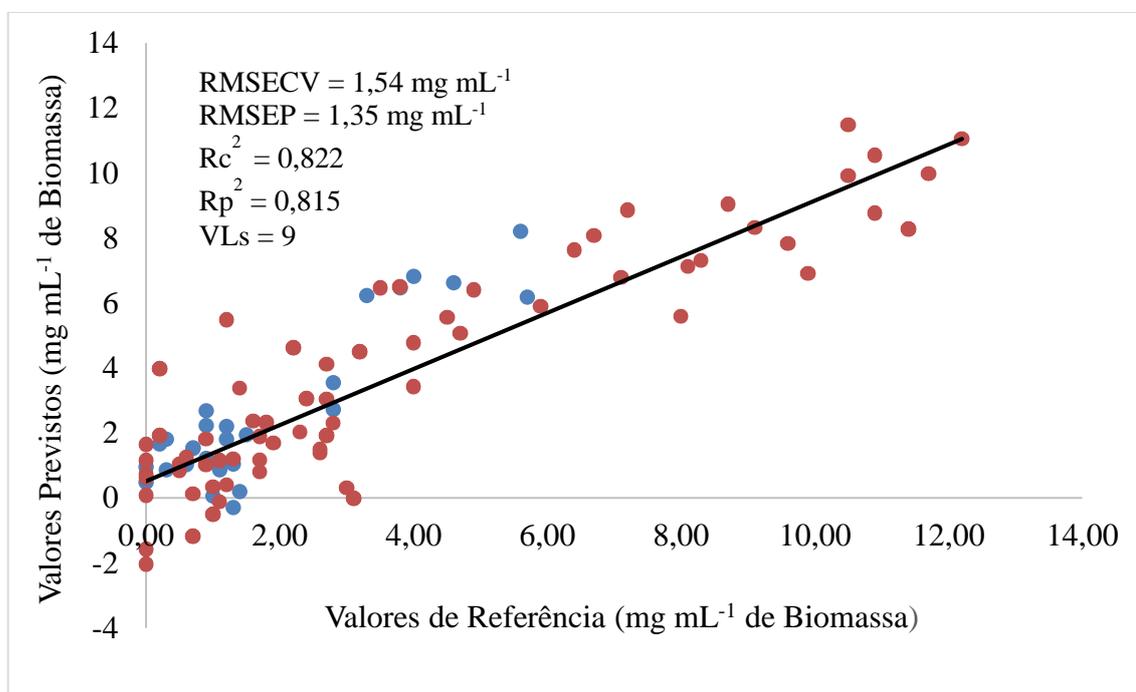
**Figura 26:** Espectro com destaque das regiões 22 e 24 selecionadas no modelo de regressão s2iPLS32.

Na Figura 27, é possível verificar que o número ideal de VLs para o modelo s2iPLS32 é 9, que apresentou um valor baixo de  $RMSECV$ , determinado por validação cruzada, sendo que para valores superiores a esse número de variáveis começa a ocorrer um aumento do erro, indicando que pode haver um sobre ajuste do modelo.



**Figura 27:** Gráfico dos valores de RMSECV por número de variáveis latentes utilizadas pelo modelo *s2iPLS32*.

Na Figura 28, tem-se o gráfico de correlação dos valores de referência, obtido pelo método de gravimetria, com os valores obtidos pelo melhor modelo de regressão através do *siPLS*.



**Figura 28:** Gráfico dos valores de referência versus valores previstos por *siPLS* para determinação de biomassa.

Comparando os modelos obtidos por regressão PLS, verificou-se que a utilização do método de seleção de variáveis por *si*PLS é mais adequada que por *i*PLS, apresentando menores discrepâncias nos erros de RMSECV e RMSEP, maiores coeficientes de determinação ( $R_c^2$  e  $R_p^2$ ) para quantificar biomassa produzida na biodegradação de blendas B10.

Com base nos resultados, o modelo de regressão por mínimos quadrados parciais para quantificação de biomassa apresentou um desempenho promissor, porém os valores de erro (RMSECV e RMSEP) apontados pelo modelo foram um pouco elevados, quando comparado com faixa em estudo (0-12 mg mL<sup>-1</sup> de peso seco). Isto pode ter ocorrido em função dos erros experimentais, pois a determinação da biomassa (peso seco) por gravimetria é uma técnica laboriosa, que envolve uma sequência de etapas: filtração, lavagem, secagem e pesagem das amostras, como foi descrito em materiais e métodos, todas essas etapas estão sujeitas a erros experimentais que são cumulativos e afetam os valores medidos de biomassa.

Uma característica importante do PLS é que o mesmo modela erros em ambas as informações medidas, tanto na obtenção do espectro quanto na quantificação da biomassa e suas contribuições para o modelo são de igual importância, os escores e pesos são dependentes de ambos os dados medidos experimentalmente.<sup>127</sup> Assim os erros associados na quantificação de biomassa e na aquisição dos espectros, contribuíram para os valores de RMSECV e RMSEP.

#### **4.6 Conclusão**

A espectroscopia de infravermelho médio associada a ferramenta de *si*PLS apresentou o melhor modelo, possivelmente pela combinação de intervalos não adjacentes. Porém os valores de RMSECV = 1,54 mg mL<sup>-1</sup> e RMSEP = 1,35 mg mL<sup>-1</sup> de peso seco apresentados indicaram uma possível influência dos erros experimentais no desempenho do modelo.

O emprego da metodologia proposta apresenta como principais vantagens, ao laborioso método gravimétrico, a redução significativa do tempo de análise e a possibilidade de ser facilmente implementada no processo analítico.

## 4.7 Utilização de Imagens Digitais para a Contagem de Colônias de Microrganismos Deteriogênicos de Biodiesel

### 4.7.1 Objetivo

Desenvolver um método de contagem computacional de colônias inicialmente da bactéria *Bacillus pumilus* e da levedura *Meyerozyma guilliermondii*, ambas isoladas a partir do óleo diesel, empregando imagens digitais.

### 4.7.2 Materiais e Métodos

#### 4.7.2.1 Cultura dos Microrganismos

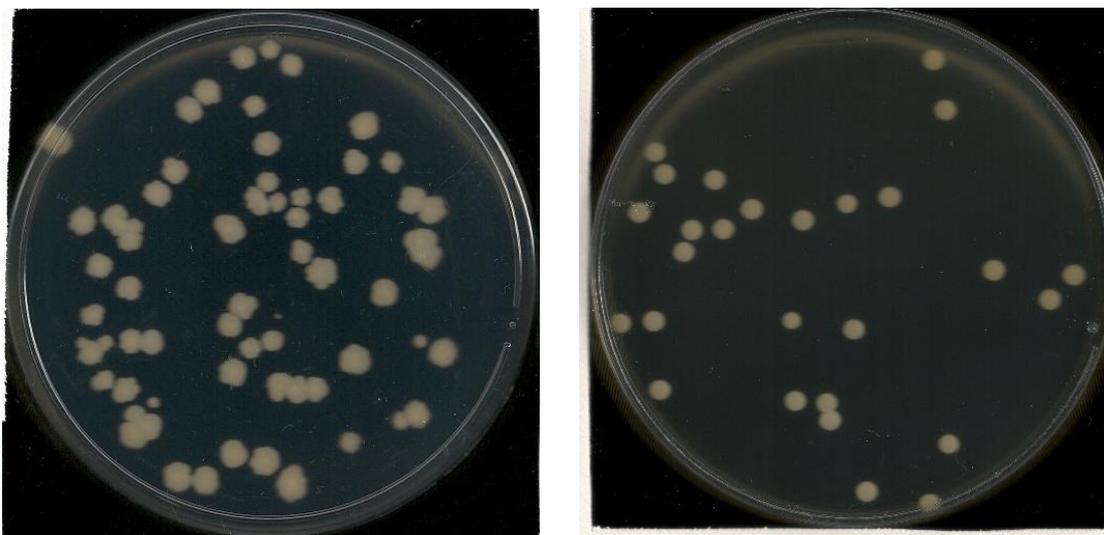
*B.pumilus* foi originalmente isolado do óleo diesel e identificado pelo sequenciamento Sanger (Código do banco de dados MG595229). O isolado *B. pumilus* foi cultivado em caldo BHI por 48 h sob 37°C de incubação, em seguida foi diluído em solução salina estéril (NaCl 0,9%), e alíquotas (0,1 mL) foram semeadas em meio Plate Count Agar (PCA) usando o spread técnica. As placas foram incubadas a 37°C por 48 h. O experimento foi realizado em triplicata. Após o período de incubação, o número de unidades formadoras de colônias em cada placa de Petri foi contado manualmente. Aqueles que não puderam ser contados de forma justa foram descartados. Ao final, foram mantidas 16 amostras, com número de colônias entre 56 e 265.

A levedura *M. guilliermondii* (Gen Bank Code MH047241) foi identificada por Boelter e colaboradores (2018). A levedura foi isolada do diesel e semeada em ágar GYMP (em g L<sup>-1</sup>: glicose 20,0, extrato de levedura 5,0, extrato de malte 20,0, fosfato de sódio monobásico 2,0 e ágar 15,0; todos da marca Himedia) por 48 h a 28°C. O *M. guilliermondii* foi usado em um experimento que avaliou seu potencial de crescimento e degradação de combustíveis.<sup>156</sup> O número de unidades formadoras de colônias em cada placa de Petri foi contado manualmente. Aqueles que não puderam ser contados de forma justa foram descartados. Ao todo, 64 amostras foram mantidas com número de colônias entre 1 e 131.

#### 4.7.2.2 Aquisição das Imagens

Imagens digitais das placas de Petri contendo cada cultura de microrganismo foram obtidas com um scanner comercial (HP Scanjet 300). O aparelho foi o mesmo utilizado por Marcelo e colaboradores (2016) <sup>137</sup>, exceto o material usado para cobrir a mesa do scanner. Neste trabalho, foi aplicada uma borracha branca de etileno-acetato de vinila (EVA) devido a sua superior capacidade de manipulação. Uma janela quadrada foi aberta no centro, grande o suficiente para uma placa de Petri. Esse sistema também contava com uma caixa de papelão com as paredes escura que cobria todo o scanner, o que melhora a qualidade das imagens adquiridas ao reduzir a luminosidade do ruído.

As imagens da colônia de *B. pumilus* foram coletadas com todas as opções automáticas do software do scanner desabilitadas e resolução definida para 300 dpi. Cada cultura foi varrida duas vezes, então girada 180 ° no sentido horário e varrida duas vezes. Um total de 64 amostras de imagens foram coletadas (armazenadas no formato PNG) e posteriormente contadas. A Figura 29 (à esquerda) mostra uma amostra da imagem coletada. Imagens de colônias de *M. guilliermondii* foram obtidas em duplicata (sem rotação). A resolução do scanner foi configurada para 1200 dpi e as imagens foram armazenadas no formato JPG. Um total de 128 imagens foram adquiridas. A Figura 29 (direita) mostra uma amostra da imagem da levedura.



**Figura 29:** Imagens da placa de Petri com colônias de bactérias (esquerda) e leveduras (direita).

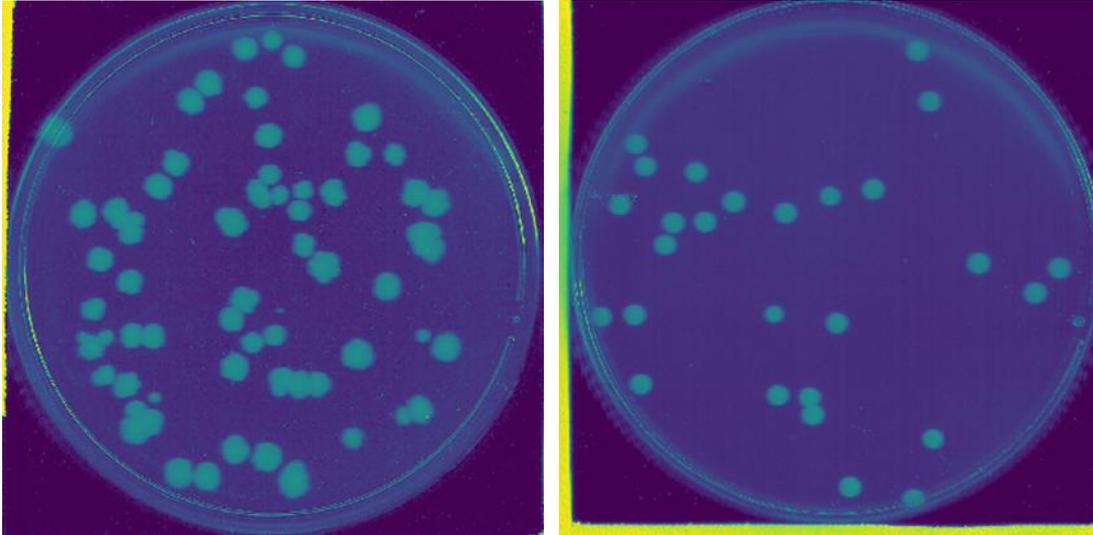
### 4.7.3 Contador Semi-Automático de Colônias

Uma etapa crucial da metodologia proposta foi o desenvolvimento de um código semi-automático para a contagem das colônias. Ele explora a morfologia semelhante das colônias de bactérias e leveduras para calcular o número de unidades formadoras de colônias (UFC) em cada amostra. Ele foi implementado em linguagem Python com a ajuda dos pacotes gratuitos scikit-image versão 0.13.1 (Walt et al. 2014)<sup>157</sup>, NumPy (versão 1.13.3), Matplotlib (versão 2.1.1) e SciPy (versão 0.19). 1-2) e utilizando um computador Intel Core i7-3612QM, dual core, quatro threads, 8 GB de RAM.

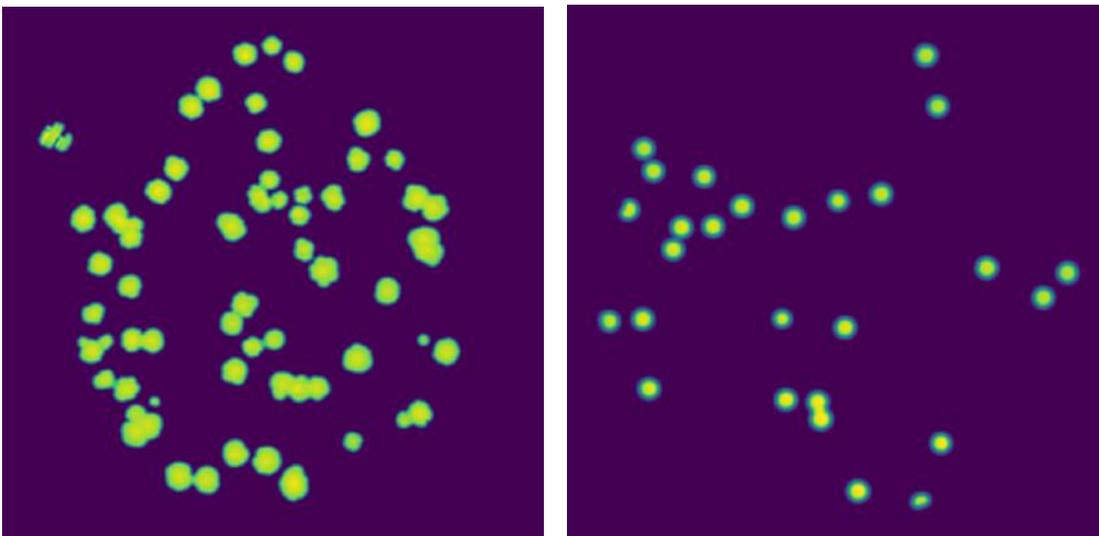
O algoritmo de contagem semi-automático desenvolvido é baseado em aspectos morfológicos semelhantes das características de interesse das colônias de *B. pumilus* e *M. guilliermondii*. Especificamente, ambos os microrganismos formam colônias arredondadas com luminosidade menos intensa nas bordas em comparação com o núcleo. A Figura 29 mostra as imagens das amostras adquiridas e selecionadas de colônias de *B. pumilus* (esquerda) e *M. guilliermondii* (direita). As principais etapas para o código de contagem são discutidas e ilustradas para esses exemplos padrão.

O algoritmo pode ser apresentado em duas etapas principais: pré-processamento da imagem (filtragem) e identificação. Na primeira etapa, a imagem original é redimensionada para  $np \times np$  pixels de imagem e achatada em tons de cinza (escala contínua de 0 a 1). A Figura 30 mostra as imagens redimensionadas e achatadas dos exemplos padrão da Figura 29. Em seguida, um filtro de corte simples é aplicado; define luminosidade zero para todos os pixels abaixo de um determinado mínimo ( $lmin$ ) e acima de um determinado máximo ( $lmax$ ), parâmetros de luminosidade ( $0 \leq lmin < lmax \leq 1$ ). Este processo filtra a maior parte do ruído, exceto a moldura da placa de Petri e algumas bordas no fundo das placas. Para removê-los, um filtro de abertura binário morfológico é aplicado usando círculos como elementos estruturantes em forma de cruz. O raio dos círculos em forma de cruz aplicados ( $ro$ ) é um parâmetro morfológico do algoritmo desenvolvido, podendo ser calibrado a partir de uma imagem de amostra.

O último filtro uniforme de tamanho  $su$  é aplicado para suavizar a luminosidade da superfície e melhorar os resultados de identificação. A Figura 31 mostra as imagens filtradas das amostras padrão.



**Figura 30:** Amostras de imagens redimensionadas e achatadas das culturas selecionadas de *Bacillus pumilus* (esquerda) e *Meyerozyma guilliermondii* (direita).

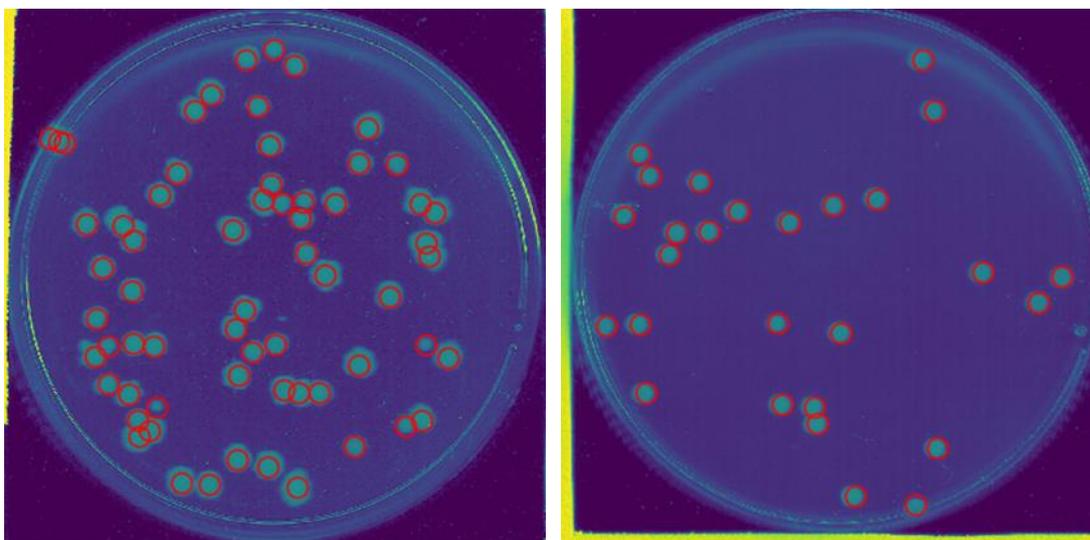


**Figura 31:** Amostras de imagens filtradas das culturas selecionadas de *Bacillus pumilus* (esquerda) e *Meyerozyma guilliermondii* (direita).

Após a etapa de filtragem, a maior parte do ruído e características sem importância são removidos, e a segunda etapa principal compreende a identificação das unidades formadoras de colônias. Aqui, o algoritmo explora o atributo de luminosidade, pois uma unidade formadora de colônia é mais intensa em seu núcleo do que em suas bordas. Pensando nisso, a estratégia de identificação envolve o cálculo dos pontos de máxima

luminosidade local. Para mitigar a redundância, uma distância mínima ( $d_{min}$ ) entre os pontos máximos é imposta.

A Figura 32 mostra as colônias identificadas nas amostras padrão. Os círculos vermelhos são centrados nos máximos locais identificados. O número de tais pontos é o número calculado de unidades formadoras de colônias.



**Figura 32:** Unidades formadoras de colônias identificadas em imagens de culturas selecionadas de *Bacillus pumilus* (esquerda) e *Meyerozyma guilliermondii* (direita). Os círculos vermelhos estão centrados nos pontos máximos locais.

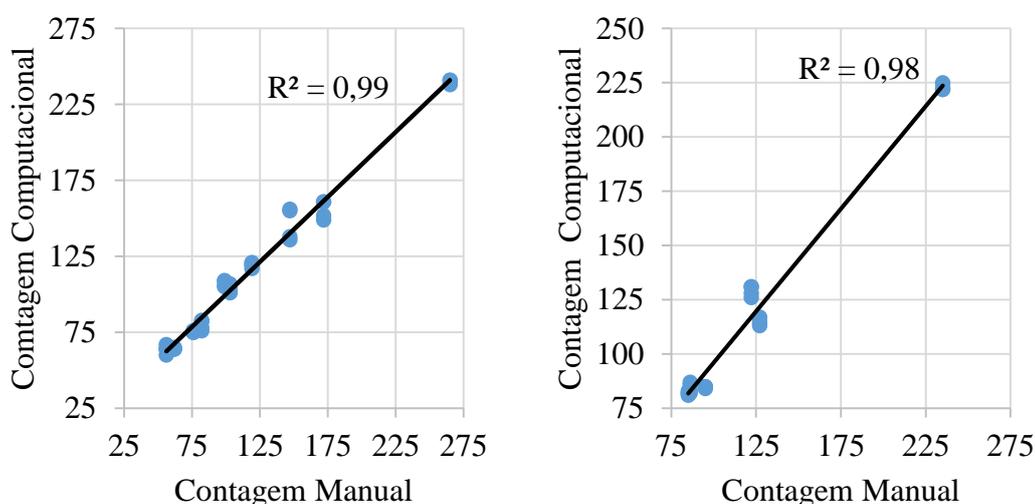
#### 4.8 Resultados e Discussão

Nesta seção, os resultados obtidos dos dois casos de teste são discutidos. Para ambos os casos, comparamos a contagem de referência (contagem manual padrão) com o algoritmo de contagem semi-automático. O algoritmo de contagem semi-automático desenvolvido requer a calibração de seis parâmetros: redimensionamento da resolução da imagem ( $np$ ), corte de luminosidade mínima ( $l_{min}$ ), corte de luminosidade máxima ( $l_{max}$ ), raio dos círculos em forma de cruz do filtro de abertura ( $ro$ ), tamanho do filtro uniforme ( $su$ ), e a distância mínima entre os pontos máximos ( $d_{min}$ ). O primeiro parâmetro é o menos sensível. Se se optar por uma resolução baixa, o contador não funcionará corretamente devido à falta de informação. Por outro lado, uma alta resolução aumenta os custos computacionais (demandas de memória e processamento). Além disso, os outros parâmetros são relativos à resolução da imagem. Para ambos os casos de teste,  $np = 1000$  foi usado. Vale ressaltar que o algoritmo não é muito sensível aos valores dos

parâmetros. Eles podem ser razoavelmente estimados a partir de uma pequena coleção de amostras e então calibrados por um método simples de tentativa e erro. Os resultados da calibração e validação são discutidos abaixo para cada microrganismo avaliado.

#### 4.8.1 *Bacillus Pumilus*

As amostras foram divididas aleatoriamente em conjuntos de calibração e validação (aproximadamente 65% e 35% do total de amostras, respectivamente). O conjunto de calibração continha 40 amostras de imagem com 56–265 colônias por imagem. O conjunto de validação continha 24 amostras de imagens com 85–222 colônias por imagem. Seguindo uma estratégia de tentativa e erro, os parâmetros calibrados para o caso de teste de *B. pumilus* foram:  $lmin = 0,32$ ;  $lmax = 0,65$ ;  $ro = 5$ ;  $su = 7$ ; e  $dmin = 7$ . Com esses parâmetros, a contagem semiautomática ajustou a contagem manual padrão com um coeficiente de regressão ( $R^2$ ) de 0,99 para o conjunto de calibração. Quando testado com o conjunto de validação,  $R^2$  foi 0,98 (Figura 33). Além disso, o procedimento de contagem semiautomática levou aproximadamente 0,5 s por amostra em um notebook comum.



**Figura 33:** Gráficos de calibração (esquerda) e validação (direita) para o caso de teste *Bacillus pumilus*.

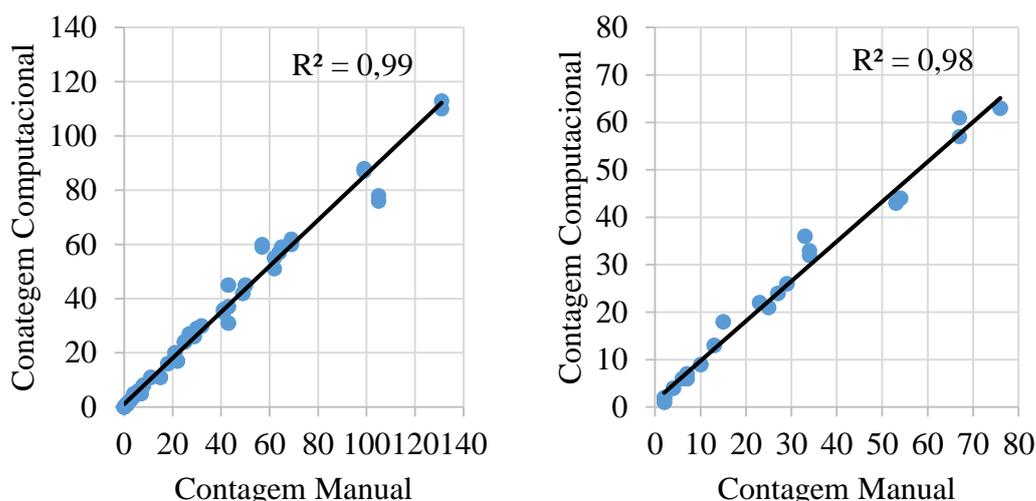
#### 4.8.2 *Meyerozyma Guilliermondii*

Semelhante a *B. pumilus*, as amostras de imagem de *M. guilliermondii* foram distribuídas aleatoriamente em conjuntos de calibração e validação (aproximadamente

66% e 33% do total de amostras, respectivamente). O conjunto de calibração continha 86 amostras de imagens com 1–131 colônias por imagem. O conjunto de validação continha 42 amostras de imagens com 2-76 colônias por imagem.

Os parâmetros calibrados no algoritmo contador foram:  $lmin = 0,32$ ;  $lmax = 0,65$ ;  $ro = 10$ ;  $su = 7$ ; e  $dmin = 15$ . Em comparação com os parâmetros de *B. pumilus*, apenas  $ro$  e  $dmin$  são diferentes, o que pode ser atribuído às diferenças nos aspectos morfológicos. As colônias de leveduras estudadas têm forma circular mais regular e luminosidade mais uniforme que as colônias de *B. pumilus*. Portanto,  $ro$  e  $dmin$  mais altos diminuem as identificações falsas quando há formação de clusters de colônias.

A Figura 34 mostra o gráfico de dispersão e a linha de regressão da contagem manual versus automática. Assim como com *B. pumilus*, os conjuntos de calibração e validação apresentaram um  $R^2$  de 0,99 e 0,98, respectivamente. O tempo computacional foi de cerca de 2 s por amostra, que é muito maior do que o das amostras de bactérias. Essa diferença é explicada pelo aumento na resolução aplicada na aquisição dessas amostras de imagens.



**Figura 34:** Gráficos de calibração (esquerda) e validação (direita) para o caso de teste *Meyerozyma guilliermondii*.

Neste estudo foi apresentado um sistema semi-automático para contagem de colônias de bactérias *Bacillus Pumilus* e da levedura *Meyerozyma guilliermondii*. Os resultados apresentados mostram um bom desempenho do algoritmo desenvolvido para a contagem de colônias. O sistema mostrou uma excelente correlação com o método de contagem tradicional nos dois casos analisados. Algumas vantagens frente as outras metodologias são o tempo de contagem de 3 s, fácil aquisição das imagens com o uso de um scanner comum e não requer conhecimento avançado de operação de software.

Uma das principais vantagens do algoritmo aqui apresentado é a serialização. Uma vez que os parâmetros são calibrados, o algoritmo pode ser aplicado em série (e potencialmente em paralelo) a um conjunto de amostras. No entanto, sua principal limitação é que a calibração depende do crescimento da colônia, o que requer um especialista para determinar o momento de digitalização das amostras. Ele também afeta os limites de menor e maior número de UFC's que o algoritmo terá um bom desempenho.

#### **4.9 Conclusão**

Este estudo apresentou um procedimento novo e robusto para contagem semiautomática de unidades formadoras de colônias de bactérias *B. pumilus* e levedura *M. guilliermondii*. Centrou-se em um aparelho de baixo custo para aquisição de imagens digitais de placas de Petri contendo uma cultura de microrganismo e um código de contagem computacional eficiente.

As imagens digitais das amostras foram adquiridas usando um scanner comercial comum. As imagens serviram de entrada para o código de contagem, que estimou o número de unidades formadoras de colônias em cada amostra, explorando aspectos morfológicos. O algoritmo se baseia em seis parâmetros, que não são muito sensíveis e podem ser calibrados por uma estratégia de tentativa e erro. O procedimento de contagem semiautomática foi testado em comparação com a contagem manual padrão. Os resultados mostraram curvas de regressão de calibração e validação com  $R^2$  de 0,99 e 0,98, respectivamente.

O método semiautomático desenvolvido tem um baixo custo computacional e pode ser facilmente paralelizado para lidar simultaneamente com muitas amostras. O aparelho de aquisição de imagem pode ser adaptado para digitalizar mais de uma placa por vez e também o algoritmo contador pode ser facilmente modificado para lidar com isso.

Além disso, pode ser adaptado para ser aplicado a outras culturas de microrganismos com aspectos morfológicos semelhantes. A bactéria *B. pumilus* e a levedura *M. guilliermondii* estão relacionadas à biodegradação do biodiesel e suas misturas com o diesel. Portanto, a metodologia eficiente e de baixo custo apresentada é uma ferramenta valiosa para analisar a biodegradação desses combustíveis.

## 5 Considerações Finais

Foram propostas três metodologias para analisar a biodegradabilidade do diesel, blendas B10 (diesel/biodiesel) e biodiesel.

A investigação de diferentes combinações de métodos de pré-processamentos na análise de espectros de infravermelho médio de amostras de diesel com diferentes teores de enxofre que foram submetidas a estocagem simulada com microrganismos, auxiliou na identificação de uma região espectral para avaliar a biodegradação em amostras de diesel.

Mediante espectros de infravermelho médio de blendas B10, foi desenvolvida uma metodologia para quantificação da biomassa formada na biodegradação causada por microrganismos, e com base nos resultados a utilização de espectroscopia no infravermelho médio e calibração multivariada é uma metodologia promissora para a quantificação de biomassa.

Por fim, a partir de imagens digitais de placas de Petri foi desenvolvida uma metodologia computacional para contagem do número de colônias de bactérias e leveduras que geralmente causam a degradação de combustíveis. Este trabalho foi publicado recentemente.<sup>158</sup>

## 6 Referências

1. Soares Moreira Cesar Borba, B.; Lucena, A. F. P.; Cunha, B. S. L.; Szklo, A.; Schaeffer, R. Diesel imports dependence in Brazil: A demand decomposition analysis. *Energy. Strategy. Rev.* **2017**, *18*, 63.
2. Issariyakul, T.; Dalai, A. K. Biodiesel from vegetable oils. *Renew. Sust. Energ. Rev.* **2014**, *31*, 446.
3. Knothe, G.; Razon, L. F. Biodiesel fuels. *Prog. Energy Combust. Sci.* **2017**, *58*, 36.
4. BRASIL Lei N° 11.097, De 13 de Janeiro de 2005, Dispõe sobre a introdução do biodiesel na matriz energética brasileira. Legislação Federal, Ministério de Minas e Energia.
5. AGÊNCIA NACIONAL DO PETRÓLEO GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP (BRASIL). Resolução de Diretoria RD N° 248/2021, de 06 de maio de 2021. Aprova a Nota Técnica Conjunta n° 10/2021/ANP que apresenta a proposta de novo modelo de comercialização de biodiesel para atendimento da mistura obrigatória ao Diesel B. Rio de Janeiro.
6. 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* **2019**, *237*, 373.
7. Verma, P.; Sharma, M. P. Review of process parameters for biodiesel production from different feedstocks. *Renew. Sust. Energ. Rev.* **2016**, *62*, 1063.
8. Shah, P.; Wee, C.; White, J.; Sanford, S.; Meier, G. Experimental Determination and Thermodynamic Modeling of Water Content in Biodiesel-Diesel Blends. *Renewable Energy Group, Inc.* **2010**.
9. Stamps, B. W.; Bojanowski, C. L.; Drake, C. A.; Nunn, H. S.; Lloyd, P. F.; Floyd, J. G.; Emmerich, K. A.; Neal, A. R.; Crookes-Goodson, W. J.; Stevenson, B. S. In situ Linkage of Fungal and Bacterial Proliferation to Microbiologically Influenced Corrosion in B20 Biodiesel Storage Tanks. *Front Microbiol* **2020**, *11*.
10. Bücker, F.; Santestevan, N. A.; Roesch, L. F.; Seminotti Jacques, R. J.; Peralba, M. d. C. R.; Camargo, F. A. d. O.; Bento, F. M. Fuel biodegradation and molecular characterization of microbial biofilms in stored diesel/biodiesel blend B10 and the effect of biocide. *Int. Biodeterior. Biodegradation* **2011**, *65*, 172.

11. Soriano, A. U.; Martins, L. F.; Santos de Assumpção Ventura, E.; Teixeira Gerken de Landa, F. H.; de Araújo Valoni, É.; Dutra Faria, F. R.; Ferreira, R. F.; Kremer Faller, M. C.; Valério, R. R.; Catharine de Assis Leite, D.; Lima do Carmo, F.; Peixoto, R. S. Microbiological aspects of biodiesel and biodiesel/diesel blends biodeterioration. *Int. Biodeterior. Biodegradation* **2015**, *99*, 102.
12. de Azambuja, A. O.; Bücker, F.; de Quadros, P. D.; Zhalnina, K.; Dias, R.; Vacaro, B. B.; Correa, C.; Ferrão, M. F.; de Oliveira Camargo, F. A.; Triplett, E.; Bento, F. M. Microbial community composition in Brazilian stored diesel fuel of varying sulfur content, using high-throughput sequencing. *Fuel* **2017**, *189*, 340.
13. Passman, F.; Dobranic, J. Relative biodegradability of b-100 biodiesel and conventional low sulfur diesel fuels. *Proceedings of the 9th international conference on the stability and handling of liquid fuels, Sidges, Spain* **2005**..
14. Gaylarde, C. C.; Bento, F. M.; Kelley, J. Microbial contamination of stored hydrocarbon fuels and its control. *Revista de Microbiologia* **1999**, *30*, 01.
15. Gassen, J.; Bento, F.; Frazzon, A.; Ferrão, M.; Marroni, I.; Simonetti, A. Growth of *Paecilomyces variotii* in B0 (diesel), B100 (biodiesel) and B7 (blend), degradation and molecular detection. *Braz. J. Biol.* **2015**, *75*, 541.
16. Ameen, F.; Moslem, M.; Hadi, S.; Al-Sabri, A. E. Biodegradation of diesel fuel hydrocarbons by mangrove fungi from Red Sea Coast of Saudi Arabia. *Saudi J. Biol. Sci.* **2016**, *23*, 211.
17. Oliveira, M. B.; Varanda, F. R.; Marrucho, I. M.; Queimada, A. J.; Coutinho, J. A. P. Prediction of Water Solubility in Biodiesel with the CPA Equation of State. *Ind. Eng. Chem. Res.* **2008**, *47*, 4278.
18. Passman, F. J. Microbial contamination and its control in fuels and fuel systems since 1980 – a review. *Int. Biodeter. Biodegr.* **2013**, *81*, 88.
19. Sørensen, G.; Pedersen, D. V.; Nørgaard, A. K.; Sørensen, K. B.; Nygaard, S. D. Microbial growth studies in biodiesel blends. *Bioresour. Technol.* **2011**, *102*, 5259.
20. Schleicher, T.; Werkmeister, R.; Russ, W.; Meyer-Pittroff, R. Microbiological stability of biodiesel–diesel-mixtures. *Bioresour. Technol.* **2009**, *100*, 724.

21. Itah, A. Y.; Brooks, A. A.; Ogar, B. O.; Okure, A. B. Biodegradation of International Jet A-1 Aviation Fuel by Microorganisms Isolated from Aircraft Tank and Joint Hydrant Storage Systems. *Bull Environ. Contam. Toxicol.* **2009**, *83*, 318.
22. Martin-Sanchez, P. M.; Gorbushina, A. A.; Toepel, J. An improved test for the evaluation of hydrocarbon degradation capacities of diesel-contaminating microorganisms. *Int. Biodeter. Biodegr.* **2018**, *126*, 216.
23. Bücken, F.; Barbosa, C. S.; Quadros, P. D.; Bueno, M. K.; Fiori, P.; Huang, C. T.; Frazzon, A. P. G.; Ferrão, M. F.; de Oliveira Camargo, F. A.; Bento, F. M. Fuel biodegradation and molecular characterization of microbial biofilms in stored diesel/biodiesel blend B10 and the effect of biocide. *Int. Biodeter. Biodegr.* **2014**, *95*, 346.
24. Aislabie, J.; Foght, J.; Saul, D. Aromatic hydrocarbon-degrading bacteria from soil near Scott Base, Antarctica. *Polar Biology* **2000**, *23*, 183.
25. Cazarolli, J. C.; Guzzato, R.; Samios, D.; Peralba, M. d. C. R.; Cavalcanti, E. H. d. S.; Bento, F. M. Susceptibility of linseed, soybean, and olive biodiesel to growth of the deteriorogenic fungus *Pseudallescheria boydii*. *Int. Biodeter. Biodegr.* **2014**, *95*, 364.
26. Cavalcanti, E. H. D. S.; Zimmer, A. R.; Bento, F. M.; Ferrão, M. F. Chemical and microbial storage stability studies and shelf life determinations of commercial Brazilian biodiesels stored in carbon steel containers in subtropical conditions. *Fuel* **2019**, *236*, 993.
27. Parthipan, P.; Elumalai, P.; Ting, Y. P.; Rahman, P. K. S. M.; Rajasekar, A. Characterization of hydrocarbon degrading bacteria isolated from Indian crude oil reservoir and their influence on biocorrosion of carbon steel API 5LX. *Int. Biodeter. Biodegr.* **2018**, *129*, 67.
28. Wang, H.; Xu, R.; Li, F.; Qiao, J.; Zhang, B. Efficient degradation of lube oil by a mixed bacterial consortium. *J. Environmental Sci.* **2010**, *22*, 381.
29. Martin-Sanchez, P. M.; Becker, R.; Gorbushina, A. A.; Toepel, J. An improved test for the evaluation of hydrocarbon degradation capacities of diesel-contaminating microorganisms. *Int. Biodeter. Biodegr.* **2018**, *129*, 89.
30. Cazarolli, J. C.; Silva, T. L.; da Camara Ribas, R. K.; da Fontoura Xavier Costa, L.; de Moura, T.; Galeazzi, C. F.; Dallé da Rosa, P.; Kenne, D. C.; Carvalho, Â. R.; Valente, P.; Pizzolato, T. M.; Correa, C.; Ferrão, M. F.; Guedes Frazzon, A. P.; Bento, F. M.

Deterioration potential of *Aureobasidium pullulans* on biodiesel, diesel, and B20 blend. *Int. Biodeter. Biodegr.* **2020**, *147*, 104839.

31. Genov, G.; Nodland, E.; Skaare, B. B.; Barth, T. Comparison of biodegradation level and gas hydrate plugging potential of crude oils using FT-IR spectroscopy and multi-component analysis. *Org. Geochem.* **2008**, *39*, 1229.

32. de Almeida, V. E.; da Costa, G. B.; de Sousa Fernandes, D. D.; Gonçalves Dias Diniz, P. H.; Brandão, D.; de Medeiros, A. C. D.; Vêras, G. Using color histograms and SPA-LDA to classify bacteria. *Anal. Bioanal. Chem.* **2014**, *406*, 5989.

33. Chiang, P.-J.; Tseng, M.-J.; He, Z.-S.; Li, C.-H. Automated counting of bacterial colonies by image analysis. *J. Microbiol. Methods* **2015**, *108*, 74.

34. Amaral, B. S.; Novaes, F. J. M.; Ramos, M. d. C. K. V.; Neto, F. R. d. A.; Gioda, A. Comparative profile of pollutants generated by a stationary engine fueled with diesel, biodiesel, and ethanol. *J. Aerosol Sci.* **2016**, *100*, 155.

35. Nespeca, M. G.; Hatanaka, R. R.; Flumignan, D. L.; Oliveira, J. E. d. Rapid and Simultaneous Prediction of Eight Diesel Quality Parameters through ATR-FTIR Analysis. *J. Anal. Methods Chem.* **2018**, *2018*, 1795624.

36. Boletim Mensal Dos Combustíveis Renováveis . Ministério de Minas e energia, Secretaria de Petróleo, Gás Natural e Combustíveis. Departamento de Renováveis. Ed., Dezembro de 2020. Disponível em [www.mme.gov.br](http://www.mme.gov.br). Acesso em: Março de 2020.

37. Zañão, L. R.; Diniz Brito dos Santos, B. C.; Sequinel, R.; Flumignan, D. L.; de Oliveira, J. E. Prediction of Relative Density, Distillation Temperatures, Flash Point, and Cetane Number of S500 Diesel Oil Using Multivariate Calibration of Gas Chromatographic Profiles. *Energy Fuels* **2018**, *32*, 8108.

38. Sachdeva, T. O.; Pant, K. K. Deep desulfurization of diesel via peroxide oxidation using phosphotungstic acid as phase transfer catalyst. *Fuel Process. Technol.* **2010**, *91*, 1133.

39. Resolução ANP n° 50, de 23/12/13, ANP - AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, publicada no Diário Oficial em 24 de Dezembro de 2013.

40. Campos, A. T.; Quintella, C. M.; Meira, M.; Luna, S. Prediction of Sulfur Content in Diesel/Biodiesel Blends Using LED-Induced Fluorescence Associated with Multivariate Calibration. *J. Braz. Chem. Soc.* **2018**, *29*, 1367.
41. Gebremariam, S. N.; Marchetti, J. M. Economics of biodiesel production: Review. *Energy Convers. Manag.* **2018**, *168*, 74.
42. Knothe, G. Biodiesel: Current Trends and Properties. *Top. Catal.* **2010**, *53*, 714.
43. Resolução N° 16, De 29 De Outubro De 2018, Conselho Nacional De Política Energética – CNPE.
44. Hoekman, S. K.; Broch, A.; Robbins, C.; Cenicerros, E.; Natarajan, M. Review of biodiesel composition, properties, and specifications. *Renew. Sust. Energ. Rev.* **2012**, *16*, 143.
45. AGÊNCIA NACIONAL DO PETRÓLEO GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP (BRASIL). Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis 2021. Rio de Janeiro. Disponível em: [www.anp.gov.br](http://www.anp.gov.br). Acesso em 05 julh. 2021.
46. Samios, D.; Pedrotti, F.; Nicolau, A.; Reiznautt, Q. B.; Martini, D. D.; Dalcin, F. M. A Transesterification Double Step Process — TDSP for biodiesel preparation from fatty acids triglycerides. *Fuel Process. Technol.* **2009**, *90*, 599.
47. Dodos, G. S.; Tsesmeli, C. E.; Zannikos, F.; A Study on Microbial Contamination of Alcohol-Blended Unleaded Gasoline. SAE International: 2016.
48. Rodríguez-Rodríguez, C. E.; Rodríguez, E.; Blanco, R.; Cordero, I.; Segura, D. Fungal contamination of stored automobile-fuels in a tropical environment. *J. Environ. Sci.* **2010**, *22*, 1595.
49. Ron, E. Z.; Rosenberg, E. Biosurfactants and oil bioremediation. *Current Opinion in Biotechnology* **2002**, *13*, 249.
50. Fathepure, B. Z. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments *Front Microbiol* **2014**, *5*, 173.
51. Das, N.; Chandran, P. Microbial degradation of petroleum hydrocarbon contaminants: an overview *Biotechnol. Res. Int.* **2011**, *2011*, 941810.

52. Yemashova, N. A.; Murygina, V. P.; Zhukov, D. V.; Zakharyantz, A. A.; Gladchenko, M. A.; Appanna, V.; Kalyuzhnyi, S. V. Biodeterioration of crude oil and oil derived products: a review. *Ver. Environ. Sci. Biotechnol.* **2007**, *6*, 315.
53. Raikos, V.; Vamvakas, S. S.; Sevastos, D.; Kapolos, J.; Karaiskakis, G.; Koliadima, A. Water content, temperature and biocide effects on the growth kinetics of bacteria isolated from JP-8 aviation fuel storage tanks. *Fuel* **2012**, *93*, 559.
54. Bento, M. F.; C. Gaylarde, C. Biodeterioration of stored diesel oil: studies in Brazil. *Int. Biodeter. Biodegr.* **2001**, *47*, 107.
55. Singh, A.; Lin, J. Diesel degradation and biosurfactant production by Gram-positive isolates. *Afr. J. Biotechnol* **2009**, *8*, 5847.
56. Dixit, H.; Lowry, M.; Mohsin, U.; Moond, M.; Kumar, S.; Ankit, C.; Gupta, S. Screening and Identification of Diesel Oil Degrading Bacterial Isolates from Petroleum Contaminated soil of Barmer. *J. Pharm. Chem. Biol. Sci.* **2018**, *6*.
57. Morales-Guzmán, G.; Ferrera-Cerrato, R.; Rivera-Cruz, M. d. C.; Torres-Bustillos, L. G.; Arteaga-Garibay, R. I.; Mendoza-López, M. R.; Esquivel-Cote, R.; Alarcón, A. Diesel degradation by emulsifying bacteria isolated from soils polluted with weathered petroleum hydrocarbons. *Appl. Soil Ecol.* **2017**, *121*, 127.
58. Imron, M. F.; Kurniawan, S. B.; Ismail, N. I.; Abdullah, S. R. S. Future challenges in diesel biodegradation by bacteria isolates: A review. *J.Clean. Prod.* **2020**, *251*, 119716.
59. Ramasamy, S.; Arumugam, A.; Chandran, P. Optimization of *Enterobacter cloacae* (KU923381) for diesel oil degradation using response surface methodology (RSM). *J. Microbiol.* **2017**, *55*, 104.
60. Pasqualino, J. C.; Montané, D.; Salvadó, J. Synergic effects of biodiesel in the biodegradability of fossil-derived fuels. *Biomass Bioenergy* **2006**, *30*, 874.
61. Meyer, D. D.; Beker, S. A.; Bucker, F.; Peralba, M. d. C. R.; Guedes Frazzon, A. P.; Osti, J. F.; Andrezza, R.; Anastácio de Oliveira Camargo, F.; Bento, F. M. Bioremediation strategies for diesel and biodiesel in oxisol from southern Brazil. *Int. Biodeter. Biodegr.* **2014**, *95*, 356.
62. Moreira de Souza, M.; Colla, T. S.; Bucker, F.; Ferrão, M. F.; Huang, C. T.; Andrezza, R.; Anastácio de Oliveira Camargo, F.; Bento, F. M. Biodegradation potential of *Serratiamarcescens* for diesel/biodiesel blends. *Int. Biodeter. Biodegr.* **2016**, *110*, 141.

63. Fregolente, P. B. L.; Fregolente, L. V.; Wolf Maciel, M. R. Water Content in Biodiesel, Diesel, and Biodiesel–Diesel Blends. *J. Chem. Eng Data* **2012**, *57*, 1817.
64. Resolução ANP n° 14, de 11/05/13, ANP - AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, publicada no Diário Oficial em 18 de Maio de 2014.
65. Passman, F. J. Chapter 1-Introduction to Fuel Microbiology Fuel and Fuel System Microbiology: Fundamentals, Diagnosis, and Contamination Control; ASTM International: West Conshohocken, PA, 2003.
66. Hill, E. C., & Hill, G. C. Strategies for resolving problems caused by microbial growth in terminals and retail sites handling biodiesels., In IASH, 2009, The 11th International Conference on Stability, Handling and Use of Liquid Fuels, Prague, Czech Republic, October, 2009.
67. AMERICAN SOCIETY FOR TESTING, MATERIALS - ASTM D 6974:09 (2013) Standard practice for enumeration of viable bacteria and fungi in liquid fuels-filtration and culture procedures. ASTM International, West Conshohocken, PA.
68. Muthukamalam, S.; Sivagangavathi, S.; Dhrishya, D.; Sudha Rani, S. Characterization of dioxygenases and biosurfactants produced by crude oil degrading soil bacteria. *Braz. J. Microbiol.* **2017**, *48*, 637.
69. Zimmer, A.; Cazarolli, J.; Teixeira, R. M.; Viscardi, S. L. C.; Cavalcanti, E. S. H.; Gerbase, A. E.; Ferrão, M. F.; Piatnicki, C. M. S.; Bento, F. M. Monitoring of efficacy of antimicrobial products during 60 days storage simulation of diesel (B0), biodeisel(B100) and blends (B7 and B10) *Fuel* **2013**, *112*, 153.
70. Mohamed, M. A.; Jaafar, J.; Ismail, A. F.; Othman, M. H. D.; Rahman, M. A. In *Membrane Characterization*; Hilal, N., Ismail, A. F., Matsuura, T., Oatley-Radcliffe, D., Chapter 1 - Fourier Transform Infrared (FTIR) Spectroscopy. Eds.; Elsevier: 2017, p 3..
71. Holler, F. J. S., D. A.; Crouch, S. R. *Princípios de Análise Instrumental*; 6 ed., 2009.
72. Coates, J., Interpretation of Infrared Spectra, A Practical Approach. In *Encyclopedia of Analytical Chemistry* **2006**, 10815.
73. Putzig, C. L.; Leugers, M. A.; McKelvy, M. L.; Mitchell, G. E.; Nyquist, R. A.; Papenfuss, R. R.; Yurga, L. Infrared Spectroscopy. *Anal. Chem.* **1994**, *66*, 26.

74. Bureau, S.; Cozzolino, D.; Clark, C. J. Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: A review. *Postharvest Biol. Technol.* **2019**, *148*, 1.
75. Felizardo, P.; Baptista, P.; Menezes, J. C.; Correia, M. J. N. Multivariate near infrared spectroscopy models for predicting methanol and water content in biodiesel. *Anal. Chim. Acta* **2007**, *595*, 107.
76. de Lira, L. F. B.; de Albuquerque, M. S.; Pacheco, J. G. A.; Fonseca, T. M.; Cavalcanti, E. H. D. S.; Stragevitch, L.; Pimentel, M. F. Infrared spectroscopy and multivariate calibration to monitor stability quality parameters of biodiesel. *Microchem J.* **2010**, *96*, 126.
77. Alves, J. C. L.; Poppi, R. J. Biodiesel content determination in diesel fuel blends using near infrared (NIR) spectroscopy and support vector machines (SVM). *Talanta* **2013**, *104*, 155.
78. de Lima, S. M.; Silva, B. F. A.; Pontes, D. V.; Pereira, C. F.; Stragevitch, L.; Pimentel, M. F. In-line monitoring of the transesterification reactions for biodiesel production using NIR spectroscopy. *Fuel* **2014**, *115*, 46.
79. Brereton, R. G.; Jansen, J.; Lopes, J.; Marini, F.; Pomerantsev, A.; Rodionova, O.; Roger, J. M.; Walczak, B.; Tauler, R. *Anal. Bioanal. Chem.* **2018**, *410*, 6691.
80. Brereton, R. G.; Jansen, J.; Lopes, J.; Marini, F.; Pomerantsev, A.; Rodionova, O.; Roger, J. M.; Walczak, B.; Tauler, R. Chemometrics in analytical chemistry—part I: history, experimental design and data analysis tools. *Anal. Bioanal. Chem.* **2017**, *409*, 5891.
81. Brereton, R. G. A short history of chemometrics: a personal view. *J. Chemom.* **2014**, *28*, 749.
82. Lavine, B.; Workman, J. Chemometrics. Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Anal. Chem.* **2008**, *80*, 4519.
83. Geladi, P. Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Spectrochimica Acta Part B: At. Spectrosc.* **2003**, *58*, 767.
84. Kumar, N.; Bansal, A.; Sarma, G. S.; Rawal, R. K. Chemometrics tools used in analytical chemistry: An overview. *Talanta* **2014**, *123*, 186.

85. Flood, M. E.; Goding, J. C.; O'Connor, J. B.; Ragon, D. Y.; Hupp, A. M. Analysis of Biodiesel Feedstock Using GCMS and Unsupervised Chemometric Methods. *J. Am. Oil Chem. Soc.* **2014**, *91*, 1443.
86. Rodrigues, É. V. A.; Silva, S. R. C.; Romão, W.; Castro, E. V. R.; Filgueiras, P. R. Determination of crude oil physicochemical properties by high-temperature gas chromatography associated with multivariate calibration. *Fuel* **2018**, *220*, 389.
87. de Godoy, L. A. F.; Pedroso, M. P.; Hantao, L. W.; Poppi, R. J.; Augusto, F. *Talanta* **2011**, *83*, 1302.
88. Zhang, X.-H.; Zhou, Q.; Liu, Z.; Qing, X.-D.; Zheng, J.-J.; Mu, S.-T.; Liu, P.-H. Comparison of three second-order multivariate calibration methods for the rapid identification and quantitative analysis of tea polyphenols in chinese teas using high-performance liquid chromatography. *J. Chromatogr. A* **2020**, *1618*, 460905.
89. Vempatapu, B. P.; Kumar, J.; Ray, A.; Chhibber, V. K.; Kanaujia, P. K. Determination of biodiesel and used cooking oil in automotive diesel/green diesel fuels through high-performance liquid chromatography. *J. Chromatogr. A* **2020**, *1629*, 461512.
90. Sharma, A.; Verma, A.; Luxami, V.; Melo, J. S.; D'Souza, S. F.; Prakash, N. T.; Prakash, R. *Energy Fuels* **2013**, *27*, 2660.
91. Flumignan, D. L.; Sequinel, R.; Hatanaka, R. R.; Boralle, N.; de Oliveira, J. E. Carbon Nuclear Magnetic Resonance Spectroscopic Profiles coupled to Partial Least-Squares Multivariate Regression for Prediction of Several Physicochemical Parameters of Brazilian Commercial Gasoline. *Energy Fuels* **2012**, *26*, 5711.
92. Heredia-Langner, A.; Cort, J. R.; Grubel, K.; O'Hagan, M. J.; Jarman, K. H.; Linehan, J. C.; Albrecht, K. O.; Polikarpov, E.; King, D. L.; Smurthwaite, T. D.; Bays, J. T. Methodology for the Development of Empirical Models Relating <sup>13</sup>C NMR Spectral Features to Fuel Properties. *Energy Fuels* **2020**, *34*, 12556.
93. Magalhães de Aguiar, L.; Bona, E.; Alberto Colnago, L.; Rodrigues Rohwedder, J. J.; M. Killner, M. H. Application of low-field and medium-resolution <sup>1</sup>H NMR spectroscopy combined with chemometric methods for automotive gasoline quality control. *Fuel* **2020**, *282*, 118684.
94. dos Santos, V. H. J. M.; Pestana, V. Z.; de Freitas, J. S.; Rodrigues, L. F. A preliminary study on traceability of biodiesel mixtures based on the raw materials profiles from

Brazilian regions and fourier transform infrared spectroscopy (FTIR). *Vib. Spectrosc.* **2018**, *99*, 113.

95. Filgueiras, P. R.; Alves, J. C. L.; Poppi, R. J. Quantification of animal fat biodiesel in soybean biodiesel and B20 diesel blends using near infrared spectroscopy and synergy interval support vector regression. *Talanta* **2014**, *119*, 582.

96. Sad, C. M. S.; da Silva, M.; dos Santos, F. D.; Pereira, L. B.; Corona, R. R. B.; Silva, S. R. C.; Portela, N. A.; Castro, E. V. R.; Filgueiras, P. R.; Lacerda, V. Multivariate data analysis applied in the evaluation of crude oil blends. *Fuel* **2019**, *239*, 421.

97. Brereton, R. G. The evolution of chemometrics. *Anal. Methods* **2013**, *5*, 3785.

98. Srivastava, S.; Mishra, G.; Mishra, H. N. Identification and differentiation of insect infested rice grains varieties with FTNIR spectroscopy and hierarchical cluster analysis *Food Chem.* **2018**, *268*, 402.

99. Cebi, N.; Yilmaz, M. T.; Sagdic, O. A rapid ATR-FTIR spectroscopic method for detection of sibutramine adulteration in tea and coffee based on hierarchical cluster and principal component analyses *Food Chem.* **2017**, *229*, 517.

100. Barra, I.; Kharbach, M.; Qannari, E. M.; Hanafi, M.; Cherrah, Y.; Bouklouze, A. Predicting cetane number in diesel fuels using FTIR spectroscopy and PLS regression *Vib. Spectrosc.* **2020**, *111*, 103157.

101. Müller, A. L. H.; Picoloto, R. S.; Mello, P. d. A.; Ferrão, M. F.; dos Santos, M. d. F. P.; Guimarães, R. C. L.; Müller, E. I.; Flores, E. M. M. Total sulfur determination in residues of crude oil distillation using FT-IR/ATR and variable selection methods. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2012**, *89*, 82.

102. Gautam, R.; Vanga, S.; Ariese, F.; Umopathy, S. Review of multidimensional data processing approaches for Raman and infrared spectroscopy. *EPJ Techn. Instrum.* **2015**, *2*, 8.

103. Engel, J.; Gerretzen, J.; Szymańska, E.; Jansen, J. J.; Downey, G.; Blanchet, L.; Buydens, L. M. C. Breaking with trends in pre-processing? *Trends Anal. Chem.* **2013**, *50*, 96.

104. Rinnan, Å.; Berg, F. v. d.; Engelsen, S. B. Review of the most common pre-processing techniques for near-infrared spectra. *Trends Anal. Chem.* **2009**, *28*, 1201.

105. Mishra, P.; Biancolillo, A.; Roger, J. M.; Marini, F.; Rutledge, D. N. New data preprocessing trends based on ensemble of multiple preprocessing techniques. *Trends Anal. Chem.* **2020**, *132*, 116045.
106. Zeaiter, M.; Rutledge, D. 3.04 - Preprocessing Methods. In *Comprehensive Chemometrics*; Brown, S. D., Tauler, R., Walczak, B., Eds.; Elsevier: Oxford, 2009, p 121.
107. Lee, L. C.; Liong, C.-Y.; Jemain, A. A. A contemporary review on Data Preprocessing (DP) practice strategy in ATR-FTIR spectrum. *Chemom. Intell. Lab. Syst.* **2017**, *163*, 64.
108. Bian, X.; Wang, K.; Tan, E.; Diwu, P.; Zhang, F.; Guo, Y. A selective ensemble preprocessing strategy for near-infrared spectral quantitative analysis of complex samples. *Chemom. Intell. Lab. Syst.* **2020**, *197*, 103916.
109. Bro, R.; Smilde, A. K. Principal component analysis. *J. Chemom.* **2003**, *17*, 16..
110. Ferreira, M.M.C. *Quimiometria Conceitos, Métodos e Aplicações*. 2015, 1ª Edição, Editora UNICAMP.
111. Zeaiter, M.; Roger, J. M.; Bellon-Maurel, V. Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods. *Trends Anal. Chem.* **2005**, *24*, 437.
112. Butler, H. J.; Smith, B. R.; Fritsch, R.; Radhakrishnan, P.; Palmer, D. S.; Baker, M. J. Optimised spectral pre-processing for discrimination of biofluids via ATR-FTIR spectroscopy. *Analyst* **2018**, *143*, 6121.
113. Monakhova, Y. B.; Tsikin, A. M.; Mushtakova, S. P. Processing of NMR, UV, and IR spectrometric data prior to chemometric simulation by independent component and principal component analysis. *J. Anal. Chem.* **2016**, *71*, 554.
114. Randolph, T. Scale-based normalization of spectral data. *Cancer biomark* **2006**, *2*, 135.
115. Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chemistry* **1964**, *36*, 1627.
116. Bro, R.; Smilde, A. K. Principal component analysis. *Anal. Methods* **2014**, *6*, 2812.

117. Jolliffe, I.; Cadima, J. Principal component analysis: A review and recent developments. *Phil. Trans. Math. Phys. Eng. Sci.* **2016**, *374*, 20150202.
118. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37.
119. Olivieri, A. Introduction to Multivariate Calibration. *A Practical Approach*, Springer International Publishing: Cham, Switzerland, **2018**.
120. Brereton, R. G. Chemometrics: Data Analysis for the Laboratory and Chemical Plant. *Wiley* **2003**.
121. Brereton, R. G. Introduction to multivariate calibration in analytical chemistry. *Analyst* **2000**, *125*, 2125.
122. Olivieri, A. C. Analytical Advantages of Multivariate Data Processing. One, Two, Three, Infinity? *Anal. Chem.* **2008**, *80*, 5713.
123. Kościelniak, P.; Wieczorek, M. Univariate analytical calibration methods and procedures. A review. *Anal. Chim. Acta* **2016**, *944*, 14.
124. Brereton, R. G. Chemometrics: Data Driven Extraction for Science, Chapter 2: John Wiley and Sons, **2018**.
125. Ferrão, M. F.; Viera, M. d. S.; Pazos, R. E. P.; Fachini, D.; Gerbase, A. E.; Marder, L. Simultaneous determination of quality parameters of biodiesel/diesel blends using HATR-FTIR spectra and PLS, iPLS or siPLS regressions. *Fuel* **2011**, *90*, 701.
126. Filgueiras, P. R.; Sad, C. M. S.; Loureiro, A. R.; Santos, M. F. P.; Castro, E. V. R.; Dias, J. C. M.; Poppi, R. J. Determination of API gravity, kinematic viscosity and water content in petroleum by ATR-FTIR spectroscopy and multivariate calibration. *Fuel* **2014**, *116*, 123.
127. Brereton, R. G. Chemometrics and Multivariate Analysis. In *digital Encyclopedia of Applied Physics*, **2009**, 1105
128. Wold, S.; Høy, M.; Martens, H.; Trygg, J.; Westad, F.; MacGregor, J.; Wise, B. M. PLS-regression: a basic tool of chemometrics. *J. Chemom.* **2009**, *23*, 67.
129. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109.

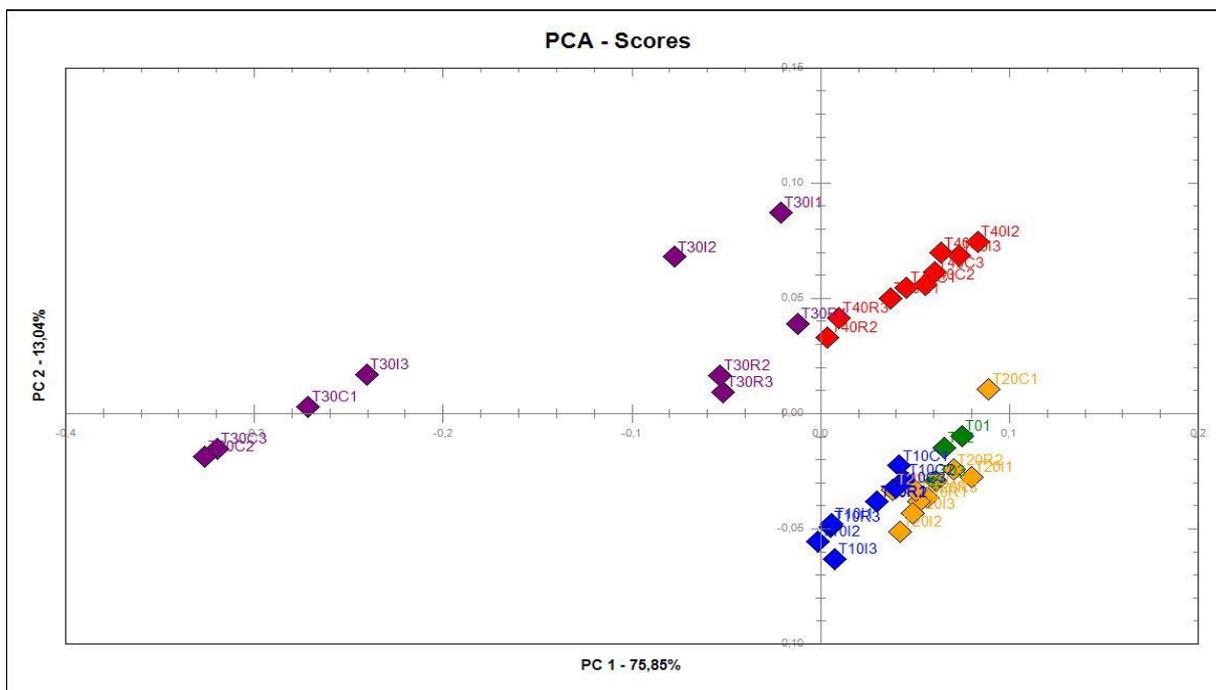
130. Facco, P.; Masiero, A.; Beghi, A. Advances on multivariate image analysis for product quality monitoring. *J. Process Control* **2013**, *23*, 89.
131. Gonzalez, R. C. Woods, R. E. Digital Image Processing. Pearson; 3ª edição, **2007**..
132. Herrero-Latorre, C.; Barciela-García, J.; García-Martín, S.; Peña-Crecente, R. M. Detection and quantification of adulterations in aged wine using RGB digital images combined with multivariate chemometric techniques. *Food Chem x*. **2019**, *3*, 100046.
133. Albizu, G.; Bordagaray, A.; Dávila, S.; Garcia-Arrona, R.; Ostra, M.; Vidal, M. Analytical control of nickel coating baths by digital image analysis. *Microchem. J.* **2020**, *154*, 104600.
134. Vidal, M.; Garcia-Arrona, R.; Bordagaray, A.; Ostra, M.; Albizu, G. Simultaneous determination of color additives tartrazine and allura red in food products by digital image analysis. *Talanta* **2018**, *184*, 58.
135. Lima, C. M. d.; Fernandes, D. D. S.; Pereira, G. E.; Gomes, A. d. A.; Araújo, M. C. U. d.; Diniz, P. H. G. D. Variable selection with a support vector machine for discriminating *Cryptococcus* fungal species based on ATR-FTIR spectroscopy. *Food Chem.* **2020**, *312*, 126060.
136. Costa, G. B.; Fernandes, D. D. S.; Almeida, V. E.; Araújo, T. S. P.; Melo, J. P.; Diniz, P. H. G. D.; Vêras, G. Digital image-based classification of biodiesel. *Talanta* **2015**, *139*, 50.
137. Marcelo, M. C. A.; Mariotti, K. C.; Ortiz, R. S.; Ferrão, M. F.; Anzanello, M. J. Scott test evaluation by multivariate image analysis in cocaine samples. *Microchem. J.* **2016**, *127*, 87.
138. Ates, H.; Gerek, O. N. An image-processing based automated bacteria colony counter. In *2009 24th International Symposium on Computer and Information Sciences* **2009**, p 18.
139. Goyal, M.; Singh, M. Machine vision based bacteria colony counter. *Int. J. Emerging Technol. Adv. Eng.* **2008**, *2*, 58-69
140. Chen, W.-B.; Zhang, C. An automated bacterial colony counting and classification system. *Inf. Syst. Front* **2009**, *11*, 349.

141. Barbedo, J. G. A. Method for Counting Microorganisms and Colonies in Microscopic Images. In *2012 12th International Conference on Computational Science and Its Applications 2012*, p 83.
142. Yoon, S. C., Lawrence, K. C., Park, B. Automatic counting and classification of bacterial colonies using hyperspectral imaging. *Food Bioprocess Technol.* **2015**, *8*, 2047–2065.
143. Zhu, G.; Yan, B.; Xing, M.; Tian, C. Automated counting of bacterial colonies on agar plates based on images captured at near-infrared light. *J. Microbiol. Methods* **2018**, *153*, 66.
144. Braga, J. W. B.; Junior, A. A. d. S.; Martins, I. S. Determination of viscosity index in lubricant oils by infrared spectroscopy and PLSR. *Fuel* **2014**, *120*, 171.
145. Zagonel, G. F.; Peralta-Zamora, P.; Ramos, L. P. Multivariate monitoring of soybean oil ethanolysis by FTIR. *Talanta* **2004**, *63*, 1021.
146. Pimentel, M. F.; Ribeiro, G. M. G. S.; da Cruz, R. S.; Stragevitch, L.; Pacheco Filho, J. G. A.; Teixeira, L. S. G. Determination of biodiesel content when blended with mineral diesel fuel using infrared spectroscopy and multivariate calibration. *Microchem. J.* **2006**, *82*, 201.
147. Guimarães, E.; Mitsutake, H.; Gontijo, L. C.; de Santana, F. B.; Santos, D. Q.; Neto, W. B. Infrared Spectroscopy and Multivariate Calibration for Quantification of Soybean Oil as Adulterant in Biodiesel Fuels. *J. Am. Oil Chem. Soc.* **2015**, *92*, 777.
148. Helfer, G.; Bock, F.; Marder, L.; Furtado, J.; Costa, A.; Ferrao, M. Chemostat: Exploratory multivariate data analysis software. *Química Nova* **2015**.
149. Silverstein, R.M.; Webster, F.X.; Kiemle, D.J.; D.L. Bryce. *Spectrometric Identification of Organic Compounds*, 2014, 8ª Edição, Editora John Wiley & Sons, Inc.
150. de Azambuja, A. O.; Cazarolli, J. C.; de Quadros, P. D.; Bücker, F.; Correa, C.; Casagrande, J. C.; Ferrão, M. F.; Celikkol, S.; Ogawa, A.; Bonifay, V.; Sunner, J. A.; Beech, I. B.; Bento, F. M. Effect of Sulfur Content on Microbial Composition and Biodegradation of a Brazilian Diesel and Biodiesel Blend (B10). *Energy Fuels* **2017**, *31*, 12305.

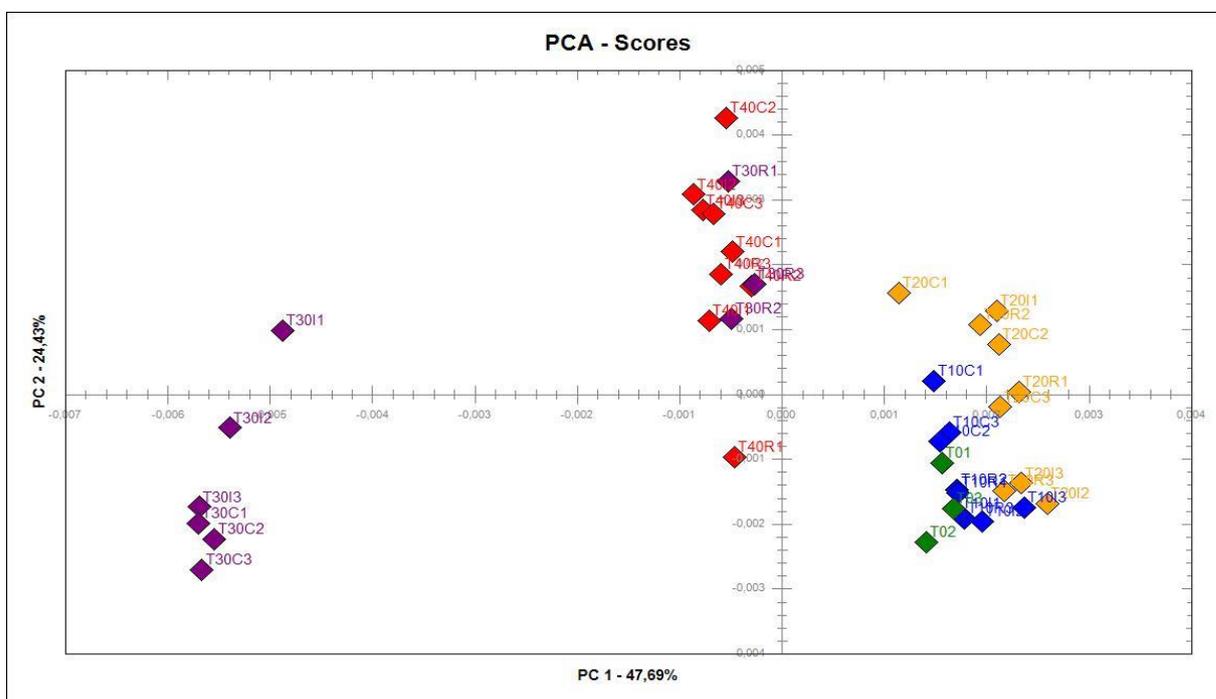
151. Nørgaard, L.; Saudland, A.; Wagner, J.; Nielsen, J. P.; Munck, L.; Engelsen, S. B. Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. *Appl. Spectrosc.* **2000**, *54*, 413.
152. Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137.
153. Mohamed Shameer, P.; Mohamed Nishath, P. Chapter 8 - Exploration and enhancement on fuel stability of biodiesel: A step forward in the track of global commercialization. In *Advanced Biofuels*; Azad, A. K., Rasul, M., Eds.; Woodhead Publishing: 2019, p 181.
154. Coates, J., Interpretation of Infrared Spectra, A Practical Approach. In *Encyclopedia of Analytical Chemistry* **2006**, 10815.
155. Abbas, O.; Rebufa, C.; Dupuy, N.; Permanyer, A.; Kister, J. PLS regression on spectroscopic data for the prediction of crude oil quality: API gravity and aliphatic/aromatic ratio. *Fuel* **2012**, *98*, 5.
156. Boelter, G.; Cazarolli, J. C.; Beker, S. A.; de Quadros, P. D.; Correa, C.; Ferrão, M. F.; Galeazzi, C. F.; Pizzolato, T. M.; Bento, F. M. Pseudallescheria boydii and Meyerozyma guilliermondii: behavior of deteriogenic fungi during simulated storage of diesel, biodiesel, and B10 blend in Brazil. *Environ. Sci. Pollut. Res.* **2018**, *25*, 30410.
157. van der Walt S, S. J., Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T. Scikit-image: image processing in Python. *PeerJ* **2:e453** **2014**.
158. Correa, C.; Konzen, P. H. A.; Carvalho, Â. R.; Giovanella, P.; Bento, F. M.; Ferrão, M. F., Use of digital images to count colonies of biodiesel deteriogenic microorganisms. *J. Microbiol. Methods* **2020**, *178*, 106063

## ANEXOS

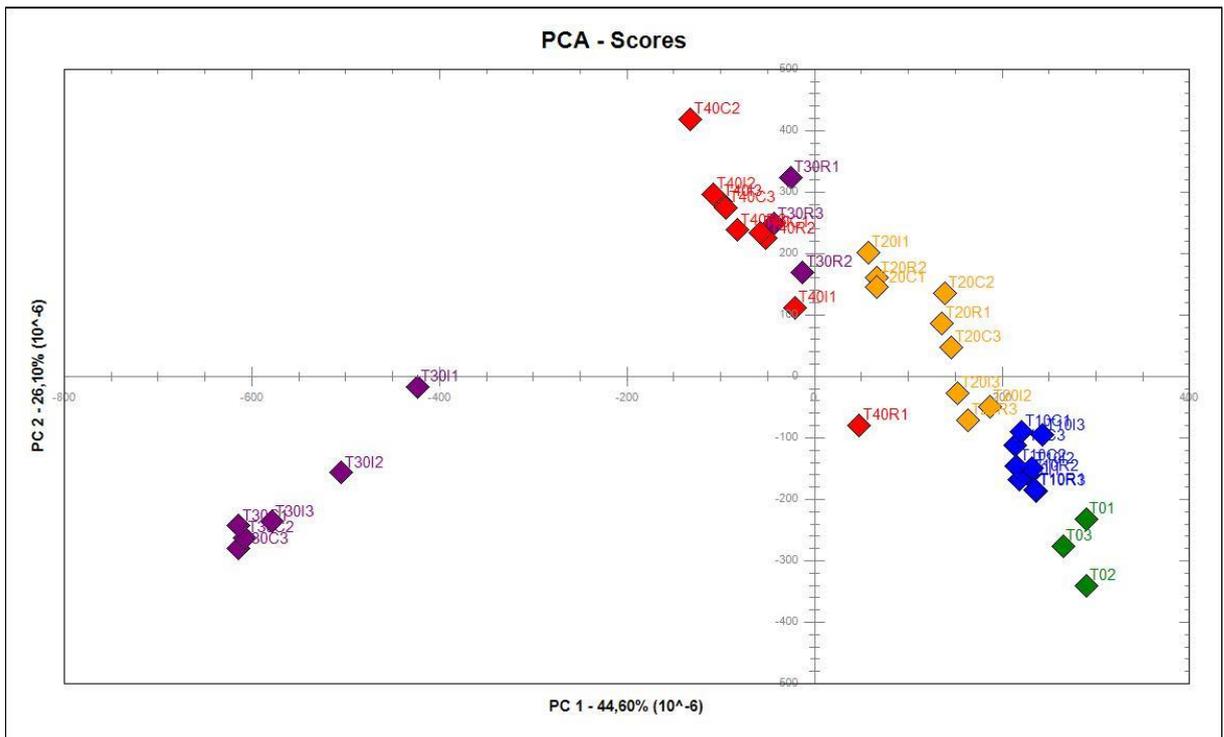
Figuras 1, 2 e 3 referentes as amostras de Diesel S50



**Figura 1:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados e centrados na média das amostras de diesel S50: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

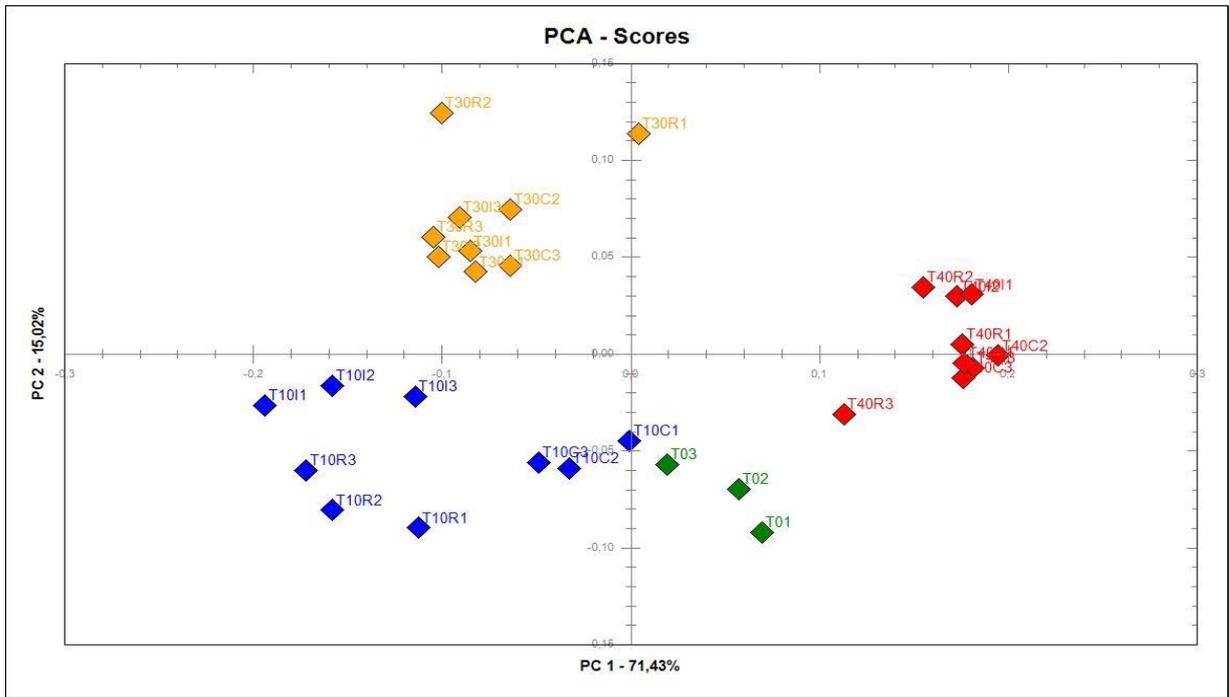


**Figura 2:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados, com primeira derivada e centrados na média das amostras de diesel S50: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

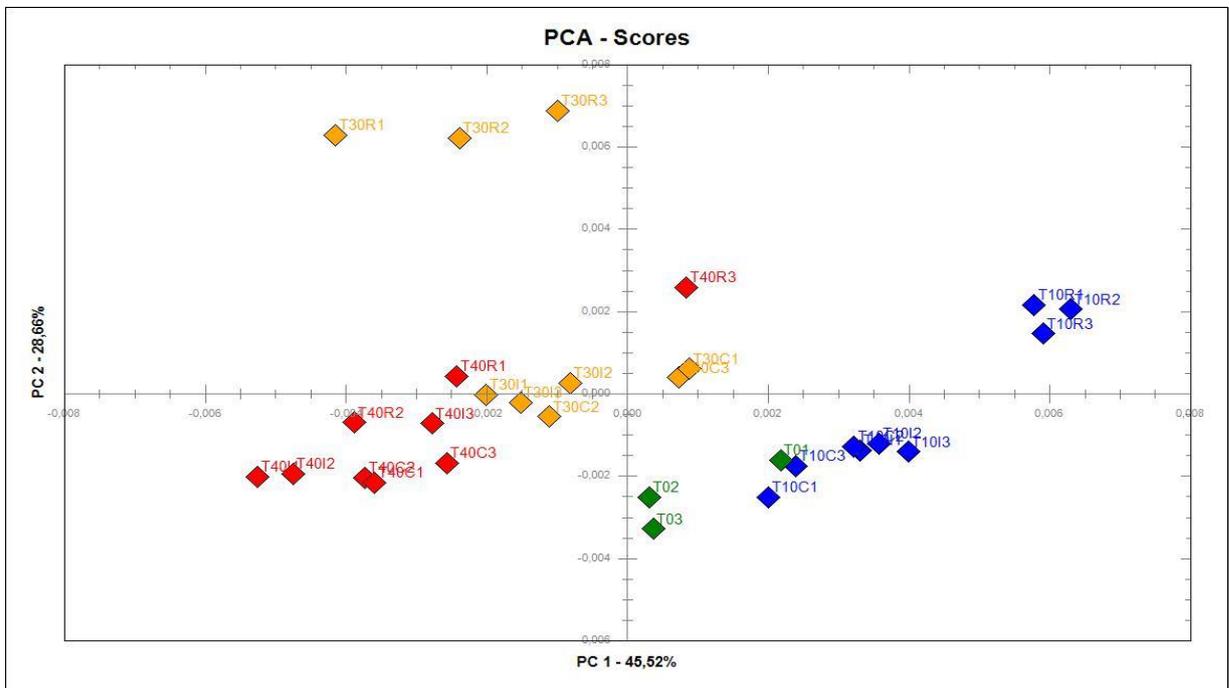


**Figura 3:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados, com segunda derivada e centrados na média das amostras de diesel S50: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

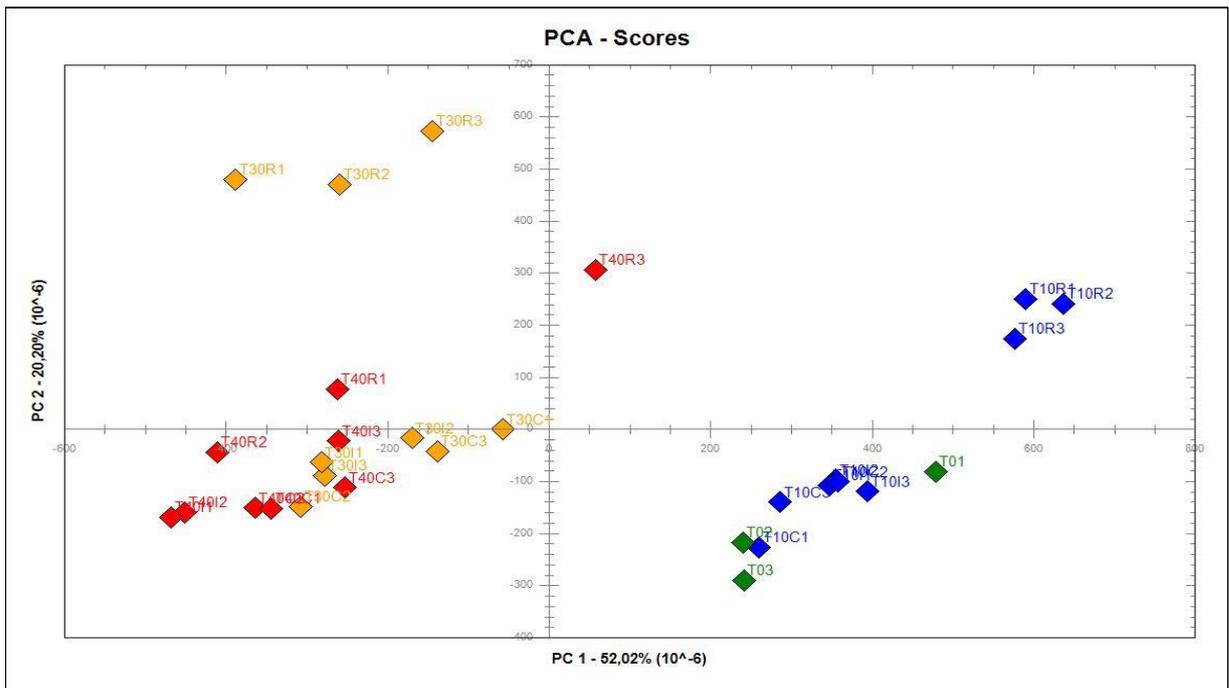
**Figuras 4, 5 e 6 referentes as amostras de Diesel S500**



**Figura 4:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados e centrados na média das amostras de diesel S500: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

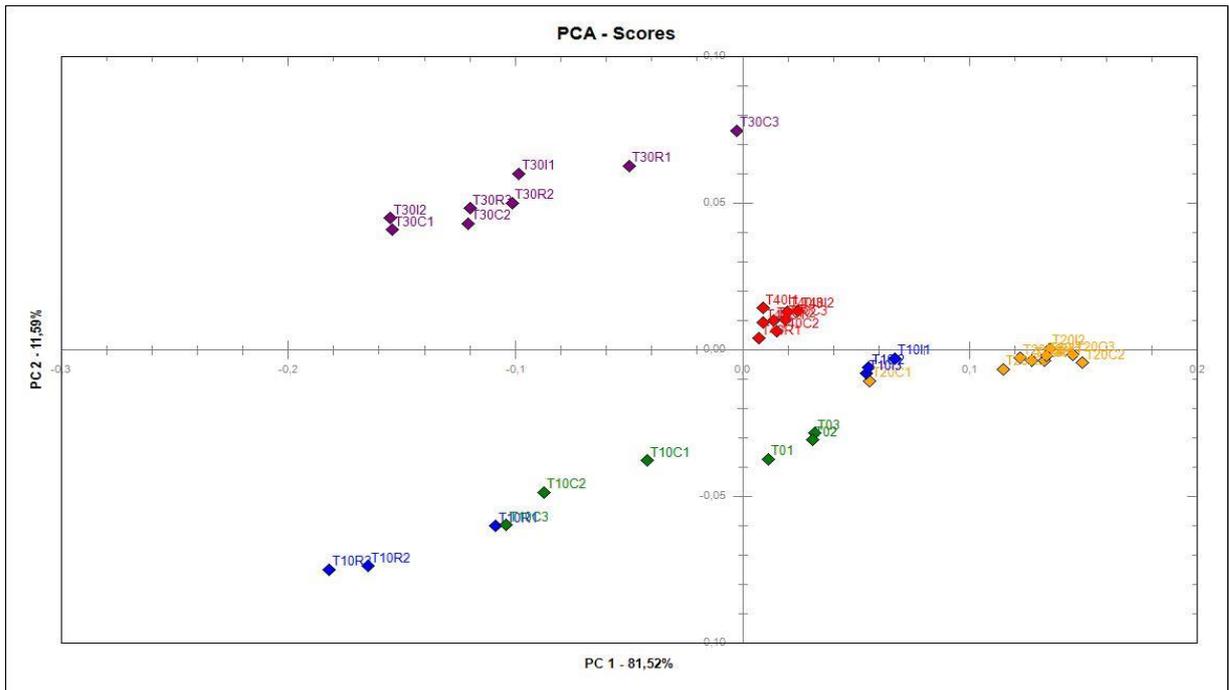


**Figura 5:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados, com primeira derivada e centrados na média das amostras de diesel S500: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

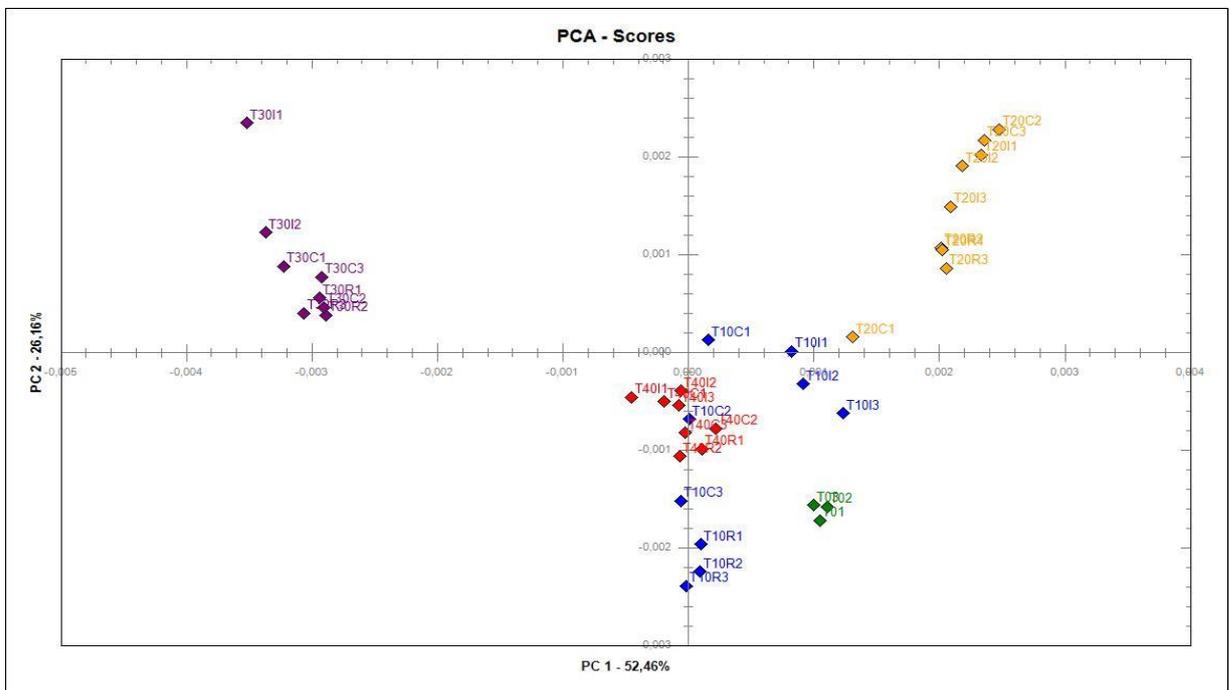


**Figura 6:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados, com segunda derivada e centrados na média das amostras de diesel S500: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

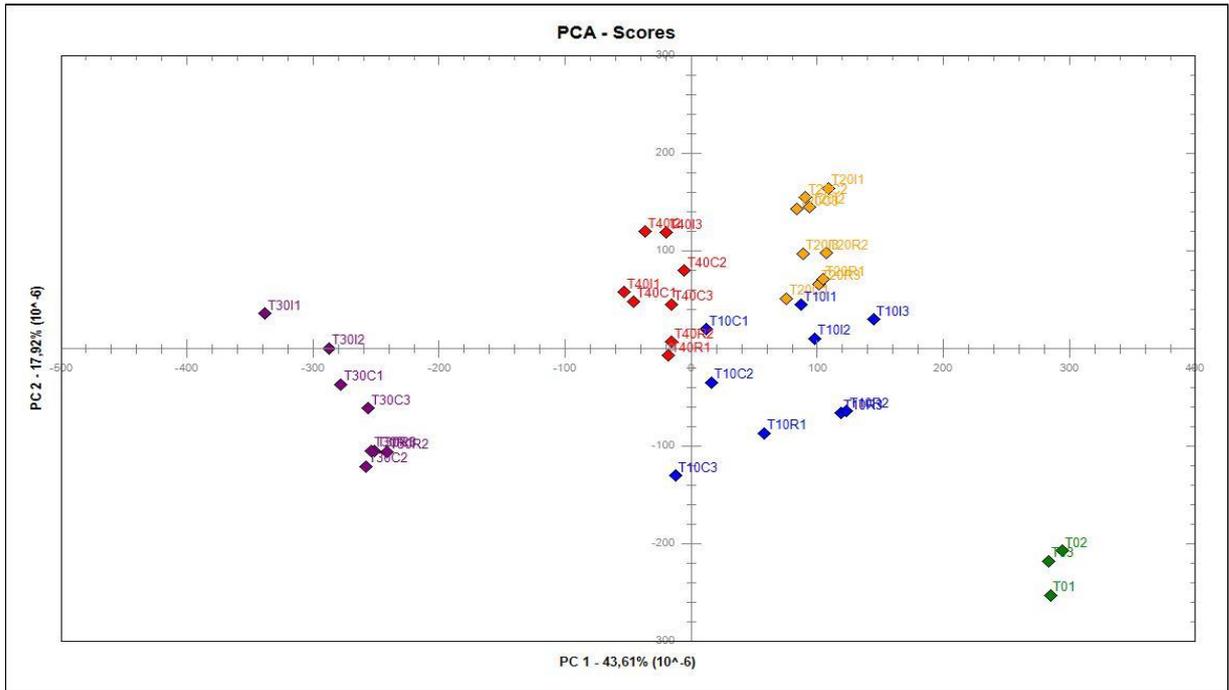
**Figuras 7, 8 e 9 referentes as amostras de Diesel S1800**



**Figura 7:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados e centrados na média das amostras de diesel S1800: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).



**Figura 8:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados, com primeira derivada e centrados na média das amostras de diesel S1800: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).



**Figura 9:** Gráfico de escores PC1xPC2 da PCA dos espectros de FTIR-HATR normalizados, com segunda derivada e centrados na média das amostras de diesel S1800: tempos T0 (verde), T10 (azul), T20 (roxo), T30 (laranja) e T40 (vermelho).

**Tabela 1:** Resultado da determinação de biomassa

<b>Amostra</b>	<b>Biomassa (mg mL<sup>-1</sup>)</b>
1	0,0
2	0,0
3	0,0
4	0,0
5	0,0
6	0,0
7	0,0
8	0,0
9	0,0
10	1,3
11	1,0
12	0,7
13	1,8
14	1,1
15	1,4
16	6,7
17	7,2
18	5,6
19	0,9
20	1,2
21	0,2
22	2,7
23	2,4
24	2,8
25	10,5
26	6,4
27	1,7
28	1,2
29	1,9
30	3,2
31	10,9
32	9,9
33	11,7
34	1,1
35	1,0
36	0,6
37	10,9
38	12,2
39	10,5
40	0,7
41	0,5
42	0,9

**Tabela 1:** Resultado da determinação de biomassa. (continuação)

<b>Amostra</b>	<b>Biomassa (mg mL<sup>-1</sup>)</b>
43	0,9
44	0,3
45	0,6
46	8,3
47	3,0
48	0,7
49	3,1
50	0,5
51	1,1
52	1,2
53	5,7
54	3,5
55	8,1
56	4,0
57	3,8
58	4,9
59	4,6
60	3,3
61	8,0
62	0,3
63	0,2
64	0,2
65	1,4
66	1,5
67	9,6
68	0,9
69	0,6
70	2,6
71	0,6
72	2,6
73	0,6
74	1,1
75	0,7
76	5,9
77	4,5
78	4,0
79	2,8
80	1,2
81	1,2
82	1,7
83	8,7
84	7,1

**Tabela 1:** Resultado da determinação de biomassa. (continuação)

<b>Amostra</b>	<b>Biomassa (mg mL<sup>-1</sup>)</b>
85	1,3
86	1,3
87	2,3
88	2,7
89	11,4
90	2,6
91	2,8
92	2,7
93	11,4
94	2,6
95	2,8
96	2,7
97	2,2
98	1,6