

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**Utilização de espectroscopia na região do infravermelho
próximo com instrumento portátil para caracterização de
petróleos e derivados**

**Portable near-infrared spectrometer to characterize crude oils and its
derivatives**

Francine Dalapícola dos Santos

Tese de Doutorado em Química

**VITÓRIA
2022**

Francine Dalapícola dos Santos

**Utilização de instrumento portátil de espectroscopia na região
do infravermelho próximo para caracterização de petróleos e
derivados**

Tese apresentada ao Programa de Pós-Graduação em Química do Centro de Ciências Exatas da Universidade Federal do Espírito Santo como requisito parcial para obtenção do Título de Doutor em Química.

Área de Concentração: Química

Linha de Pesquisa: Química do Petróleo e Biocombustíveis.

Orientador: Prof. Dr. Paulo Roberto Filgueiras

Co-orientador(a): Prof. Dr. Wanderson Romão

**VITÓRIA
2022**

Utilização de instrumento portátil de espectroscopia na região
do infravermelho próximo para caracterização de petróleos e
derivados

Francine Dalapícola dos Santos

Tese submetida ao Programa de Pós-Graduação em Química da Universidade
Federal do Espírito Santo como requisito parcial para a obtenção do grau de
Doutor em Química.

Aprovado(a) em 21/03/2022 por:

Prof. Dr. Paulo Roberto Filgueiras
Universidade Federal do Espírito Santo
Orientador

Prof. Dr. Wanderson Romão
Universidade Federal do Espírito Santo
Co-orientador

Prof. Dr. Eustáquio Vinícius Ribeiro de Castro
Universidade Federal do Espírito Santo

Prof. Dr. Eneas Nascimento
Universidade Federal do Espírito Santo

Prof. Dr. Elcio Cruz de Oliveira
Pontifícia Universidade Católica do Rio de Janeiro

Dra. Luciana Assis Terra
Agilent Technologies

Universidade Federal do Espírito Santo
Vitória, Março de 2022

A todos que buscam conhecimento.

AGRADECIMENTOS

Agradeço à minha família Zeni, Sergio e Franciele pelo incentivo de todas as formas aos meus estudos, pelo amor e confiança sempre. Se eu chego aqui, foi por nunca terem limitado o que eu poderia ser e aonde eu poderia chegar.

Aos amigos feitos durante toda minha trajetória terrestre, mas principalmente por aqueles que se tornaram minha segunda família: Satie, Ana Caroline, Beatriz, Natã, Samanta, Nayara, Livia e Raphaela.

A todos que colaboraram com o Laboratório de Caracterização e Processamento Primário. Vou sempre me sentir parte dessa equipe, pois é onde eu fiz morada por boa parte da minha construção científica. Nominalmente agradeço à Dra. Cristina Sad e o Prof. Dr. Valdemar Lacerda Jr. por terem me acompanhado no início dessa jornada e me incentivado a fazer pesquisa.

Aos alunos do Laboratório de Quimiometria e Laboratório Petrolômica e Forense pelos quatro anos de ensinamentos compartilhados, principalmente a Flávia, Márcia, Pedro Henrique, Gabrielly, Ellisson, Mariana, Matheus, Layla e Iago por também terem sido parceiros de trabalho.

Aos professores Dr. Paulo Roberto Filgueiras e Wanderson Romão por acompanharem cada dificuldade e alegrias enfrentadas na execução desse trabalho, pela orientação e por estarem sempre disponíveis para que esse projeto conseguisse ser realizado.

Ao prof. Dr. Elcio Cruz de Oliveira por trazer a problemática inicial e disponibilizar as amostras. Foi um desafio enorme, mas o aprendizado foi do mesmo tamanho.

À equipe do NCQP-LABPETRO pela disponibilização dos equipamentos e materiais necessários para a realização deste projeto.

Às agências de fomento CAPES/FAPES pelo incentivo concedido através da bolsa de estudos.

E aos membros da banca por aceitarem o convite pra participarem nessa importante etapa pra atestar e contribuir com a qualidade desse trabalho. Não é ciência, se não houver revisão por pares.

“Nada na vida deve ser temido, apenas compreendido. Agora é o tempo de entendermos mais para que tenhamos menos medo.”

Marie Curie

LISTA DE FIGURAS

Figura 4.1. a) Produção científica anual entre 2000 a 2021 com os termos de entrada “NIR” e “crude oil” com um total de 141 artigos. b) Produção científica anual entre 2000 a 2021 com os termos de entrada “FTIR” e “crude oil” com um total de 141 artigos.	30
Figura 4.2. a) Palavras-chave mais frequentes de acordo com os termos de entrada “NIR” e “crude oil”. b) Palavras-chave mais frequentes de acordo com os termos de entrada “FTIR” e “crude oil”.	31
Figura 4.3. a) Fontes mais relevantes de acordo com os termos de entrada “NIR” e “crude oil”. b) Fontes mais relevantes de acordo com os termos de entrada “FTIR” e “crude oil”.	33
Figura 4.4. a) Produção científica dos países de acordo com os termos de entrada “NIR” e “crude oil”. b) Produção científica dos países de acordo com os termos de entrada “FTIR” e “crude oil”.	34
Figura 4.5. Diagrama de um espectrofotômetro de feixe duplo. A aplicação do feixe duplo favorece a detecção de forma simultânea da amostra em questão e da referência utilizada, portanto, qualquer erro do tipo aleatório presente no sistema é compensado.	43
Figura 4.6. Diagrama de um espectrômetro com Transformada de Fourier. Um feixe de uma fonte de luz é dividido em dois feixes pelo interferômetro. Esses dois feixes percorrem em caminhos distintos até se encontrarem na mesma região espacial do detector, formando um padrão de interferência. O espelho móvel é responsável pela modulação do sinal óptico a partir do seu deslocamento horizontal que altera o padrão de interferência.	47
Figura 5.1. a) Espectrômetro MicroNIR e acessórios acoplados sobre ele: b) Gota e c) Vial Holder (com vial de 4 mL inserido).	53
Figura 5.2. a) Visão superior do acessório (painel refletor na tampa à direita, janela amostral retangular com tampa transparente protetora sobre ela à esquerda); b) Visão lateral do sistema fechado; c) Amostragem; d) Visão lateral da janela amostral coberta apenas pela tampa protetora.	54
Figura 5.3. Modo Vial: a) Acessório vazio acoplado sobre MicroNIR; b) Painel refletor em vial de vidro transparente de 4 mL; c) Vial de 4 mL com tampa rosqueada preenchido com amostra de petróleo e inserido dentro do acessório Vial.	55
Figura 5.4. a) Espectrômetro de infravermelho FT-IR/FT-NIR Spectrum 400 Perkin Elmer com acessório de refletância total atenuada (UATR – IR) e b) Cella de vidro de 35 mm vazia e com c) peça transflatora utilizadas na análise	56
Figura 5.5. Espectros brutos de infravermelho na região do infravermelho próximo adquiridos pelo instrumento portátil MicroNIR para os Petróleo A, B e C no Modo (a) Vial; (c) Gota 1; (e) Gota 2; E tratados com primeira derivada (b) Vial; (d) Gota Antigo; (f) Gota Novo.	59
Figura 5.6. Gráficos dos modelos PCA construídos a partir dos espectros adquiridos para o modo vial (a – c) com 88.62% da variância explicada (E.V.), e para o modo gota 1 com 99.19% E.V. (d – f) para os grupos (●) Petróleo A, (▲) Petróleo B e (◆)	

Petróleo C; Gráfico dos escores para (a,d) PC1 versus PC2, (b,e) PC1 x PC2 x PC3; (c,f) Pesos das variáveis (comprimentos de onda) para as três PC.	60
Figura 5.7. Gráficos do modelo PCA para o modo gota 1 com 99.85% E.V. (a – c) para os grupos (●) Petróleo A, (▲) Petróleo B e (♦) Petróleo C; Gráficos dos scores para a) PC1 versus PC2, b) PC1 x PC2 x PC3; c) Loadings das variáveis (comprimentos de onda) para as três PC.	62
Figura 5.8. Gráfico dos scores do modelo PLS-DA construído com dados NIR obtidos no equipamento portátil para a probabilidade das amostras pertencerem à classe do (a) Petróleo A (amostras em azul); (b) B (amostras em vermelho); (c) C (amostras em verde). Amostras de treinamento (o) e teste (Δ).	63
Figura 5.9.a -c. Espectros de infravermelho na região do infravermelho próximo adquiridos pelo instrumento de bancada para os Modos 1, 2 e 3, respectivamente.	64
Figura 5.10. Gráfico dos scores e loadings, respectivamente, dos modelos PCA construídos para o (a,d) Modo 1, (b,e) Modo 2 e (c, f) Modo 3. (○) Petróleo A; (□) Petróleo B e; (▲) Petróleo C.	65
Figura 6.1. Resumo das principais etapas da metodologia construída	71
Figura 6.2. Espectros brutos das amostras de petróleos adquiridos com os espectrômetros (a) Micronir e (b) FT-NIR	74
Figura 6.3. Curvas de calibração dos modelos PLS °API, TET e TNB, respectivamente, para (a-c) Micronir I e (d-f) FT-NIR I	78
Figura 6.4. Curvas de calibração dos modelos PLS para estimar TET (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.	81
Figura 7.1. Espectro NIR médio dos petróleos e óleo de motor (a) bruto e (b) pré-tratado com primeira derivada.	91
Figura 7.2. Resposta calculada para a probabilidade das amostras pertencerem a classe (a) Petróleo, (b) Óleo de Motor dos modelos PLS-DA.	92
Figura 7.3. Espectros NIR dos 53 blends (a) sem tratamento e pré-tratados com (b) centralização e (c) primeira derivada e normalização.	94
Figura 7.4. Curvas de calibração dos modelos (a,c,e) PLS e (b,c,f) SVR para estimar os teores de (a,b) Óleo leve, (c,d) Óleo pesado, (e,f) Óleo de motor.	96
Figura 7.5. Espectros (a, b) brutos e (c,d) pré-tratados com primeira derivada para (a,c) nafta e gasolina e (b,d) diesel e querosene.	97
Figura 7.6. Resposta calculada para a probabilidade das amostras pertecerem a cada classe (a) nafta, (b) gasolina, (c) diesel, (d) querosene para o modelo PLS-DA com 5 VL.	99
Figura 7.7. Curvas de calibração dos modelos PLS para estimativa do teor de (a) nafta e (b) querosene construídos com 4 VL em um intervalo de confiança de 95%.	101
Figura 0.1. Curvas de calibração dos modelos PLS Sat, Aro, Res e Asf, respectivamente, para (a-d) Micronir I e (e-h) FT-NIR I.	134
Figura 0.2. Curvas de calibração dos modelos PLS para estimar °API com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.	135
Figura 0.3. Curvas de calibração dos modelos PLS para estimar TNB (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.	136
Figura 0.4. Curvas de calibração dos modelos PLS para estimar Sat (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.	137
Figura 0.5. Curvas de calibração dos modelos PLS para estimar Aro (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.	138

Figura 0.6.Curvas de calibração dos modelos PLS para estimar Res (% m/m) com os dados **(a)** Micronir I, **(b)** Micronir II, **(c)** FT-NIR I, **(d)** FT-NIR II e **(e)** FT-NIR III.139

Figura 0.7. Curvas de calibração dos modelos PLS para estimar Asf (% m/m) com os dados **(a)** Micronir I, **(b)** Micronir II, **(c)** FT-NIR I, **(d)** FT-NIR II e **(e)** FT-NIR III.140

LISTA DE TABELAS

Tabela 1.1. Artigos utilizando equipamentos portáteis em petróleos e derivados entre 2018 – 2022	16
Tabela 4.1. Divisão do espectro infravermelho (BURNS; CIURCZAK, 2007)	24
Tabela 4.2. Fontes utilizadas na espectroscopia de infravermelho (SKOOG <i>et al.</i> , 2014)	43
Tabela 4.3. Detectores utilizados na espectroscopia de infravermelho (SKOOG <i>et al.</i> , 2014)	44
Tabela 5.1. Parâmetros de desempenho do modelo PLS-DA com 5 variáveis latentes.	63
Tabela 5.2. Parâmetros de desempenho dos modelos PLS-DA para os Modos 1, 2, 3 para espectros NIR obtidos com equipamento de bancada.	66
Tabela 6.1. Informações sobre as análises de referência utilizadas para caracterização das amostras de petróleo.....	71
Tabela 6.2. Desempenho dos modelos PLS para as sete propriedades construídos com os dados NIR dos instrumentos portátil e de bancada, respectivamente.	75
Tabela 6.3. Parâmetros de desempenho dos modelos PLS para estimar sete propriedades físico-químicas em petróleos com os conjuntos de dados NIR na faixa espectral de 10.000 cm ⁻¹ a 6.000 cm ⁻¹ (Micronir II, FT-NIR II e FT-NIR III).	79
Tabela 6.4. Parâmetros de desempenho dos modelos de regressão (PLS e SVR) para a estimativa das propriedades físico-químicas construídos com os dados Micronir I.	82
Tabela 6.5. Parâmetros de desempenho dos modelos de regressão SVR (Micronir I) e PLS (FT-NIR I) para a estimativa das sete propriedades físico-químicas.	83
Tabela 7.1. Parâmetros de desempenho do modelo PLS-DA para as classes Petróleo e Óleo de motor.....	93
Tabela 7.2. Parâmetros de desempenho para os modelos SVR e PLS para quantificar óleo leve, óleo pesado e óleo de motor nas misturas ternárias.....	95
Tabela 7.3. Parâmetros de desempenho para o modelo PLS-DA para as classes Nafta, Gasolina, Diesel e Querosene.	98
Tabela 7.4. Parâmetros de desempenho para os modelos PLS para estimar teor de nafta em gasolina e querosene em diesel.	100

RESUMO

Esse trabalho desenvolveu metodologias analíticas baseadas na análise multivariada de dados (classificação e regressão) utilizando espectros NIR obtidos com um espectrômetro portátil. O foco foi caracterizar petróleos quanto as suas propriedades físico-químicas, classificar por origem e discriminar adulterantes em petróleos e combustíveis. Utilizou-se um espectrômetro NIR portátil e comparou-se o desempenho dos modelos obtidos construindo a mesma metodologia em um espectrômetro de bancada. As amostras de petróleos e combustíveis (nafta, gasolina, diesel e querosene) foram cedidas pelo Cenpes e Transpetro, compreendendo, então, amostras do campo de produção brasileiro. Os modelos utilizados para resolução dos problemas de classificação foram PCA e PLS-DA. Na construção da metodologia para petróleos, o modelo PCA não possibilitou a discriminação das três origens das amostras de petróleo, mas obteve-se boa resposta com o modelo PLS-DA. Os parâmetros de desempenho (especificidade, sensibilidade e precisão) na etapa de predição obteve 100% de taxa de classificação correta com os dados obtidos com equipamento portátil, enquanto o equipamento de bancada apresentou taxa de erro de classificação para todas as classes. Para a estimativa das propriedades físico-químicas ($^{\circ}$ API, TET, TNB, SARA) de um conjunto de 182 petróleos, utilizou-se PLS e SVR. Com a utilização do SVR a estimativa das propriedades com equipamento portátil se torna equivalente ao equipamento de bancada, apesar de o espectrômetro NIR portátil não adquirir a faixa espectral entre 6.000 e 4.000 cm^{-1} , aonde estão localizadas bandas de absorção descritivas sobre a composição química das amostras (combinações e primeiros sobretons das ligações SH, NH, CH, OH). Por fim, na discriminação e quantificação de adulterantes em petróleos e combustíveis, o modelo PLS-DA conseguiu discriminar petróleos de óleos de motor usados com sensibilidade, especificidade e precisão acima de 94% e conseguiu-se quantificar com PLS esse adulterante com RMSEP de 4,8 % m/m. Enquanto para combustíveis, o modelo permitiu taxa de classificação de 100%. Enquanto, para quantificação, obteve-se para as misturas de gasolina e nafta, LD é 1,3% m/m e LQ 4,4% m/m com RMSEP 1,4% m/m. Da mesma forma, para as misturas de diesel e querosene, o modelo PLS permitiu a quantificação do querosene com LD 2,8 % m/m, LQ 9,3 % m/m com RMSEP de 11,4 % m/m. Dessa forma, o espectrômetro portátil se mostrou eficiente e com resultados competitivos para todas aplicações. Torna-se, assim, uma nova ferramenta a ser explorada em aplicações para o estudo de petróleos e seus derivados.

Texto do resumo.

Palavras-chave: Petróleo. Infravermelho próximo. Portabilidade. Análise multivariada de dados.

ABSTRACT

This work developed analytical methodologies based on multivariate data analysis (classification and regression) using NIR spectra obtained with a portable spectrometer. The focus was to characterize oils in terms of their physicochemical properties, classify them by origin and discriminate adulterants in oils and fuels. A portable NIR spectrometer was used and the performance of the models obtained by building the same methodology on a benchtop spectrometer was compared. The oil and fuel samples (naphtha, gasoline, diesel and kerosene) were provided by Cenpes and Transpetro, thus comprising samples from the Brazilian production field. The models used to solve the classification problems were PCA and PLS-DA. In the construction of the methodology for petroleum, the PCA model did not allow the discrimination of the three origins of the petroleum samples, but a good response was obtained with the PLS-DA model. The performance parameters (specificity, sensitivity and precision) in the prediction step obtained a 100% correct classification rate with the data obtained with portable equipment, while the benchtop equipment presented a classification error rate for all classes. To estimate the physicochemical properties ($^{\circ}$ API, TET, TNB, SARA) of a set of 182 oils, PLS and SVR were used. With the use of SVR, the estimation of properties with portable equipment becomes equivalent to benchtop equipment, although the portable NIR spectrometer does not acquire the spectral range between 6,000 and 4,000 cm^{-1} , where descriptive absorption bands are located on the chemical composition. samples (combinations and first overtones of SH, NH, CH, OH bonds). Finally, in the discrimination and quantification of adulterants in oils and fuels, the PLS-DA model was able to discriminate oils from used motor oils with sensitivity, specificity and precision above 94% and it was possible to quantify this adulterant with RMSEP of 4 .8% w/w. While for fuels, the model allowed a classification rate of 100%. While, for quantification, it was obtained for the mixtures of gasoline and naphtha, LD is 1.3% m/m and LQ 4.4% m/m with RMSEP 1.4% m/m. Likewise, for the mixtures of diesel and kerosene, the PLS model allowed the quantification of kerosene with LD 2.8% m/m, LQ 9.3% m/m with RMSEP of 11.4% m/m. Thus, the portable spectrometer proved to be efficient and with competitive results for all applications. It thus becomes a new tool to be explored in applications for the study of petroleum and its derivatives.

Keyword: Crude oil; Near-infrared; Portability; Multivariate data analysis.

SUMÁRIO

1. INTRODUÇÃO.....	15
2. OBJETIVOS.....	18
2.1. Objetivo geral	18
2.2. Objetivos Específicos	18
3. FUNDAMENTAÇÃO TEÓRICA	19
3.1. Espectroscopia na região do infravermelho próximo	19
4. REVISÃO.....	21
Petróleos através das lentes da espectroscopia na região do infravermelho próximo	21
4.1. Introdução	21
4.2. Espectroscopia de Infravermelho	23
4.3. Utilização da espectroscopia NIR em petróleo.....	24
4.4. Utilização da espectroscopia MIR em petróleo	26
4.5. Comparação entre o MIR e o NIR	29
4.6. Quimiometria.....	34
4.6.1. Modelos de Regressão.....	34
4.6.2. Modelos de Classificação	39
4.7. Instrumentação	41
4.7.1. Instrumentos Dispersivos	42
4.7.2. Instrumentos FTIR	45
4.8. Conclusão	48
5. APLICAÇÃO 1	50
Comparação entre equipamento portátil e de bancada na classificação de petróleos	50
5.1. Introdução	50
5.2. Metodologia.....	51
5.2.1. Amostras	51
5.2.2. Ensaios de infravermelho próximo.....	52
5.2.3. Análise dos dados	56
5.3. Resultados e discussão	57

5.3.1. Análise exploratória para os diferentes modos de aquisição dos espectros	57
5.3.2. Comparação com instrumento de bancada para a classificação das amostras de petróleos	64
5.4. Conclusão	67
6. APLICAÇÃO 2	69
Caracterização de petróleos com espectrômetro NIR portátil	69
6.1. Introdução	69
6.2. Metodologia.....	70
6.2.1. Amostras.....	71
6.2.2. Espectroscopia NIR	71
6.2.3. Análise de dados	72
6.3. Resultados e discussão	73
6.4. Conclusão	84
7. APLICAÇÃO 3	85
Discriminação de petróleos e combustíveis utilizando um espectrômetro NIR portátil	85
7.1. Introdução	85
7.2. Metodologia.....	88
7.2.1. Espectroscopia NIR	88
7.2.2. Modelo qualitativo	89
7.2.3. Modelo quantitativo.....	89
7.3. Resultados e discussão	90
7.3.1. Conjunto de dados 1: Petróleos e Óleo de motor	90
7.3.2. Conjunto de dados 2: Combustíveis	97
7.4. Conclusão	101
8. CONCLUSÕES.....	103
9. REFERÊNCIAS BIBLIOGRÁFICAS	106
ANEXOS	Erro! Indicador não definido.

1. INTRODUÇÃO

O uso do equipamento portátil de espectroscopia de infravermelho na região do próximo (*Micronir®*) traz qualidades para o estudo e mapeamento de furtos de petróleo e combustíveis. Além da portabilidade, alia o não-preparo das amostras, a rapidez na obtenção dos dados e a confiabilidade da quimiometria (O'BRIEN *et al.*, 2012). Assim, torna-se uma boa ferramenta de campo.

A necessidade de tal aplicação se justifica pela grande quantidade de crimes. Em 2014, apenas um caso foi reportado em território nacional, porém esse crime escalonou em anos seguintes. Entre 2016 e 2018, a Transpetro informou que houve 560 casos de furto ou tentativa de furtos em dutos. Só em 2018 foram registrados 261 casos. Nos últimos quatro anos as perdas da estatal somaram em torno de R\$ 600 milhões, o equivalente a 42 milhões de litros de petróleo e derivados (REUTERS, 2019). Em âmbito mundial, são movimentados US\$ 133 bilhões no mercado ilegal (assalto a navios, roubos em dutos e de caminhões, adulteração de combustíveis) de petróleo e combustíveis, de acordo com uma pesquisa realizada por Yale (RAMALHO, 2019).

Em agosto de 2018 foi notícia a “Operação Bandeira Suja” na qual o Ministério Público (MP) apurava o roubo de combustíveis da refinaria Gabriel Passos (Regap) com ações ocorrendo em Minas Gerais, São Paulo e Bahia. Entre os integrantes da quadrilha presos estavam ex-funcionários da Petrobrás que tinham o conhecimento de como burlar o esquema de segurança da Petrobrás – que detecta a diminuição de pressão – e o robô que faz vistoria nos tubos (RICCI, 2018). Já em dezembro do mesmo ano, uma tentativa de furto no oleoduto da Petrobras em Magé, na Baixada Fluminense, provocou o vazamento de 60 mil litros de óleo no rio Estrela, que deságua na Baía de Guanabara (GLOBONEWS, 2018). Mais de 30 notícias sobre casos e operações diferentes foram encontradas nos jornais brasileiros.

Assim, a indústria busca estar na vanguarda de tecnologias de processamento e análise que sejam mais eficientes: mais rápidas, confiáveis e seguras. A espectroscopia na região do infravermelho próximo (NIR) está consolidada nesse sentido, contudo se baseia em equipamentos de bancada. (PANTOJA *et al.*, 2017) Entre os anos de 2016 a 2021 são encontrados 256 trabalhos sobre a utilização de espectroscopia na região NIR para análise de petróleo e combustíveis. No entanto,

nas últimas décadas, avanços instrumentais permitiram a miniaturização de espectrômetros, tornando-os ergonômicos, resistentes e adaptáveis para análise de diferentes tipos de amostras (CAPITÁN-VALVLEY; PALMA, 2011; CROCOMBE, 2018). Tamanha versatilidade facilita análises em campo. Em outros ramos industriais já se percebe essa atualização, (TEIXEIRA DOS SANTOS *et al.*, 2013; ZONTOV *et al.*, 2016) mas no campo do petróleo inicia-se os estudos somente agora (SANTOS *et al.*, 2021).

Equipamentos portáteis podem ser definidos como versões miniaturizadas de equipamentos laboratoriais que permitam serem transportados até o campo de trabalho e possibilitem a coleta dos dados. (CROCOMBE, 2018) O primeiro estudo considerando um sistema portátil se tratava de um sistema em sonda que se conectava a um notebook para quantificar contaminação por hidrocarbonetos em águas. (BUERCK *et al.*, 2001) Pela forma como descreveram, o equipamento poderia ser transportado para diferentes localidades – ou seja, portátil – porém, no estudo, os autores procederam com a instalação de um desvio numa tubulação de tratamento de água, no qual a sonda era inserida para adquirir os dados das águas contaminadas e, para a limpeza do sistema, o desvio era fechado e limpo com água deionizada – ou seja, um sistema em linha. Somente após uma década, foi observado o início de trabalhos com equipamentos portáteis comerciais, com o estudo de Cooper e colaboradores (2011) que estimou diferentes parâmetros de qualidade de mais de 800 amostras de combustíveis de aviação, com a coleta de dados ocorrendo apenas em dois dias, e adquiriam resultados similares aos obtidos pelos métodos convencionais. (COOPER *et al.*, 2011)

A **Tabela 1.1** traz os estudos refinando a pesquisa para a utilização de equipamentos portáteis em petróleos e derivados. A maior parte dos estudos utiliza o mesmo equipamento MicroNIR[™] Pro 1700 da Viavi Solutions com o método de análise multivariada por quadrados mínimos parciais (do inglês, *partial least squares*, PLS) sendo o mais aplicado. Desses estudos, apenas um se concentrou em petróleo bruto (SANTOS *et al.*, 2021), os outros se concentram em quantificar ou classificar amostras de combustíveis ou quantificar resíduos hidrocarbonetos em solos contaminados.

Tabela 1.1. Artigos utilizando equipamentos portáteis em petróleos e derivados entre 2018 – 2022

Referência	Matriz	Modelos	Equipamento
(SANTOS <i>et al.</i> , 2021)	Petróleo, gasolina, nafta, querosene, diesel	PLS-DA, PLS e SVM	MicroNIR™ Pro 1700, Viavi Solutions Inc
(TOSATO <i>et al.</i> , 2020)	Gasolina, diesel	PCA	MicroNIR™ Pro 1700, Viavi Solutions Inc
(CAVALCANTE DA SIVLA <i>et al.</i> , 2019)	Gasolina	PLS	MicroNIR™ Pro 1700, Viavi Solutions Inc
(CORREA PABÓN; SOUZA FILHO; OLIVEIRA, 2019)	Solos contaminados com petróleo	MESMA	ASD FieldSpec® 3 Hi-Res portable spectroradiometer
(DOUGLAS <i>et al.</i> , 2018)	Solos contaminados com petróleo	PLS, RF	ASD LabSpec2500® Vis–NIR spectrophotometer, MaLvern Panalytical
(CÁO <i>et al.</i> , 2018)	Gasolina, diesel	PLS	MicroNIR™ Pro 1700, Viavi Solutions Inc

Assim, percebe-se claramente o ganho da utilização da espectroscopia na região do infravermelho próximo na caracterização de derivados, logo há espaço para trabalhos utilizando petróleo. Visto que a projeção é que instrumentos portáteis sejam cada vez mais comuns (CAPITÁN-VALLVEY; PALMA, 2011; CROCOMBE, 2018). Esse trabalho desenvolverá metodologias analíticas baseadas na análise multivariada de dados obtidos em um espectrômetro portátil focada na caracterização de petróleos e adulteração de derivados.

2. OBJETIVOS

2.1. Objetivo geral

Avaliar petróleos e derivados pela técnica de espectroscopia na região do infravermelho próximo utilizando equipamento portátil para criação de modelos quimiométricos.

2.2. Objetivos Específicos

- Otimizar as condições instrumentais para análise de amostras de petróleo e derivados pela técnica de espectroscopia na região do infravermelho próximo utilizando equipamento portátil;
- Construir modelos quantitativos para estimar propriedades físico-químicas (densidade API, SARA, teor de enxofre, teor de nitrogênio básico) com espectros de infravermelho na região do próximo;
- Discriminar adulterantes em petróleos e combustíveis através de modelos qualitativos utilizando espectros de infravermelho na região do próximo.

3. FUNDAMENTAÇÃO TEÓRICA

3.1. Espectroscopia na região do infravermelho próximo

A luz sempre despertou o interesse dos humanos, afinal um dos grandes marcos na evolução da nossa espécie é a descoberta do fogo. Os primeiros filósofos já descreviam suas observações sobre a interação da luz ao passar por materiais transparentes como a água e objetos vítreos. Porém, o estudo da luz passou a ser sistematizado a partir dos anos 1600 com Willebrord Snell, Christiaan Huygens, Isaac Newton entre outros (BURNS; CIURCZAK, 2007). Assim, surge o ramo da espectroscopia.

Espectroscopia é o estudo da interação entre matéria e energia eletromagnética. A espectroscopia de infravermelho se refere, então, ao estudo do fenômeno da matéria interagindo com a radiação no intervalo de comprimento de onda de 780 nm a 1.000 μm – ou, em número de onda, 12.800 a 10 cm^{-1} . Essa região é dividida em três partes, as quais foram nomeadas de acordo com a sua proximidade do espectro visível: próximo (12.800 a 4.000 cm^{-1}), médio (4.000 a 400 cm^{-1}) e distante (400 a 10 cm^{-1}) (BURNS; CIURCZAK, 2007; SPRINGER, 2021).

Diferentes processos ocorrem quando há a interação entre a energia eletromagnética e a matéria, por exemplo, refração, difração, absorção, reflexão. Na espectroscopia do infravermelho observa-se o fenômeno da absorção de energia que leva a alteração da configuração molecular do analito. Na região NIR se observa bandas de transições eletrônicas, assim como também sobretons e bandas de combinações dos modos vibracionais.

De forma simplificada, utiliza-se o modelo do oscilador harmônico para compreender a vibração da ligação de uma molécula diatômica. Nesse modelo, um átomo (com massa m_1) se desloca em relação ao outro átomo (com massa m_2) com o qual faz ligação em uma determinada frequência, definida pela força da ligação e pelas massas dos átomos envolvidos, conforme a Lei de Hooke, demonstrada pela $\nu =$

$$\frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$$

Equação 3.1.

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$$

Equação 3.1

onde **k** é a constante da força de ligação e **μ** a massa reduzida:

$$\mu = \frac{m_1 m_2}{m_1 + m_2}$$

Equação 3.2

Dessa forma, a energia potencial do sistema para a posição dos átomos com a alteração na configuração molecular (deslocamento dos átomos envolvidos, **q**) é dada pela $E = \frac{1}{2} k q^2$

Equação 3.3. Assim, a energia entre diferentes níveis energéticos seria a mesma. E aí está a simplificação pois à medida que dois átomos se aproximam durante o movimento de vibração, as forças de repulsão de Coulomb ocorrerão entre os átomos. Assim, a energia potencial aumenta mais intensamente do que no modelo do oscilador harmônico. Além disso, para o movimento oposto, se a distância interatômica aumentar significativamente, haverá a dissociação da molécula, o que também acarreta numa distorção no modelo do oscilador harmônico.

$$E = \frac{1}{2} k q^2$$

Equação 3.3

Logo, o modelo que melhor descreve a absorção de energia pelas moléculas é o modelo do oscilador anarmônico, no qual a energia dos níveis vibracionais se dá pela $E = h\nu(v + 1/2) - x_m(v + 1/2)^2$

Equação 3.4, na qual **h** é a constante de Planck, **x_m** a constante de anarmonicidade, **v** o número quântico vibracional.

$$E = h\nu(v + 1/2) - x_m(v + 1/2)^2$$

Equação 3.4

4. REVISÃO

Petróleos através das lentes da espectroscopia na região do infravermelho próximo

4.1. Introdução

O estudo de petróleos se deu pela necessidade de compreender o que ocasionava características tão distintas aos óleos crus, pois se percebia o quanto o material de partida influenciava as qualidades dos derivados produzidos a partir destes (ASTM INTERNATIONAL, 1998). Os primeiros estudos químicos envolvendo o petróleo datam do ano de 1831 quando Robert Christison desenvolveu uma série de experimentos a fim de examinar a composição do petróleo originário de Yangon localizado em Mianmar entre Bangladesh e Tailândia (CHRISTISON, 1835). Vale ressaltar que técnicas espectroscópicas como a espectroscopia na região do infravermelho ainda não tinham sido desenvolvidas, portanto, os experimentos realizados variavam desde experimentos gravimétricos até experimentos de destilação com o intuito de separar e identificar os diferentes componentes presentes no petróleo. Em um desses estudos, uma das etapas de destilação produziu uma fração a qual foi denominada petrolina. Alguns meses depois da apresentação do trabalho, Christison porém teve ciência de um estudo desenvolvido pelo Dr. Reichenbach (1830) com uma fração que apresentava as mesmas propriedades da petrolina, porém com a denominação de parafina. Assim, Christison retratou-se afirmando que a fração descrita por ele correspondia à parafina, nome que perdura até hoje.

À medida que a exploração do petróleo foi se estendendo pelo crescente interesse econômico, surge a necessidade de criar procedimentos para prever e padronizar o comportamento desse material nas diversas etapas de exploração, isto é, extração, refino, transporte e armazenamento. Estes procedimentos baseiam-se no desenvolvimento de métodos analíticos e normas técnicas internacionais que estruturam o modelo industrial de produção (ASTM INTERNATIONAL, 1998; RICH, 2019). Desde então a produção e comércio de petróleos têm sido normalizados pela determinação de 26 propriedades físico-químicas (RIAZI, M., 2007; SPEIGHT, 2015).

Algumas dessas propriedades (densidade, curva de ponto de ebulição verdadeiro, teor de nitrogênio e teor de enxofre) influenciam diretamente na qualidade e no preço do óleo (MORO *et al.*, 2021). Contudo, essas informações a respeito das propriedades físico-químicas, embora denotem o comportamento dos petróleos durante o processo de produção, não são suficientes para responder sobre a composição química da matriz. Além disso, a determinação dessas propriedades físico-químicas é trabalhosa e muitas vezes demanda meses para que algum resultado possa ser encontrado.⁹ Alternativamente, os métodos espectroscópicos tem sido capazes de trazer informações a respeito do óleo bruto em nível molecular. Aliada a informação obtida, os ensaios de espectroscopia na região do infravermelho demandam pouco ou nenhum preparo de amostra, alta frequência analítica (centenas de espectros podem ser adquiridos por dia) e exigem um volume amostral bem menor (mililitros) quando comparados às técnicas clássicas (LOVATTI *et al.*, 2019).

As bandas de absorção na faixa do infravermelho são o registro das mudanças nos estados vibracionais das ligações químicas das moléculas presentes na amostra (SILVERSTEIN *et al.*, 2014). Assim, contribui com informações a respeito da composição pela relação entre as bandas de absorção e as diferentes funções orgânicas presentes. O petróleo é majoritariamente composto por hidrocarbonetos, mas suas características são fortemente influenciadas, também, por outras classes de compostos orgânicos presentes como ácidos, nitrogenados e sulfurados (SPEIGHT, 2015). O registro e criação de bancos de dados possibilitou a investigação da sinergia das variáveis e assim começou o trabalho da quimiometria em estimar as propriedades físico-químicas a partir dos dados espectroscópicos, uma das principais aplicações do NIR na avaliação do petróleo.¹² Pois o sinal analítico instrumental dificilmente dará informação quantitativa de interesse diretamente, dada à complexidade do petróleo. Dessa forma, seria difícil afirmar a presença de possíveis grupos orgânicos de forma direta, necessitando da estatística multivariada para extrair este tipo de informação (MORO *et al.*, 2021).

O presente trabalho busca abordar de forma crítica temas relevantes e aspectos modernos do estudo da espectroscopia na região do infravermelho próximo (NIR) aplicada, a princípio, na elucidação do petróleo e comparando às informações obtidas com a espectroscopia na região do infravermelho médio (MIR), focando no tratamento de dados e nos avanços instrumentais. Apesar dos fundamentos

vibracionais do infravermelho serem bem estabelecidos, a quimiometria e a miniaturização dos equipamentos avançaram de forma rápida em diversas áreas de aplicação do NIR, portanto a compilação desses resultados e das inovações tecnológicas são essenciais para pesquisadores da área de petróleo.

4.2. Espectroscopia de Infravermelho

O infravermelho próximo compreende a região do espectro eletromagnético entre 4.000 a 12.800 cm^{-1} (**Tabela 4.1**). A absorção de energia nessa faixa proporciona a observação dos sobretons e de bandas de combinações que ocorrem nos níveis energéticos vibracionais das ligações químicas. Para que seja possível registrar bandas na região do infravermelho próximo (NIR), o grupo de átomos da ligação química precisa apresentar transições fundamentais que absorvem radiação infravermelha na faixa entre 2.500 a 3.300 cm^{-1} e que apresentem alteração no momento dipolo (BURNS; CIURCZAK, 2007; SILVERSTEIN *et al.*, 2014).

A região do infravermelho médio (MIR) está compreendida entre 200 a 4.000 cm^{-1} e nela são observadas as transições fundamentais, bandas intensas e bem definidas que possibilitam a relação com os grupos funcionais de uma molécula e por isso tão utilizada na identificação de funções orgânicas (SILVERSTEIN *et al.*, 2014).

Os átomos presentes nas ligações de hidrogênio com o carbono (assim como as ligações entre hidrogênio e enxofre, oxigênio ou nitrogênio) quando absorvem a radiação infravermelho se deslocam no movimento de vibração produzindo uma alteração no momento dipolar da molécula, por isso é possível utilizar ambas regiões do espectro infravermelho para o estudo de compostos orgânicos, como no caso do petróleo (PANTOJA *et al.*, 2017; SILVERSTEIN *et al.*, 2014).

Já a região do infravermelho distante (FAR) está compreendida entre 200 a 10 cm^{-1} e é largamente aplicada na área medicinal devido a capacidade que essa região tem como uma modalidade terapêutica para tratamentos que envolvam radiação. Os comprimentos de onda dessa região são absorvidos pelo corpo como uma suave radiação que aquece a epiderme (camada mais externa da pele), esse aquecimento ocorre por meio dos modos de vibração e rotação das ligações químicas, principalmente da molécula de água, além de ressoar com as frequências das células

da pele (LARKIN, 2011). Por outro lado, o FAR não é largamente aplicado na estimativa de propriedades físico-químicas de petróleos devido à baixa interação que essa matriz tem nessa região espectral.

Tabela 4.1. Divisão do espectro infravermelho (BURNS; CIURCZAK, 2007)

Região	Comprimento de onda (nm)	Número de onda (cm ⁻¹)	Frequência (Hz)
NIR	780 a 2.500	12.800 a 4.000	3,8 x 10 ¹⁴ a 1,2 x 10 ¹⁴
MIR	2.500 a 50,000	4.000 a 200	1,2 x 10 ¹⁴ a 6,0 x 10 ¹²
FAR	50,000 a 100,000	200 a 10	6,0 x 10 ¹² a 3,0 x 10 ¹¹

4.3. Utilização da espectroscopia NIR em petróleo

Um estudo precursor para a área foi o de William Abney e Edward R. Festing (1881) que pelo interesse em fotografia conseguiram revelar as linhas espectrais de diferentes compostos em papel fotográfico e constataram que os grupos funcionais que apresentavam ligação com átomo de hidrogênio, com atenção especial para grupos hidrocarbonetos, tinham alta absorção na região entre 700 e 1.200 nm, enquanto outros grupos orgânicos não eram visíveis nessa faixa (ABNEY; FESTING, 1881; ADETUNJI; DRONSFIELD; MORRIS, 2015). Com base nestas observações, as pesquisas buscavam correlacionar as informações espectrais com grupos funcionais, focando em melhorias na instrumentação para ampliar a faixa de trabalho, qualidade dos resultados e, como consequência, surgiram os primeiros equipamentos que registravam espectros, com destaque para os trabalhos de William Coblentz nos anos 1900 (ADETUNJI; DRONSFIELD; MORRIS, 2015; PASQUINI, 2003).

Porém, com o avanço instrumental, identificou-se que a correlação com grupos funcionais era mais apropriada para a espectroscopia na região do infravermelho médio, visto que os espectros NIR apresentam bandas alargadas e por vezes sobrepostas, dessa forma houve um consequente declínio na utilização do infravermelho próximo (PASQUINI, 2018). Os estudos nas décadas seguintes passaram a se concentrar na vibração de estiramento da ligação C-H em torno de 3.000 nm, região MIR, e buscaram expandir o entendimento sobre petróleo pelo estudo de misturas de hidrocarbonetos (BROWN; LYNCH; AHMADJIAN, 1975). No

trabalho de Brattain e colaboradores há os esforços da época da tentativa de quantificar os componentes por modelos baseados na lei de Beer (BRATTAIN; RASMUSSEN; CRAVATH, 1943). Todavia, observa-se o problema nessa hipótese pelo trabalho de Hughes e Martin (1957) que descreve divergências dos modelos para a quantificação dos grupos metil, metileno, alquilbenzeno em diferentes misturas e principalmente em situações em que havia compostos poliaromáticos (ASTM, 1958).

O trabalho de Hughes e Martin traz ao fim uma discussão sobre a possibilidade da solução estar na espectroscopia NIR, mas com o adendo de que não havia instrumentação que desse o suporte necessário para essa investigação (ASTM, 1958). Contemporaneamente, há o trabalho de Owen Wheeler (1959) que revisou e catalogou todos os compostos orgânicos sobre os quais já haviam sido identificados os sobretons – ligações CH, OH, NH, C=O, SH, CCl, OOH (WHEELER, 1959). Nele se observa o pensamento da época em que a técnica de espectroscopia NIR teria aplicabilidade na identificação de compostos e quantificação baseada em modelos univariados. A linha de raciocínio levava em conta estudos recentes sobre corantes, pois a faixa revisada por Wheeler compreendia de 500 nm a 2.000 nm, região entre o espectro visível e infravermelho próximo (VIS-NIR). Por isso, a perspectiva apresentada por Hughes e Martin é tão interessante.

A espectroscopia NIR ressurgiu com o setor da agricultura nos EUA, em torno dos anos 1960, que procurava novos métodos analíticos para analisar amostras sólidas (NORRIS, K. H.; BUTLER, 1961). Com o estudo para determinar umidade em alimentos como grãos, verduras e carne, os autores provaram empiricamente que a determinação seria mais rápida e confiável. Em 1996, o trabalho de Karl Norris e Joe Hart abriu o campo da espectroscopia NIR quando quantificaram umidade em farinha pela combinação da informação de duas regiões espectrais (a banda do primeiro sobreton da ligação OH em 1.940 nm e a banda de combinações da ligação CH em 2.080 nm) criando um modelo que evidenciou a utilidade do NIR combinada à análise multivariada (NORRIS, Karl H.; HART, 1996).

A década de 90 marcou a chegada da espectroscopia NIR no campo industrial com a disponibilidade de equipamentos que se baseiam em reflectância difusa, com a popularização da quimiometria e o aumento do número de computadores pessoais (GELADI, 2003; PASQUINI, 2018). O primeiro estudo na área de petróleo foi o de Hidajat e Chong que construíram modelos PLS (do inglês, *Partial Least Squares*) para

estimar os parâmetros densidade e curva de Ponto de Ebulição Verdadeiro (PEV) e observaram nos gráficos de *scores* a possibilidade de classificar petróleos por sua origem (HIDAJAT; CHONG, 2000). No estudo de Hoeil Chung e colaboradores (1999) colocaram como perspectiva futura, a utilização de equipamentos portáteis (CHUNG; CHOI; KU, 1999). Assim chegamos ao século 21, em que o desenvolvimento de softwares de tratamento de dados tornou a quimiometria mais acessível, principalmente por facilitar a utilização dos usuários sem conhecimento refinado sobre matemática, estatística e programação, possibilitando a criação e integração às linhas de produção (GELADI, 2003).

4.4. Utilização da espectroscopia MIR em petróleo

Como descrito no tópico anterior, a região MIR encontrou inicialmente mais aplicações devido a resolução inerente a essa região, permitindo a identificação de grupos funcionais com mais facilidade.

A partir de 1927, o Instituto Americano de Petróleo (API) investiu em uma série de projetos para investigar a composição dos petróleos e seus derivados (RICH, 2019). O Projeto 44 tratou da tentativa de catalogar todos os possíveis constituintes por métodos espectroscópicos. Para a espectroscopia de infravermelho, a faixa investigada era de 14.000 a 2.000 nm e foram tabelados mais de 500 compostos entre parafinas, aromáticos e não-hidrocarbonetos (ROSSINI, 1951).

Já Williams (1936) utilizou o infravermelho médio para identificar polímeros na borracha, sendo ainda capaz de observar os efeitos da vulcanização desse material, o autor também destaca que o infravermelho apresenta algumas vantagens experimentais frente ao RAMAN por não haver necessidade de criar soluções de borracha em solventes como o dissulfeto de carbono ou o tetracloreto de carbono (WILLIAMS, 1936). Gordy e Stanford (1940) utilizaram o infravermelho médio para evidenciar ligações de hidrogênio a átomos de enxofre e nitrogênio (SH, NH e NH₂) (GORDY; STANFORD, 1940).

A região MIR ficou, então, conhecida pela capacidade de descrever a estrutura química de uma amostra, tanto na área do petróleo como em outras áreas da química. Porém, é na década de 1970 que o MIR se estabelece como uma técnica laboratorial

rotineira para identificação de compostos, devido o surgimento da técnica FTIR que proporcionou para os aparelhos, rapidez nas análises e aumento da relação sinal/ruído (SKOOG; HOLLER; CROUCH, 2018). Mais detalhes sobre a instrumentação da espectroscopia de infravermelho são discutidos na seção 5.

Em 1988, Solomon e colaboradores desenvolveram um trabalho para avaliar e caracterizar as frações mais pesadas do petróleo como os asfaltenos (visto que essas frações são fundamentais para prever o comportamento dos hidrocarbonetos durante o processo de refino) por meio da integração da análise termogravimétrica (TG) com a espectroscopia FTIR, que permite o estudo da evolução dos hidrocarbonetos mais pesados pela TG e o monitoramento dos produtos gasosos pelo FTIR (SOLOMON *et al.*, 1988).

Um ano depois, O'Grady e colaboradores (1989) utilizaram o FTIR para estudar a inclusão de petróleo em minerais. Um fluido incluído é uma bolha microscópica de líquido ou gás preso no interior de um mineral, essas inclusões trazem consigo informações a respeito das condições que existiram durante a formação daquele mineral (O'GRADY *et al.*, 1989). Os autores foram capazes de encontrar bandas em $2337 - 1300 \text{ cm}^{-1}$ que estavam presentes no espectro do fluido incluído, mas estavam ausentes no espectro do óleo bruto, essas bandas correspondem, respectivamente, ao CO_2 e CH_4 . Além disso, bandas acima de 3000 cm^{-1} no espectro incluído indicam a presença de aromáticos, já o espectro bruto não fornece nenhuma indicação da presença desses compostos. Os resultados espectroscópicos fornecem informação a respeito dos constituintes voláteis que, geralmente, são perdidos em amostras de óleo bruto, mas que quando capturados dentro de um mineral, permanecem no sistema.

A aplicação da quimiometria em estudos envolvendo o MIR só surgiu no início do século 20 e abriu novas portas e aplicações dessa técnica. Erade e colaboradores (2000) aplicaram modelos preditivos como o PLS e o PCR em dados espectrais de FTIR para avaliar o número de octanas de naftas oriundas do refino do petróleo (GARCÍA-MENCÍA *et al.*, 2000). O principal objetivo era avaliar a estabilidade dessas técnicas quimiométricas ao introduzir novas amostras no conjunto de calibração durante nove meses. Os resultados mostraram que os modelos foram capazes de manter a robustez, cujos valores de erro de predição foram um terço da reprodutibilidade das técnicas ASTM, indicando que a espectroscopia de infravermelho e a quimiometria podem ser futuras substituições das metodologias

ASTM, visto que essas metodologias demandam muito tempo e alto volume de amostra (FILGUEIRAS *et al.*, 2014; FOLLI *et al.*, 2020).

Já em 2010, observamos uma das primeiras aplicações de metodologias de seleção de variáveis com o trabalho de Parisotto e colaboradores, no qual os autores determinaram o número de ácidos totais (TAN) em resíduos de destilação do petróleo, ou seja, na parte mais pesada do óleo, a partir de um conjunto de dados de ATR-FTIR (PARISOTTO *et al.*, 2010). Os modelos de calibração foram construídos usando o iPLS, biPLS e o siPLS (NØRGAARD *et al.*, 2000) e a métrica avaliativa do modelo foi o RMSEP, porém a aplicação dessas técnicas não geraram vantagens significativas frente à utilização do espectro MIR inteiro. Ademais, os autores comentam que a espectroscopia e a quimiometria são capazes de determinar o número de acidez total e podem substituir metodologias mais trabalhosas, como a titulação potenciométrica, em análises rotineiras de resíduos de destilação do petróleo.

Em 2014, dois trabalhos se destacam ao utilizar o MIR, o primeiro é o trabalho de Filgueiras e colaboradores (2014) que determinou o grau API, a viscosidade cinemática e o teor de água em petróleo por meio da espectroscopia ATR-FTIR e calibração multivariada (FILGUEIRAS *et al.*, 2014). O trabalho também discute a respeito da comparação em modelos lineares e não lineares; para a determinação de grau API, o SVR (modelo não linear aplicado) mostrou-se mais preciso se comparado ao PLS (modelo linear aplicado). Já para a viscosidade cinemática e teor de água, houve a equivalência entre os modelos. Por fim, os autores concluem que a associação do ATR-FTIR e da calibração multivariada é eficiente na determinação de propriedades físico-químicas do petróleo, trazendo simplicidade no preparo de amostra além de determinar de forma simultânea as três propriedades com apenas um espectro, ao contrário do que é esperado para as metodologias ASTM.

O segundo trabalho de 2014 em destaque é o de Hemmateenejad e Dorostkar (2014) que utilizaram a quimiometria e o FTIR para identificar e classificar o petróleo conforme a sua origem geográfica (HEMMATEENEJAD; DOROSTKAR, 2014). Os autores discutem que a espectroscopia de infravermelho associada à quimiometria proporciona uma rápida e confiável metodologia analítica, principalmente por ser aplicada em amostras de qualquer estado físico: soluções, filmes finos, sólidos, gases etc. E, comparada a outras técnicas, as vantagens aumentam, pois o infravermelho não necessita de técnicas cromatográficas custosas, solventes perigosos e é não

destrói a amostra. Por fim, os autores concluem que a PCA não foi suficiente para discriminar as amostras, mas a aplicação da técnica ECVA (do inglês, *Extended Canonical Variates Analysis*) foi capaz de classificar todas as amostras corretamente, sendo que a região de 1281 – 1475 cm⁻¹ proporcionou modelos mais robustos, no que tange a classificação das amostras, se comparado ao uso do espectro inteiro.

Por fim, Guzmán-Osorio e colaboradores (2021) propuseram uma metodologia a fim de classificar o petróleo conforme a sua origem e integridade utilizando o FTIR como método espectroscópico e a quimiometria como forma de análise dos dados (GUZMÁN-OSORIO *et al.*, 2021). Os autores utilizaram a análise hierárquica (HCA) para agrupar amostras conforme os grupos funcionais, em seguida aplica-se uma análise discriminante de conglomerados, afim de encontrar classificações canônicas das funções, que poderia gerar um método de classificação mais geral. A metodologia proposta pelos autores associada ao FTIR foi capaz de realizar quatro tipos de classificação de origem/integridade do óleo: marinho, não marinho, degradados e transformados termicamente, cuja precisão foi de 100%.

De forma a concluir, a espectroscopia de infravermelho associada a calibração multivariada é uma metodologia eficiente na discriminação e identificação do petróleo devido a região média e suas características citadas nos tópicos anteriores, mas também na quantificação de propriedades físico-químicas devido a região média e próxima como citado no tópico anterior. Além disso, o infravermelho requer pouco ou nenhum método de preparação da amostra, é rápida se comparada com técnicas convencionais (como as ASTM) para determinar a origem do óleo e pouca quantidade de amostra é requerida.

4.5. Comparação entre o MIR e o NIR

De forma a comparar numericamente os trabalhos realizados utilizando a região do MIR e do NIR, a **Figura 4.1** traz o gráfico da produção científica anual a partir do início do século 21. Nela observamos o crescimento da produção de trabalhos utilizando o NIR e o MIR na área de petróleo. Essa figura faz parte de um estudo bibliométrico que será apresentado durante o trabalho. Esse estudo foi construído conforme a metodologia de Aria e Cuccurullo (2017) utilizando o software R versão

4.0,5 (R Core Team, 2021) com as palavras-chave “NIR”, “FTIR” e “crude oil” (ARIA; CUCCURULLO, 2017; R CORE TEAM, 2021).

Estudos sobre a estimativa das propriedades de petróleos, tais como teor das classes carbônicas, densidade e PEV, expandindo o trabalho de Hidajat e Chong (2000), por utilizar variações no modelo PLS para encontrar as áreas espectrais que apresentavam maior importância, ganharam espaço conforme a quimiometria foi sendo expandida dentro da área de petróleo (HIDAJAT; CHONG, 2000; LAXALDE *et al.*, 2011, 2014; PASQUINI; BUENO, 2007).

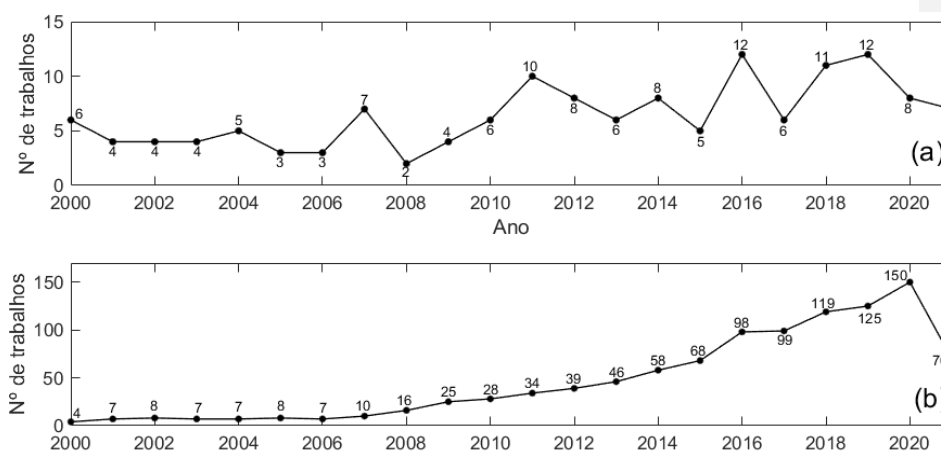


Figura 4.1. a) Produção científica anual entre 2000 a 2021 com os termos de entrada “NIR” e “crude oil” com um total de 141 artigos. b) Produção científica anual entre 2000

Por meio das **Figura 4.1.a** e **Figura 4.1.b**, podemos comparar a produção científica das técnicas NIR e MIR (Para a região MIR utilizamos a sigla FTIR que corresponde ao equipamento mais utilizado, o espectrômetro infravermelho com transformada de Fourier (em inglês, *Fourier Transform Infrared*), mas que se popularizou como sinônimo da técnica e mais resultados são encontrados utilizando esse termo.) aplicadas em petróleo. Notamos que o MIR (**Figura 4.1.b**) teve um aumento exponencial a partir de 2006, chegando a 150 artigos em 2020, Este crescimento pode ser devido a popularização do computador sendo responsável pela redução do custo no tratamento dos dados através da quimiometria (GELADI, 2003). Por outro lado, para a espectroscopia NIR, há poucos artigos publicados no mesmo

período de tempo (141) e não há uma tendência clara, ou seja, parece haver uma variação aleatória da quantidade de artigos publicados anualmente. Se levarmos em conta que o MIR apresenta transições fundamentais, as bandas vibracionais estão relacionadas com algum grupo funcional e, portanto, permitem a identificação da mistura. Por outro lado, no NIR as absorções são devidos a combinações de sobretons, que não auxiliam na identificação da mistura, mas permitem a quantificação, em particular se associada a análise multivariada (MORO *et al.*, 2021). Essas aplicações do MIR e do NIR ficam mais claras se observarmos a **Figura 4.2** que trazem quais as palavras-chave mais utilizadas pelos autores em conjunto com as palavras-chave “NIR e crude oil” e “FTIR e crude oil”, respectivamente.

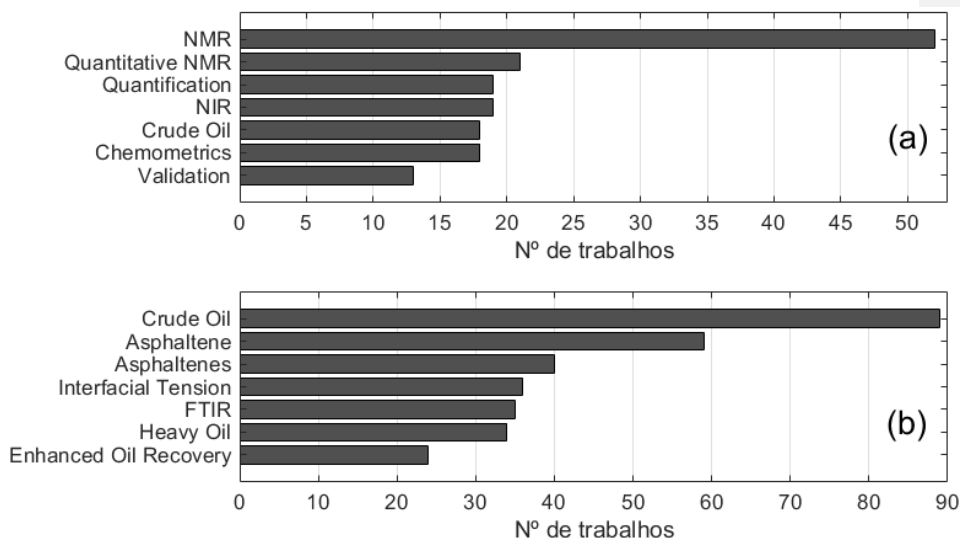


Figura 4.2. a) Palavras-chave mais frequentes de acordo com os termos de entrada “NIR” e “crude oil”.
b) Palavras-chave mais frequentes de acordo com os termos de entrada “FTIR” e “crude oil”.

Pela **Figura 4.2.a** a afirmação dada no parágrafo anterior a respeito do NIR é corroborada, pois as palavras-chave que o acompanha nos trabalhos são *NMR*, *chemometrics*, *validation*, *quantification* etc, além disso, notamos uma larga comparação com a técnica de Ressonância Magnética Nuclear já que essa técnica também é conhecida por ser largamente aplicada na quantificação das propriedades físico-químicas do petróleo (BALABIN; LOMAKINA; SAFIEVA, 2011; BALABIN;

SAFIEVA; LOMAKINA, 2010). Por outro lado, para o MIR, **Figura 4.2.b**, notamos palavras-chave que trazem a ideia de identificação composicional como *Asphaltene*, *adsorption*, *biosurfactant*, etc.

Ademais, Chung e colaboradores (1999) previram que o campo dos equipamentos portáteis estava se iniciando (CAPITÁN-VALLVEY; PALMA, 2011; CHUNG; CHOI; KU, 1999; CROCOMBE, 2018). E nesse ponto os equipamentos de espectroscopia NIR superaram os equipamentos de espectroscopia MIR, não por apresentarem uma melhor resolução. Nesse quesito, os equipamentos de espectroscopia de infravermelho com transformada de Fourier ainda apresentam melhores resultados, dado seu maior aproveitamento da potência luminosa e consequentemente maior precisão (FRIEDRICH *et al.*, 2014). Por sua vez, a robustez é mais satisfatória na instrumentação portátil que favoreceu a construção de espectrômetros NIR (CROCOMBE, 2018).

O primeiro estudo considerando um sistema mais próximo do portátil foi o de Buerck e colaboradores (2001) que construíram um sistema em sonda que se conectava a um notebook para quantificar contaminação por hidrocarbonetos em águas (BUERCK *et al.*, 2001). Pela forma como descreveram, realmente o equipamento poderia ser transportado para diferentes localidades – ou seja, portátil – porém, no estudo, os autores procederam com a instalação de um desvio numa tubulação de tratamento de água, no qual a sonda era inserida para adquirir os dados das águas contaminadas e, para a limpeza do sistema, o desvio era fechado e limpo com água deionizada, o que caracteriza um sistema em linha. Somente após uma década, foi observado o início de trabalhos com equipamentos portáteis comerciais, com o estudo de Cooper e colaboradores (2011) que estimou diferentes parâmetros de qualidade de mais de 800 amostras de combustíveis de aviação, com a coleta de dados ocorrendo apenas em dois dias, e adquiriam resultados similares aos obtidos pelos métodos convencionais (COOPER *et al.*, 2011). Por fim, Santos e colaboradores (2021) utilizaram um espectrômetro NIR portátil para classificar e quantificar óleo cru e combustíveis. O modelo de classificação desenvolvido foi capaz de discriminar o óleo cru de óleos de motor usados com sensibilidade, especificidade e precisão acima de 94%, já o modelo de regressão foi capaz de quantificar uma mistura ternária contendo óleos leves e pesados e óleo de motor com uma precisão de 4,8 % m/m (SANTOS *et al.*, 2021).

Por fim, podemos destacar outros fatores da bibliometria como as **Figura 4.3** e **Figura 4.4** que trazem informação a respeito das fontes mais relevantes e da produção científica mundial.

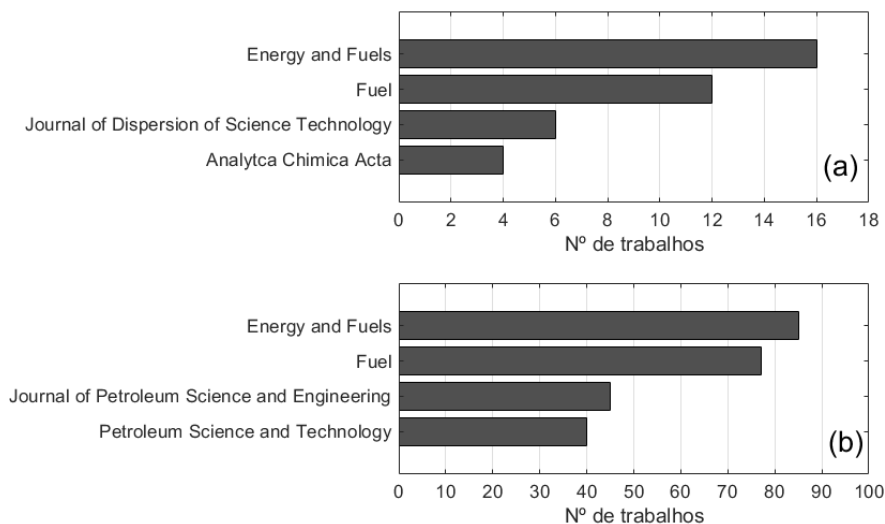


Figura 4.3.a) Fontes mais relevantes de acordo com os termos de entrada “NIR” e “crude oil”. b) Fontes mais relevantes de acordo com os termos de entrada “FTIR” e “crude oil”.

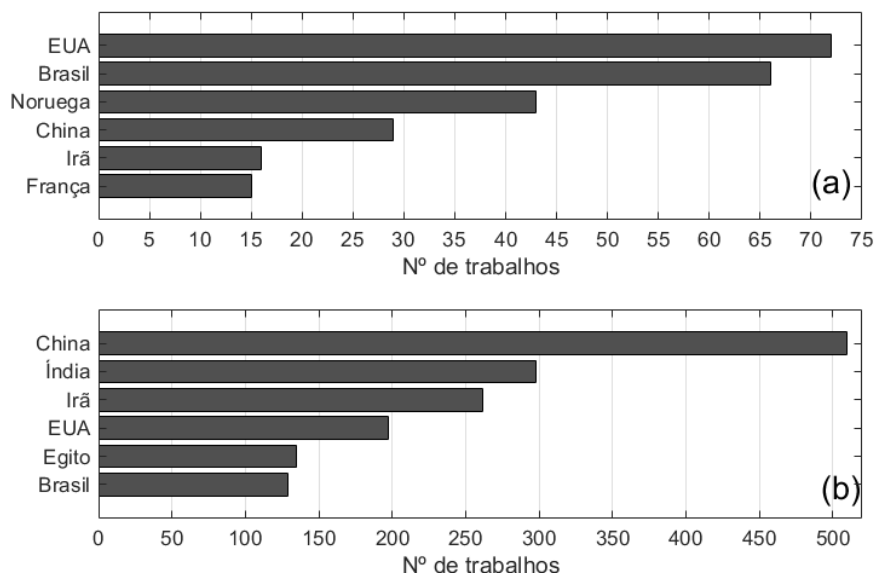


Figura 4.4. a) Produção científica dos países de acordo com os termos de entrada "NIR" e "crude oil".
b) Produção científica dos países de acordo com os termos de entrada "FTIR" e "crude oil".

Pelas **Figura 4.3.a** e **Figura 4.3.b**, notamos que a *Energy e Fuels* e a *Fuel* apresentam a maior quantidade de trabalhos publicados tanto no MIR quanto no NIR. E pelas **Figura 4.4.a** e **Figura 4.4.b**, notamos que os países com maiores publicações na espectroscopia de infravermelho aplicado a petróleo são a China, o Brasil, os EUA e o Irã por apresentarem grandes quantidades de trabalhos publicados para o NIR e para o MIR.

4.6. Quimiometria

A quimiometria surge devido a necessidade de se estudar as técnicas químicas (espectroscopia, cromatografia e espectrometria) que traziam consigo uma maior quantidade de informação e complexidade nas suas medições. Esse estudo quimiométrico é feito por meio da estatística multivariada, visto que a estatística univariada se mostra ineficiente e muito trabalhosa, quando estamos tratando dados de centenas ou até milhares de amostras (FERREIRA, 2015).

Na revisão de Moro e colaboradores (2021) estão tabeladas todas propriedades físico-químicas de petróleos que já foram modeladas com dados NIR e MIR. Os modelos quimiométricos mais aplicados para petróleo são modelos de regressão (MORO *et al.*, 2021).

4.6.1. Modelos de Regressão

Para esses modelos, o objetivo é encontrar a princípio uma reta que seja capaz de descrever e prever o comportamento do petróleo conforme a propriedade físico-química avaliada. Inicialmente os modelos aplicados em petróleo foram modelos lineares como o PLS, cujos resultados se mostraram promissores principalmente porque essa técnica é de rápida e fácil aplicação. Andersson (2009) desenvolveu um

trabalho afim de comparar nove algoritmos diferentes do PLS, nos próximos parágrafos discutiremos uma forma geral de calculá-lo (ANDERSSON, 2009).

Partirmos da forma geral de um modelo linear:

$$Y = Xb + e \quad \text{Equação 4.1}$$

É comum em dados espectrais haver covariância entre as variáveis, logo é uma boa ideia escrever a matriz X em função da seguinte combinação linear:

$$T = XV \quad \text{Equação 4.2}$$

$$V = [v_1 \ v_2 \ \dots \ v_A] \quad \text{Equação 4.3}$$

Onde, os vetores v_i representam os pesos e T representa os escores. T pode ser utilizado na calibração do vetor y :

$$y = Tq^T + e \quad \text{Equação 4.4}$$

Onde, q é o vetor de pesos para o vetor y . Se combinarmos a $Y = Xb + e$

XV

$$y = Tq^T + e$$

4.4, obtemos:

Equação 4.1 $T =$

Equação 4.2

Equação

$$y = XVq^T + e \quad \text{Equação 4.5}$$

$$b = Vq^T \quad \text{Equação 4.6}$$

Os vetores residuais e das equações $Y = Xb + e$

$$\text{Equação 4.1, Equação 4.4 } ey = XVq^T + e$$

Equação 4.5 são equivalentes. Dessa forma, ao invés de resolvermos a $Y = Xb + e$

Equação 4.1 para b , isto é, encontrar o coeficiente angular da reta que descreve a relação entre a matriz X e a matriz Y , podemos resolver a **Equação 4.5** para q , visto que um modelo PLS busca minimizar $e^T e$ dada as características dos vetores v_i . Para um modelo contendo A fatores, ou seja, A vetores v , a $T = XV$

Equação 4.2 seria uma matriz de rank máximo $I \times A$ e q um vetor $1 \times A$ e como o sistema é sobredeterminado, isto é, $A < I$, a solução é feita determinando o valor de q , por outro lado, se resolvêssemos a $Y = Xb + e$

Equação 4.1 direta, deveríamos resolver um sistema indeterminado que possui infinitas soluções. Logo, a partir da abordagem de resolver para q , o problema se torna a especificação de V .

Há várias formas de resolver a **Equação 4.5**, a primeira é caso $V^T X^T X V$ tem uma estrutura relativamente simples, podemos simplesmente utilizar a sua inversa, ou seja, a inversa pode ser calculada sem problemas de arredondamento:

$$q^T = (V^T X^T X V)^{-1} V^T X^T y \quad \text{Equação 4.7}$$

Outra abordagem para resolver a **Equação 4.5** é utilizar a pseudoinversa, que é definida da seguinte forma:

$$A^+ = (A^T A)^{-1} A^T \quad \text{Equação 4.8}$$

Onde, o sobrescrito $+$ indica a pseudoinversa da matriz em questão. Portanto, a $q^T = (V^T X^T X V)^{-1} V^T X^T y$ **Equação 4.7** pode ser descrita da seguinte forma se considerarmos $A = XV$:

$$q^T = (XV)^+ y \quad \text{Equação 4.9}$$

A pseudoinversa pode facilmente ser calculada por meio de pacotes de *softwares*, que geralmente é construída com uma função interna padrão desse *software*. Dessa forma, podemos generalizar as equações para fazer o cálculo de regressão como:

$$b = V(V^T X^T X V)^{-1} V^T X^T y \quad \text{Equação 4.10}$$

$$b = V(XV)^+ y \quad \text{Equação 4.11}$$

Além disso, os valores previstos podem ser calculados da seguinte forma:

$$\hat{y} = (XV)(XV)^+ y \quad \text{Equação 4.12}$$

Um algoritmo que é muito utilizado para cálculos do PLS é o NIPALS (do inglês, *Nonlinear Iterativa Partial Least Squares*), as equações desse método estão descritas abaixo:

$$v = X^T y \quad \text{Equação 4.13}$$

$$w = v / \sqrt{v^T v} \quad \text{Equação 4.14}$$

$$t = Xw \quad \text{Equação 4.15}$$

$$p = X^T t / t^T t \quad \text{Equação 4.16}$$

$$q = t^T y / t^T t \quad \text{Equação 4.17}$$

$$b = W(p^T W)^{-1} q \quad \text{Equação 4.18}$$

Inicialmente, calculamos o vetor peso v por meio da $v = X^T y$ **Equação 4.13** de forma que esse vetor seja o coeficiente de regressão de y em X , em seguida normalizamos esse vetor peso (w), vide $w = v / \sqrt{v^T v}$

Equação 4.14. Depois calculamos, pela $t = Xw$

Equação 4.15, o vetor de

escore $\hat{w}$ de tal forma que esse vetor seja o coeficiente de regressão de $\hat{w}$ em X . Utilizando esse vetor t podemos calcular o vetor de pesos p pela $p = X^T t / t^T t$

Equação 4.16e o

coeficiente de regressão q pela $q = t^T y / t^T t$

Equação 4.17. Em seguida, podemos calcular o vetor de regressão b utilizando a $b = W(P^T W)^{-1} q$

Equação 4.18, note que W e P representam matrizes dos vetores w e p , respectivamente, visto que esse processo é repetido para n variáveis latentes.

Por fim, a importância desse algoritmo se dá devido a facilidade que há em calcular a inversa de $P^T W$, pois essa matriz é bidiagonal superior.

Wilt e colaboradores (1998) utilizaram o PLS associado com o MIR para estimar o teor de asfaltenos no petróleo (WILT; WELCH; RANKIN, 1998). Alguns anos depois, há a primeira aplicação do NIR associado a quimiometria para estimar propriedades físico-químicas do petróleo com o trabalho de Aske e colaboradores (2001) que estimou o teor de SARA (saturados, aromáticos, resinas e asfaltenos) em petróleo com o auxílio do MIR, do NIR e do PLS (ASKE; KALLEVIK; SJÖBLOM, 2001). Os erros de predição para o modelo construído a partir dos dados de NIR foram 2,8; 2,4; 1,4 e 1,0 % (m/m) para saturados aromáticos, resinas e asfaltenos, respectivamente. Esses valores estão na mesma faixa dos valores reportados pela determinação direta utilizando o HPLC, mostrando que o NIR é uma alternativa rápida e confiável para estimar o teor de SARA em petróleo.

Moro e colaboradores (2021) destacaram que a relação entre a propriedade físico-química e a resposta espectral permitiu que houvesse uma estimativa de diversas propriedades como a densidade na escala API (ou °API), mas por outro lado, propriedades físico-químicas com o teor de metais não são facilmente exploradas pela espectroscopia por não haver uma correlação direta ou indireta com a resposta espectral (BAIRD; OJA, 2016; MORO *et al.*, 2021). É interessante também que nessa revisão sistemática, os autores tabelam todos os artigos que utilizaram espectroscopia MIR e NIR em petróleos. Entre os 19 artigos presentes na tabela, apenas um trabalho não constrói modelos PLS (FALLA *et al.*, 2006). Para todos os outros artigos citados é utilizado esse algoritmo seja como metodologia principal ou para comparação com

o algoritmo alvo. Ou seja, confirma o algoritmo PLS como principal estratégia para estimativa de propriedades físico-químicas em petróleos.

Os métodos OPLS, iPLS, siPLS e CarsPLS também são modelos lineares aplicados em estudos de petróleo associado ao NIR, porém, além da construção do modelo em si, há inicialmente, um processo de seleção de variáveis que busca encontrar uma região ou quais as variáveis mais adequadas para aquela propriedade físico-química específica. Sendo assim, os modelos matemáticos ficam mais simples, mais rápidos de serem executados e mais robustos. Balabin e colaboradores (2011) fizeram uma comparação de 16 métodos distintos de seleção de variáveis (além de entrarem em detalhes sobre o funcionamento de cada técnica aplicada) em estudos envolvendo o NIR aplicado em dados de biodiesel, utilizando dois modelos lineares para comparar os resultados: PLS e MLR (do inglês, *Multiple Linear Regression*) (BALABIN; SMIRNOV, 2011). Os autores também compararam os resultados com outras técnicas espectroscópicas como a RMN, o Raman e o UV-Vis. Por fim, há, também, a comparação com outros tipos de combustíveis como a gasolina, o diesel e o bioetanol.

Laxalde e colaboradores (2011) utilizaram o algoritmo genético (do inglês, *Genetic Algorithm*, GA) para estimar o teor de SARA em petróleo associado a espectroscopia NIR (LAXALDE *et al.*, 2011). Os erros de predição foram, respectivamente, para saturados, aromáticos, resinas e asfaltenos: 1,51; 1,59; 0,77 e 1,26 % m/m. Para o teor de saturados, aromáticos e resinas, os resultados se mostraram superiores aos da literatura, apenas para asfaltenos que os resultados foram equiparáveis.

Conforme a área foi se estabelecendo e se desenvolvendo, a aplicação de modelos não-lineares ganhou espaço, principalmente o modelo SVR (do inglês, *Support Vector Regression*). Filgueiras e colaboradores (2014) comparou o PLS com o SVR na predição de propriedades físico-químicas do petróleo com o auxílio do MIR. Os resultados mostraram que um estudo não-linear desenvolveu melhores resultados na modelagem de °API, porém para viscosidade cinemática e teor de água, os resultados do modelo linear e não-linear foram equivalentes (FILGUEIRAS *et al.*, 2014).

Folli e colaboradores (2020) fizeram uma comparação entre o MIR, o NIR e a RMN com o auxílio da combinação do algoritmo ASA (do inglês, *Angular Search*

Algorithm), VIF (do inglês, *Variable Inflation Factor*) e do SVR para estimar número de acidez total (TAN), teor de nitrogênio básico (TNB) e teor de enxofre (SC) (FOLLI *et al.*, 2020). A validação do SVR foi feita a partir dos valores de erro quadrático médio de calibração (do inglês, *Root Mean Square Error of Calibration*, RMSEC). Os autores comentam que a eliminação de *outliers* (amostras que destoam da resposta média) melhorou a performance do SVR, além disso o modelo de seleção ASA-VIF foi capaz de selecionar 0,3% do espectro do MIR, 0,4% do espectro para a RMN de ^1H e 0,9% para o espectro de NIR. Frente aos resultados, os autores concluíram que para o teor de enxofre, todas as três técnicas (NIR, MIR e RMN) apresentaram resultados semelhantes, para o teor de nitrogênio básico, o NIR mostrou-se superior e, por fim, para o TAN, o MIR apresentou os melhores resultados.

4.6.2. Modelos de Classificação

Para estudos de classificação do petróleo conforme as propriedades físico-químicas, começamos a atingir uma defasagem de estudos; na literatura foi encontrado apenas um artigo que trata desse tipo de classificação: Mohammadi e colaboradores (2020) realizaram um estudo a respeito da classificação do óleo a partir do grau API, o modelo utilizado foi o PLS-DA e os resultados foram promissores atingindo 100% de precisão, sensibilidade e especificidade, porém os autores utilizaram o MIR (como discutido nos tópicos anteriores, o MIR apresenta transições fundamentais que favorece a identificação de grupos funcionais e posterior classificação do óleo), sendo assim o NIR ainda não apresenta aplicações diretas para a classificação do óleo (MOHAMMADI *et al.*, 2020).

O PLS-DA, modelo utilizado pelos autores, tem a mesma premissa que o modelo PLS: reduzir a informação química em variáveis latentes que maximizam a covariância entre informação espectral e a informação das propriedades físico-químicas, porém para a matriz y (ou vetor y caso a avaliação seja feita com uma propriedade de cada vez) os valores são discretos (0, 1, 2, etc) que descrevem a classes que aquela amostra se encontra (FERREIRA, 2015).

Essa divisão de classes pode ser feita por meio de probabilidades, mas nos estudos de petróleo é mais comum que essa divisão seja feita pelo próprio analista

seguindo os limites impostos pelas refinarias, por exemplo, para teor de enxofre, amostras contendo valores acima de 0,5 % (m/m) são considerados prejudiciais, pois podem causar problema como corrosão e/ou contaminação de catalisadores, portanto, há uma divisão clara entre duas classes: uma contendo valores abaixo de 0,5 (chamada de doce) e a outra contendo valores acima ou iguais a 0,5 (chamada de azedo) (MORO *et al.*, 2021). Lovatti e colaboradores (2019) e Lovatti e colaboradores (2020) utilizaram essa premissa para separar classes (LOVATTI *et al.*, 2020, 2019). No trabalho de 2019, os autores classificaram o petróleo conforme o ponto de fluidez máximo, essa divisão foi feita separando as amostras cujos valores eram menores ou iguais que $-9\text{ }^{\circ}\text{C}$ e valores maiores que $9\text{ }^{\circ}\text{C}$. Já no trabalho de 2020, os autores classificaram o petróleo conforme o TAN e essa divisão foi feita separando as amostras com valores menores ou iguais a $0,3\text{ mgKOH}\cdot\text{g}^{-1}$ e valores maiores que $0,3\text{ mgKOH}\cdot\text{g}^{-1}$. Enquanto no segundo trabalho, os autores utilizaram dois métodos de reconhecimento de padrões: a análise por componentes principais (PCA) e o k vizinhos próximos (do inglês, k Nearest Neighbors, kNN) para identificar e separar em grupos amostras similares (LOVATTI *et al.*, 2020). A PCA foi utilizada para diminuir as dimensões do conjunto de dados e o kNN foi utilizado para identificar amostras semelhantes a partir do espectro MIR. Os autores ainda conseguiram encontrar relações entre as propriedades físico-químicas a partir do plot duplo de *score* e pesos da PCA, ou seja, a partir da disposição das amostras, os autores encontraram relação diretas ou indiretas entre as propriedades físico-químicas. Por fim, a compatibilidade entre os dois modelos foi de 66,7% em média.

4.7. Instrumentação

A instrumentação da espectroscopia de infravermelho pode ser dividida em três grupos: espectrômetros dispersivos (também chamados de espectrofotômetros), espectrômetros com transformada de Fourier (FTIR) e fotômetros de filtro. Os dois primeiros são utilizados para a construção do espectro de infravermelho e posterior identificação qualitativa da amostra, já o terceiro é empregado em trabalhos quantitativos, porém não é largamente utilizado em petróleo (SKOOG *et al.*, 2014).

Do ponto de vista histórico, o infravermelho foi a primeira espectroscopia estrutural utilizada pelos químicos orgânicos. Nas décadas de 1930 e 1940, essa técnica era experimentalmente desafiadora e apresentava poucos trabalhos. O desenvolvimento de aparelhos IV disponíveis comercialmente teve início no ano de 1946 quando a American Cyanamid Stamford colaborou com uma empresa especializada em aparelhos óticos chamada Perkin-Palmer, o espectrômetro IV desenvolvido apresentava um prisma de distância focal curta (LARKIN, 2011).

Trabalhos iniciais que procuravam encontrar bandas referentes as ligações que caracterizavam uma função química como o de Gordy e Stanford (1940) beneficiaram o avanço da espectroscopia de infravermelho, principalmente por contribuírem para a construção de um quadro que descreve as bandas mais importantes e quais os intervalos de comprimento de onda que as bandas estão presentes (GORDY; STANFORD, 1940). Esse quadro é utilizado como base para o ensino da técnica, sendo encontrada em todos os livros de espectroscopia de infravermelho (LARKIN, 2011).

Porém, a explosão de trabalhos só aconteceu nas décadas de 1950 e 1960 quando os aparelhos se tornaram cada vez mais baratos e comercialmente disponíveis, principalmente a instrumentação que continha um feixe duplo (LARKIN, 2011). No próximo tópico haverá uma discussão mais detalhada a respeito da técnica, contemplando desde a instrumentação que compreende o NIR até a instrumentação que compreende o MIR (FTIR).

4.7.1. Instrumentos Dispersivos

Um espectroscópio é um instrumento ótico empregado na identificação visual de linhas atômicas. O aparelho consiste na presença de um monocromador e o comprimento de onda da radiação emitida é calculado a partir do ângulo formado pelo feixe incidido e o feixe dispersado. Já um fotômetro consiste numa fonte, num filtro, num transdutor elétrico e num detector que irá ler e processar o sinal recebido. Um espectrômetro é caracterizado como um instrumento capaz de fornecer informação a respeito da intensidade da radiação em função da frequência e/ou do comprimento de onda. Sendo assim, um espectrofotômetro é definido como um equipamento equipado

com uma ou mais fendas e um transdutor fotoelétrico que permite a determinação da razão entre dois feixes de luz em função do comprimento de onda e/ou frequência. Vale ressaltar que todos esses equipamentos utilizam dispersores ou filtros para isolar uma seção do espectro. Para o infravermelho próximo, os instrumentos são caracterizados como instrumentos dispersivos e de feixe duplo (LARKIN, 2011; SKOOG; HOLLER; CROUCH, 2018). A esquematização dessa técnica encontra-se na **Figura 4.5**.

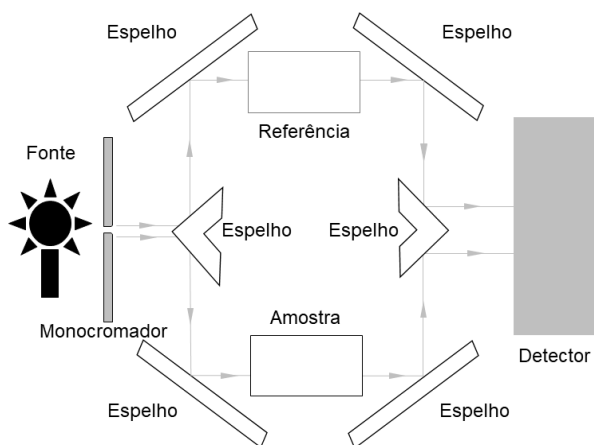


Figura 4.5. Diagrama de um espectrofotômetro de feixe duplo. A aplicação do feixe duplo favorece a detecção de forma simultânea da amostra em questão e da referência utilizada, portanto, qualquer erro do tipo aleatório presente no sistema é compensado.

A fonte desse instrumento é composta de lâmpadas capazes de emitir radiação na região do infravermelho (SKOOG *et al.*, 2014). Na **Tabela 4.2** está descrito as principais fontes contínuas utilizadas para o infravermelho. Essas fontes são construídas a partir de sólidos inertes que passarão por um processo de aquecimento, como a fonte do tipo Globar que ao sofrer um aquecimento a 1.500 °C por meio da eletricidade, emite radiação no infravermelho.

Tabela 4.2. Fontes utilizadas na espectroscopia de infravermelho (SKOOG *et al.*, 2014)

Fonte	Comprimento de	Região do infravermelho
-------	----------------	-------------------------

	onda (nm)	
Lâmpada de tungstênio/halogênio	240 a 2.500	Compreende o infravermelho próximo
Lâmpada de tungstênio	350 a 2.200	Compreende o infravermelho próximo
Fonte de Nernst	400 a 20,000	Compreende todas faixas do infravermelho
Fio de níquel-crômio	750 a 20,000	Compreende todas faixas do infravermelho
Globar	1.200 a 40,000	Compreende todas faixas do infravermelho

Como descrito na **Tabela 4.2**, a radiação é emitida em uma janela muito ampla, portanto os aparelhos trazem uma metodologia capaz de selecionar ou restringir a radiação que chega no analito. Essa metodologia é empregada por um monocromador ou filtro. Os monocromadores selecionam o comprimento de onda por meio de uma rede de difração ou um prisma (equipamentos mais antigos empregavam o uso do prisma), essa rede apresenta ranhuras para gerar diferentes reflexões da radiação e pode ser movida horizontalmente para manipular a seleção dos comprimentos de onda (SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

Já um filtro para região do infravermelho é composto de uma camada fina de um material dielétrico recoberto em ambos os lados com um filme metálico fino. As reflexões que ocorrem dentro do filtro podem gerar interferências construtivas ou destrutivas, selecionando ou eliminando diferentes comprimentos de onda, por isso o nome filtro de interferência é empregado. Por outro lado, os filtros, se comparados com os monocromadores, encontram uma aplicação limitada no infravermelho, pois é empregada em desde o ultravioleta até o infravermelho em 14 μm , compreendendo a região do próximo inteira e próximo da metade da região do médio. Portanto, essa metodologia encontra aplicação apenas para instrumentos de infravermelho na região do próximo (LARKIN, 2011; SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

Assim que a radiação emitida atravessa e interage com o analito, o que foi emitido ou absorvido é detectado por um detector, que é definido como um dispositivo que indica a ocorrência de algum fenômeno físico. Para os instrumentos modernos, o detector codifica a informação em sinais elétricos; o termo aplicado para esses detectores é transdutor. Essa informação pode ser de diferentes naturezas como: pH, massa, temperatura, intensidade da luz. Na **Tabela 4.3** está representado os

diferentes tipos de detectores que atuam na região do infravermelho (SKOOG *et al.*, 2014).

Tabela 4.3. Detectores utilizados na espectroscopia de infravermelho (SKOOG *et al.*, 2014)

Tipo	Comprimento de onda (nm)	Região do infravermelho
Detectores térmicos		
Termopares	600 a 20,000	Infravermelho próximo e parte do médio
Células pneumáticas	600 a 40,000	Infravermelho próximo e grande parte do médio
Células piroelétricas	1000 a 20,000	Parte do infravermelho próximo e médio

Em geral, para o infravermelho, são empregados os detectores térmicos, pois estes são capazes de trabalhar com mais facilidade na região do infravermelho. Os detectores térmicos apresentam uma superfície negra capaz de absorver a radiação infravermelha e aquecer, esse aquecimento é detectado pelo transdutor e transformado em sinal elétrico. A variação de temperatura é muito pequena, sendo assim o ambiente de trabalho deve ser controlado para que erros provenientes do ambiente sejam introduzidos ao sistema (LARKIN, 2011; SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

Por fim, para que o sinal proveniente do detector seja ampliado e manipulado, os instrumentos mais modernos trazem consigo processadores de sinais e dispositivos de leitura, que podem alterar o sinal, filtrá-lo etc. Os laboratórios que apresentam um espectrômetro na região do infravermelho, geralmente, contêm um computador acoplado ao sistema que, por meio do software fornecido pela empresa fabricante do espectrofotômetro, é capaz de realizar operações matemáticas como a diferenciação, integração etc (LARKIN, 2011; SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

4.7.2. Instrumentos FTIR

Até a década de 1970, os espectrômetros dispersivos eram os únicos instrumentos manufaturados que faziam medições na região do infravermelho. A princípio, os espectrômetros FTIR não possuíam bom custo/benefício, não eram instrumentos de bancada (inviabilizando o seu uso em laboratório) e requeriam ajustes do ponto de vista mecânico de forma frequente, portanto, a sua aplicação, inicialmente, era muito restrita a estudos nas quais as características do FTIR eram necessárias (alta velocidade, alta resolução, confiabilidade e razão sinal/ruído). Esses instrumentos só encontraram uma aplicação mais ampla a partir da década de 1990, cujos espectrômetros sofreram uma redução espacial, mais robustos e o preço dos modelos mais simples se tornou competitivo com instrumentos dispersivos (SKOOG; HOLLER; CROUCH, 2018).

Os instrumentos com transformada de Fourier não apresentam nenhum elemento dispersivo, logo todos os comprimentos de onda são medidos simultaneamente por meio de um interferômetro. Para a obtenção do espectro de uma amostra, primeiro é obtido um espectro de fundo (ou branco), normalmente é utilizado o solvente em questão, água do ambiente ou dióxido de carbono. Em seguida é obtido o espectro da amostra, a razão entre o espectro de fundo e o espectro amostral é calculada e a absorbância (ou transmitância) *versus* comprimento de onda (é também utilizado o número de onda, definido pelo inverso do comprimento de onda) é reportada. Para evitar a absorção de umidade e CO₂, alguns instrumentos purgam o ambiente do aparelho com gases inertes ou ar seco livre de CO₂ (SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

O FTIR apresenta diversas vantagens inerentes ao instrumento. A primeira é a chamada vantagem de Jacquinot, que é derivada do entendimento que há pouca aparelhagem compondo um instrumento FTIR, portanto, a intensidade da radiação que chega no detector é alta se comparada a aparelhos de dispersão, gerando uma relação sinal/ruído elevada. A segunda é a capacidade de resolução e reprodutibilidade para diferentes comprimentos de onda do equipamento, que encontra aplicações em sistemas cuja sobreposição de bandas dificultaria a identificação das funções presentes na amostra – comum para petróleo (SKOOG; HOLLER; CROUCH, 2018). De forma a exemplificar essa alta resolução, Thorne (1991) apresenta um espectro de emissão do aço e, em um intervalo de apenas 299,85 – 300,75 nm, é possível observar linhas bem separadas para três elementos

distintos: Cr, Ni e Fe, sendo ainda possível distinguir os elementos conforme o número de oxidação *i.e.* Fe(I), Fe(II), Cr(I), Cr(II), Ni(I) e Ni(II) (THORNE, 1991). A terceira, e última, vantagem ocorre devido ao alcance simultâneo das medições de radiação no detector, sendo assim é possível obter um espectro completo em 1 segundo ou menos (SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

Na **Figura 4.6** está expresso um esquema de funcionamento de um espectrômetro com interferômetro de Michelson que, aliado a construção matemática da transformada de Fourier, permitiu o surgimento dos instrumentos FTIR. O interferômetro produz um sinal que varia de forma senoidal no detector, conforme o espelho se move a partir de uma luz monocromática e pela combinação de diversos comprimentos de onda, produz um sinal que pode ser decomposto em uma soma de senoides cujas frequências são conhecidas (SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

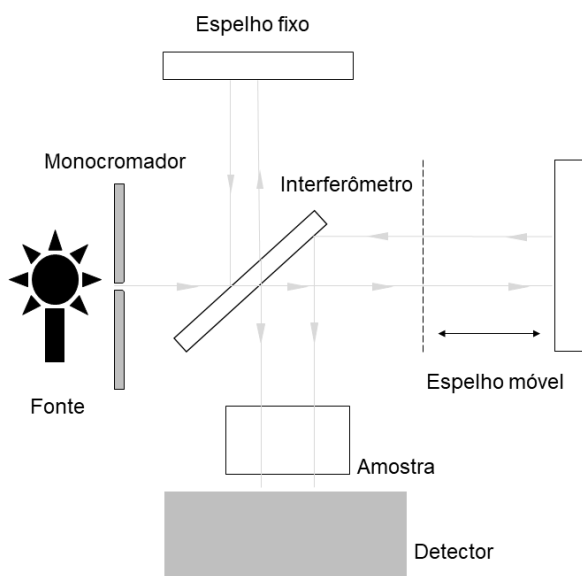


Figura 4.6. Diagrama de um espectrômetro com Transformada de Fourier. Um feixe de uma fonte de luz é dividido em dois feixes pelo interferômetro. Esses dois feixes percorrem em caminhos distintos até se encontrarem na mesma região espacial do detector, formando um padrão de interferência. O espelho móvel é responsável pela modulação do sinal óptico a partir do seu deslocamento horizontal que altera o padrão de interferência

Comparando o funcionamento dos aparelhos dispersivos (**Figura 4.5**) e aparelhos FTIR (**Figura 4.6**) fica clara a vantagem de Jacquinot que os instrumentos FTIR apresenta. Vemos a existência de apenas uma fonte e um detector, os outros componentes são espelhos, diferente de um aparelho dispersivo que apresenta diversos componentes.

Por outro lado, como citado anteriormente, apesar da existência do interferômetro já datar quase um século e a metodologia matemática datar de dois séculos, a aparelhagem só ganhou espaço em estudos rotineiros quando alguns avanços aconteceram, tais como: o baixo custo de computadores de alta velocidade e métodos computacionais apropriados para lidar com o enorme volume de cálculos que ocorrem dados espectrais brutos. Por fim, pacotes de programas como o MATLAB e até mesmo o Excel favorecem a análise de dados provenientes dessa técnica (SKOOG *et al.*, 2014; SKOOG; HOLLER; CROUCH, 2018).

Por fim, vale ressaltar que os instrumentos com transformada de Fourier não são largamente aplicados para outras regiões espectrais como a região do UV, do visível e do infravermelho próximo (NIR), pois as limitações sinal/ruído presente nessas regiões não são oriundas do ruído do detector, mas sim devido ao ruído de disparo do feixe e da cintilação da fonte. Em contrapartida ao ruído do detector, as magnitudes do ruído de disparo de cintilação aumentam conforme a intensidade da radiação aumenta (NIKI *et al.*, 1978).

A espectroscopia de infravermelho com transformada de Fourier encontrou as suas primeiras aplicações, na química, como um meio para entender a cinética e o mecanismo de reação e elucidação estrutural (HOLLE-MORITZ; STACH; HOLE-MORITZ, 1980; NIKI *et al.*, 1978; SENICH; MACKNIGHT, 1980).

Ademais, a portabilidade desses instrumentos é de bastante interesse para a análise rápida e de baixo custo. Pasquini e colaboradores (2018) trazem uma tabela contendo 18 instrumentos manufaturados por diversas empresas explorando as vantagens de cada um deles (PASQUINI, 2018).

4.8. Conclusão

A utilização da espectroscopia na região do infravermelho próximo se consolidou na estimativa de propriedades físico-químicas de petróleos. A associação da quimiometria por meio do PLS e do SVR forneceram resultados satisfatórios para estimar grau API, ponto de fluidez, teor de enxofre, TAN e viscosidade. Métodos de seleção de variáveis auxiliam nesses estudos, pois excluem variáveis não relevantes ou selecionam variáveis que são relevantes, facilitando a captura da região espectral para estudar a propriedade de interesse. Além disso, como o NIR não apresenta transições fundamentais e, por consequência, muita sobreposição de bandas, essa técnica não é indicada para estudos qualitativos do petróleo. Por fim, a projeção é que instrumentos portáteis sejam cada vez mais comuns, pois conforme fica evidente na história, os estudos começam com a análise de derivados e esses estão delimitando seu espaço com modelos competitivos na predição de diferentes propriedades físico-químicas em combustíveis de aviação, diesel, gasolina, e já temos sinais de que será possível expandir esse campo para petróleos.

5. APLICAÇÃO 1

Comparação entre equipamento portátil e de bancada na classificação de petróleos

5.1 Introdução

Petróleos são misturas complexas de hidrocarbonetos e compostos orgânicos com heteroátomos de diferentes pesos moleculares que fornecem heterogeneidade ao produto que varia de acordo com a sua natureza e origem (PASQUINI; BUENO, 2007; RIAZI, M. R., 2005). O monitoramento do petróleo bruto é essencial para que o processamento da matéria ao longo de toda cadeia ocorra com alto grau de qualidade e segurança. A indústria busca estar na vanguarda de tecnologias de processamento e análise que sejam mais eficientes: mais rápidas, confiáveis e seguras.

A espectroscopia de infravermelho próximo (NIR) apresenta tais vantagens. Combina rapidez nas análises, com espectros sendo coletados em questão de segundos e pouco ou nenhum preparo das amostras, baixo custo de manutenção de equipamentos, mínima produção de resíduos e confiabilidade dos resultados pela combinação com a quimiometria (PANTOJA *et al.*, 2017; PASQUINI, 2018). E o avanço dos diodos emissores de luz permitiu a miniaturização dos equipamentos conferindo portabilidade e resistência aos equipamentos (CAPITÁN-VALLVEY; PALMA, 2011; CROCOMBE, 2018). Além do mais, equipamentos portáteis apresentam a vantagem de uma variedade de dispositivos para leitura, dependendo da característica da amostra (TEIXEIRA DOS SANTOS *et al.*, 2013; ZONTOV *et al.*, 2016). Essa versatilidade facilita análises em campo e o uso da tecnologia nos últimos anos está em ascensão (CROCOMBE, 2018).

Os espectrofotômetros de bancada apresentam maior resolução espectral e por isso geralmente ganham em avaliações de desempenho quando comparados aos equipamentos portáteis (CROCOMBE, 2018; MALEGORI *et al.*, 2017). Logo, não pretendem substituir tais equipamentos, mas serem complementares. Sabendo de suas diferenças é importante a construção e otimização das metodologias adaptando para os equipamentos portáteis. Assim, encontramos estudos como de Sun *et al.*

(2016) (SUN *et al.*, 2016) que utilizou seis espectrômetros NIR portáteis para avaliar o poder de classificação de tais instrumentos para diferentes medicamentos de diferentes lotes e marcas utilizando diferentes modelos de classificação lineares e não lineares e comprovando não só o grau de exatidão de tais instrumentos para fármacos, mas a facilidade na transferência de modelos entre diferentes aparelhos. Enquanto, Hoffmann *et al.* (2016) comprovou a eficiência na transferência de modelos obtidos com dados de equipamentos de bancada para um equipamento portátil para a quantificação de solventes em misturas ternárias, assim como para a classificação de diferentes polímeros. Já Malegori *et al.* (2017) compararam a performance de um equipamento portátil e um equipamento de bancada para análise de dois parâmetros de acidez de acerolas, um tipo de fruta nativa do Brasil, utilizando os frutos inteiros ao longo de 14 dias de amadurecimento. Os autores obtiveram resultados considerados satisfatórios com o uso do algoritmo não linear Support Vector Machine com R^2 maior que 0,7 e comprovaram que não havia diferenças estatísticas significativas entre os dois instrumentos (MALEGORI *et al.*, 2017).

Nesse artigo mostramos o desenvolvimento da metodologia para análise de petróleo em um espectrômetro portátil e comparamos com os dados obtidos em equipamento de bancada, um espectrômetro de infravermelho com transformada de Fourier (FT-IR). Utilizamos 93 amostras de petróleo de três origens distintas (Petróleo A, B e C). Os dados foram discutidos com base na PCA e nos modelos PLS-DA.

5.2. Metodologia

5.2.1. Amostras

As amostras foram obtidas com empresas do setor de energia e se referem a amostras de petróleo da bacia sedimentar brasileira. Foram utilizadas 93 amostras de petróleo de três origens distintas (30, 33 e 30 amostras de petróleo das refinarias A, B e C, respectivamente). Todo o procedimento de amostragem e manuseamento ocorreu de acordo com a ASTM D5854 (ASTM INTERNATIONAL, 2019). A amostragem ocorreu em triplicata.

5.2.2. Ensaios de infravermelho próximo

5.2.2.a. Equipamento portátil

O instrumento utilizado foi um espectrômetro modelo MicroNIR™ Pro 1700 (**Figura 5.1.a**) com softwares MicroNIR™ Pro versão 2.1 e 3.0 da Viavi Solutions Inc. (FRIEDRICH *et al.*, 2014).

Este equipamento dispõe de diferentes acessórios para se adaptar às amostras, de forma que estas não sejam destruídas para a aquisição dos espectros, e para facilitar a utilização do equipamento em campo (CORREIA *et al.*, 2018). Em nosso estudo, avaliamos a utilização dos dois acessórios, *Droplet* e *Vial Holder*, que permitem análise no modo reflectância para líquidos. Estes acessórios são acoplados sobre o MicroNIR (**Figura 5.1.b,c**).

O acessório *Droplet* se refere a um acessório prateado que permite a amostragem de uma pequena alíquota (aproximadamente 0,2 mL) (**Figura 5.1.b**), enquanto o *Vial Holder* corresponde a um acessório metálico coberto por verniz preto no qual são encaixados *vials* de 4 mL (**Figura 5.1.c**).



Figura 5.1. a) Espectrômetro MicroNIR e acessórios acoplados sobre ele: b) *Droplet* e c) *Vial Holder* (com frasco de 4 mL inserido).

Para a aquisição de espectros, primeiramente é necessário a aquisição do sinal na ausência de radiação (corrente escura) e espectro da referência (material que reflita 99.9% da luz). No início do projeto, o software que tínhamos a nossa disposição era o MicroNIR™ Pro 2.1. Nele a corrente escura é feita com as lâmpadas de tungstênio acesas. Assim, executamos os procedimentos aqui denominados Gota 1 e *Vial*. Depois, porém, nos foi disponibilizado a versão 3.0 do software. Assim, adaptamos as configurações com o acessório Gota criando o procedimento Gota 2.

A forma de aquisição dos espectros para os acessórios *Droplet* e *Vial Holder* estão ilustradas nas **Figura 5.2** e **Figura 5.3**, respectivamente.

Para o modo Gota 1, eram adquiridos, a cada 10 min de análise, os espectros da corrente escura e de referência - com painel refletor (polímero branco com 99% de refletância difusa) que se encontra inserido na tampa do acessório (**Figura 5.2.a**) - com as duas tampas do equipamento fechadas (**Figura 5.2.b**) e com as luzes acesas. Após, as tampas eram abertas e a amostra era colocada, com auxílio de micropipeta, sobre a janela do equipamento (**Figura 5.2.c**). Era posicionada, então, a tampa transparente, que impede que a amostra entre em contato com o painel refletor (**Figura 5.2.d**) e, depois, a segunda tampa que retorna o sistema completamente fechado. Os espectros eram então adquiridos. A faixa de trabalho do equipamento está entre 1676 nm (5967 cm^{-1}) e 908 nm (11013 cm^{-1}). Outros parâmetros importantes: tempo de integração de 8 ms, aquisição de 100 varreduras e 125 pontos por varredura pelo software versão 2.1. Para o modo Gota 2, a amostragem e aquisição dos espectros ocorre da mesma maneira, porém com o software MicroNIR™ versão 3.0 a corrente escura passou a ser adquirida com as luzes apagadas. Aumentou-se também o número de varreduras para 200, de forma a minimizar ruídos.

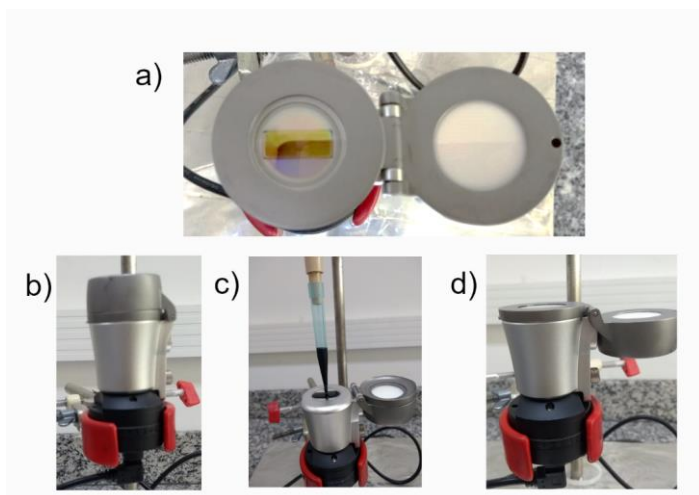


Figura 5.2. a) Visão superior do acessório (painel refletor na tampa à direita, janela amostral retangular com tampa transparente protetora sobre ela à esquerda); b) Visão lateral do sistema fechado; c) Amostragem; d) Visão lateral da janela amostral coberta apenas pela tampa protetora.

No Modo *Vial* utilizou-se os mesmos parâmetros com 100 varreduras. A corrente escura adquirida com acessório vazio – **Figura 5.3.a** – e luzes acesas. O espectro de referência foi feito com as luzes acesas com painel refletor fornecido pelo fabricante, um disco do polímero de 10 mm de altura e 12.5 mm de diâmetro dentro de um *vial* de 4 mL de vidro transparente (**Figura 5.3.b**). A corrente escura e de referência eram adquiridas a cada 10 min. Os espectros das amostras eram adquiridos com *vials* de 4 mL de vidro transparente preenchidos completamente com as amostras (**Figura 5.3.c**).

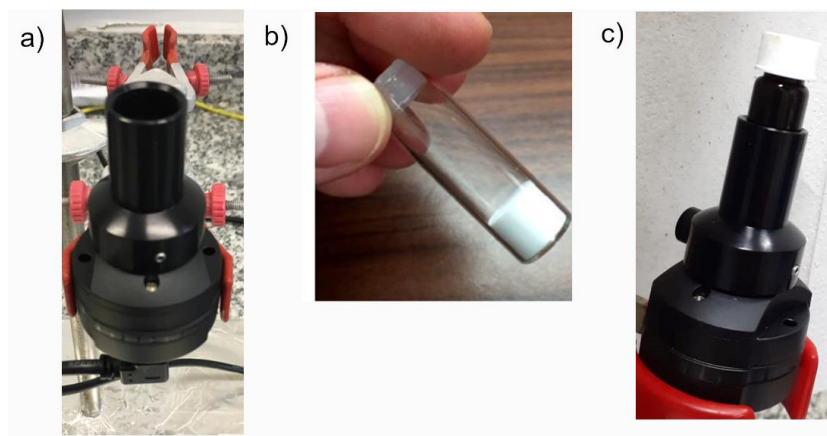


Figura 5.3. Modo *Vial* : a) Acessório vazio acoplado sobre MicroNIR; b) Painel refletor em *vial* de vidro transparente de 4 mL; c) *Vial* de 4 mL com tampa rosqueada preenchido com amostra de petróleo e inserido dentro do acessório *Vial* .

5.2.2.b. Equipamento de bancada

Espectrômetro de infravermelho FT-IR/FT-NIR Spectrum 400 Perkin Elmer com acessório de refletância total atenuada (UATR – IR), **Figura 5.4.a**. Parâmetros utilizados: faixa de 12.300 a 4.000 cm^{-1} , resolução de 8 cm^{-1} e 32 varreduras. Os espectros de referência eram adquiridos com cela de vidro de 35mm de diâmetro em vidro transparente com peça metálica transflatora (**Figura 5.4.b,c**), posicionando-os sobre a saída do feixe de infravermelho e fechando o sistema com cúpula do equipamento. Após, o sistema era aberto para se adicionar uma alíquota de 0,2 mL de amostra na cela, com auxílio de micropipeta, então, posicionava-se o transflator sobre a amostra, o sistema era novamente fechado e o espectro era adquirido.

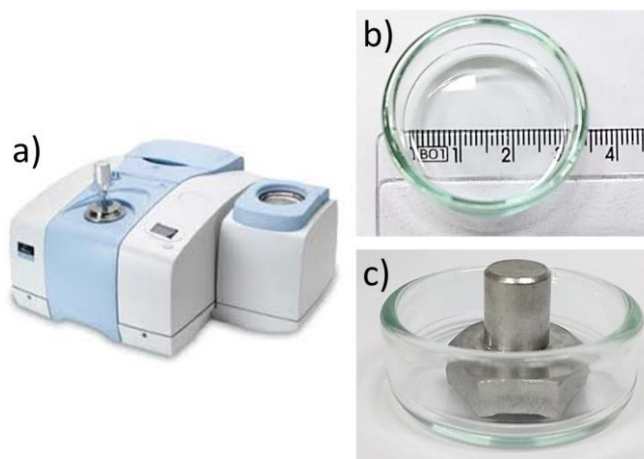


Figura 5.4. a) Espectrômetro de infravermelho FT-IR/FT-NIR Spectrum 400 Perkin Elmer com acessório de refletância total atenuada (UATR – IR) e b) Cella de vidro de 35 mm vazia e com c) peça transfectora utilizadas na análise

5.2.3. Análise dos dados

Investigou-se os pré-tratamentos mais utilizados para dados NIR (PANTOJA *et al.*, 2017) (isolados e combinados) e o que apresentou melhor correção foi escolhido e discutido nesse trabalho. Os conjuntos de espectros foram pré-processados com primeira derivada pelo algoritmo de Savitzky-Golay (SAVITZKY; GOLAY, 1964) com janela de sete pontos e ajuste polinomial de segundo grau. Inicialmente conduziu-se a análise exploratória com PCA. A aplicação da classificação se conduziu a seguir pelos modelos PLS-DA. As amostras foram separadas em subconjuntos de treinamento e teste para construção do modelo de classificação e validação dele, respectivamente. O modelo PLS-DA foi avaliado de acordo com os parâmetros de desempenho: sensibilidade, especificidade e exatidão.

$$\text{Sensibilidade} = \frac{TP}{TP+FN} 100$$

Equação 5.1

$$\text{Especificidade} = \frac{TN}{TN+FP} 100$$

Equação 5.2

$$\text{Exatidão} = \frac{TP+TN}{TP+TN+FP+FN} 100$$

Equação 5.3

onde FP e FN são os números de amostras falso positivo e falso negativo respectivamente; TP e TN são os números de amostras verdadeiro positivo e negativo, respectivamente (BALLABIO; CONSONNI, 2013c).

A sensibilidade representa a capacidade do modelo construído em identificar a amostra de uma determinada origem sabendo que a amostra realmente pertence a esta origem, ou seja, capacidade em identificar verdadeiros positivos. A especificidade é a capacidade do teste em identificar os verdadeiros negativos e a exatidão representa a concordância entre o valor esperado e o adquirido no teste.

5.3. Resultados e discussão

5.3.1. Análise exploratória para os diferentes modos de aquisição dos espectros

Como descrito na metodologia, foram testados três diferentes procedimentos (Modo *Vial*, Gota 1 e Gota 2). Para todos espectros brutos (**Figura 5.5.a, c, e**) é observada uma grande variação da linha de base espectral, característica comum em espectros NIR de petróleos pela presença dos asfaltenos (PANTOJA *et al.*, 2017). Não necessariamente há informação química relevante nessas diferenças de intensidade entre as amostras, sendo sempre necessário o pré-processamento dos espectros afim de corrigir os efeitos de espalhamento de luz (PANTOJA *et al.*, 2017; PASQUINI, 2018). Ainda assim, é interessante notar duas amostras (amostras 6 e 7) dos petróleos A, estas apresentam espectros com menor intensidade do que o restante do conjunto para o *Vial* (**Figura 5.5.a**) e Gota 2 (**Figura 5.5.e**).

Mas mais grave é a qualidade espectral para o modo *Vial* (**Figura 5.5.a**) que apresenta muito ruído, sofrendo tanto o efeito do espalhamento da luz que dificulta a análise clara das bandas e se assemelha mais aos espectros de solos (DOUGLAS *et al.*, 2018; NG; MALONE; MINASNY, 2017).

Os espectros para o modo gota 1 apresentaram valores de absorbância mais intensos, em torno de 1.5, e não mantém um perfil constante entre os espectros (**Figura 5.5.c**), o que não ocorre com os espectros para o modo gota 2 (**Figura 5.5.e**). Nesses dois conjuntos de espectros, observa-se o perfil de bandas largas

característico da região do infravermelho próximo. Pontua-se que as faixas entre 900 e 1000 nm, próximo de 1200 nm e próximo a 1400 nm. Nessas faixas encontram-se, respectivamente, o segundo sobretom da ligação O-H e as bandas de combinação do primeiro sobretom da ligação O-H, o segundo sobretom C-H, e primeiro sobretom O-H e as bandas de combinação do primeiro sobretom C-H (CORREA PABÓN; SOUZA FILHO, 2019; LAMMOGLIA; FILHO, 2011; PANTOJA *et al.*, 2017).

O método de Savitzky-Golay da primeira derivada foi aplicado para a correção da variação aleatória e deslocamento na linha de base nos espectros brutos (**Figura 5.5.b, d, f**). Os conjuntos de espectros pré-tratados apresentam menor variação nas intensidades das absorbâncias (**Figura 5.5.a, c, e**). No entanto, duas amostras de petróleos A para o modo *Vial* continuam distintas das demais (**Figura 5.5.b**). Observa-se também que os espectros para o modo gota 1 continuaram apresentando grande dispersão na faixa entre 900 e 1000 nm (**Figura 5.5.d**). Enquanto os espectros para o modo gota 2 apresentaram maior homogeneidade no conjunto (**Figura 5.5.f**).

Ainda que espectros NIR não tragam um detalhamento de informações que possibilite a resolução de problemas químicos pela observação dos espectros, ressalta-se a importância da análise espectral pois pode sinalizar amostras anômalas e mudanças na aquisição dos espectros, conforme ocorreu em nosso estudo. O desempenho dos modelos é consequência da qualidade na aquisição dos dados.

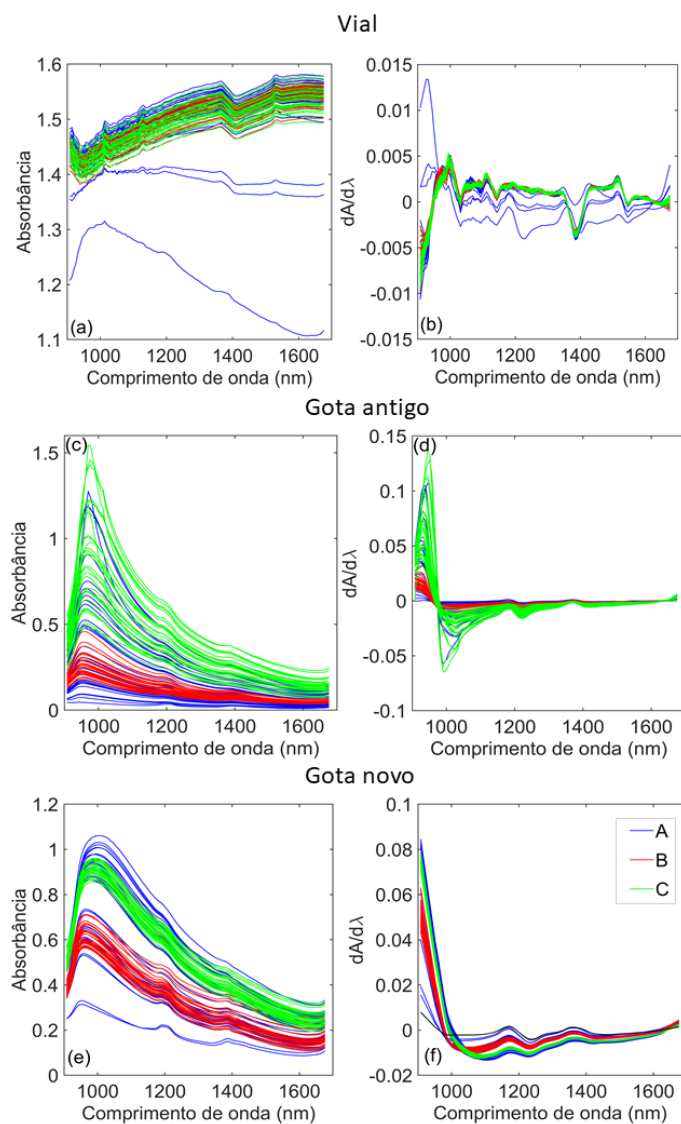


Figura 5.5. Espectros brutos de infravermelho na região do infravermelho próximo adquiridos pelo instrumento portátil MicroNIR para os Petróleo A, B e C no Modo **(a)** Vial ; **(c)** Gota 1; **(e)** Gota 2; E tratados com primeira derivada **(b)** Vial ; **(d)** Gota Antigo; **(f)** Gota Novo.

Dado o nível de informação obtidas nos espectros, realizamos a análise exploratória com a construção dos modelos PCA para investigar aonde estaria a

informação que possibilitasse a discriminação das amostras por origem. Investigamos às combinações com as cinco primeiras componentes principais. Discutimos anteriormente que os espectros no modo gota 1 apresentaram um perfil muito mais semelhante a espectros de petróleos (**Figura 5.5.c,d**) quando comparados ao modo *vial* (**Figura 5.5.a,b**). Logo foi uma surpresa não ser possível observar a menor separação entre os grupos das amostras para o modo gota 1 (**Figura 5.6.d,e**). Enquanto que, apesar da má qualidade dos espectros (ruídos e declive nos espectros) para o modo *Vial*, parte das amostras do grupo Petróleo A é distinguível para $PC3 > 2$, assim como parte das amostras do grupo Petróleo B em $PC3 < -2$ (**Figura 5.6.a,b**). Observa-se também que o comportamento anômalo das amostras 6 e 7 se preserva com a construção do modelo PCA, localizadas em $PC1 < -0,02$. Enquanto todas as outras amostras se localizam em $PC1 > -0,02$ (**Figura 5.6.a,b**). Nos gráficos de *scores* do conjunto modo gota 1, **Figura 5.6.d,e**, observa-se também algumas amostras separadas do resto do conjunto, porém conforme sinalizado não correspondem às mesmas amostras.

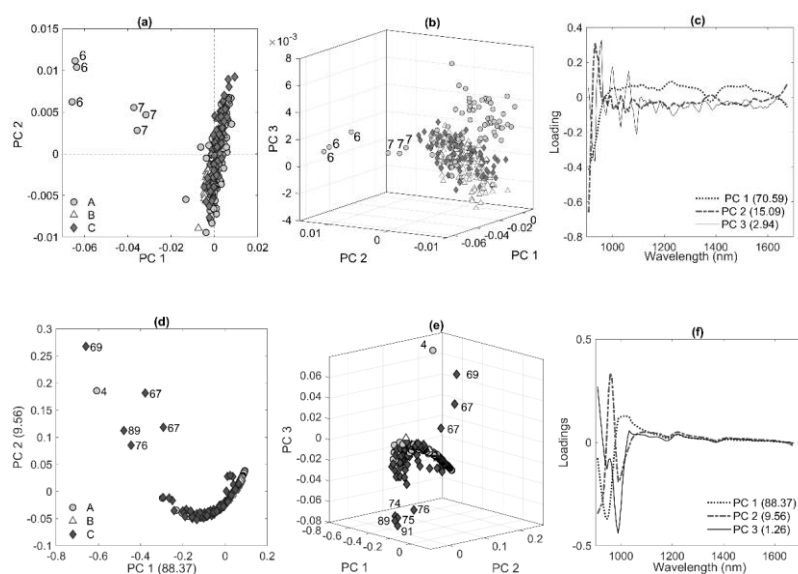


Figura 5.6. Gráficos dos modelos PCA construídos a partir dos espectros adquiridos para o modo *vial* (a – c) com 88.62% da variância explicada (E.V.), e para o modo gota 1 com 99.19% E.V. (d – f) para os grupos (●) Petróleo A, (▲) Petróleo B e (◆) Petróleo C; Gráfico dos *escores* para (a,d) PC1 versus PC2, (b,e) PC1 x PC2 x PC3; (c,f) Pesos das variáveis (comprimentos de onda) para as três PC.

Ainda que a análise dos gráficos de *scores* não tenha sido esclarecedora para a discriminação dos grupos de amostras, é interessante que na **Figura 5.6.c,f** as variáveis que apresentam o maior peso estão na faixa entre 900-1100 nm, região que sofre com o efeito do espalhamento de luz pelos asfaltenos (LAMMOGLIA; FILHO, 2011; PASQUINI; BUENO, 2007). Por conta desse problema é que chegamos à forma de aquisição gota 2. Pois os espectros de referência eram adquiridos com as luzes acesas para o modo *vial* e gota 1. Assim ao analisar as amostras, os sinais analíticos poderiam estar sendo sobrepostos pelos ruídos e após o pré-tratamento estar sendo suprimidos, uma consequência ao se fazer o alisamento dos espectros (FERREIRA, 2015). Nesse novo modo, aumentamos também a quantidade de corridas para diminuir ruídos. Não apresentamos também um 'modo *vial* 2' pois para a aquisição de espectros com esse acessório, o efeito do espalhamento continuou prejudicando a qualidade dos espectros, o que resultava no mesmo perfil ao já apresentado na **Figura 5.5.a,b** e, assim, não continuamos as análises. Para esse novo conjunto de espectros também construímos o modelo PCA (**Figura 5.7**).

Na **Figura 5.7.a** se observa que a otimização no modo de análise possibilitou a discriminação do Petróleo B e C pelo gráfico de *scores* do modelo PCA com 99,6% da E.V. para as duas primeiras componentes (**Figura 5.7.a**). E assim como observado na **Figura 5.6**, o gráfico dos *loadings* (**Figura 5.7.c**) traz a informação de que as variáveis responsáveis por essa separação são influenciadas pela região inicial do espectro, abaixo de 1100 nm. Porém, mesmo analisando todas as outras componentes principais – como no caso da **Figura 5.7.b** em que confrontamos PC1 x PC2 X PC3 – não é possível separar o grupo do Petróleo A. Além disso, na **Figura 5.7.a** observa-se duas amostras deste grupo afastadas do restante do conjunto, estas amostras são as mesmas de menor viscosidade (amostras 6 e 7) que foram pontuadas na observação dos espectros **Figura 5.5.e**.

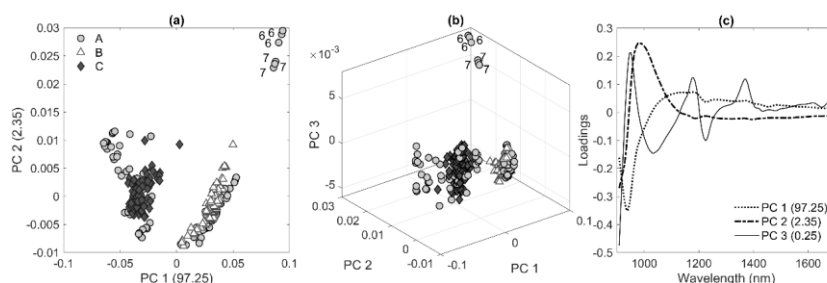


Figura 5.7. Gráficos do modelo PCA para o modo gota 1 com 99.85% E.V. (a – c) para os grupos (●) Petróleo A, (▲) Petróleo B e (◆) Petróleo C; Gráficos dos scores para a) PC1 versus PC2, b) PC1 x PC2 x PC3; c) Loadings das variáveis (comprimentos de onda) para as três PC.

Não sendo possível a discriminação quanto as três origens com o método não supervisionado, construiu-se então o modelo PLS-DA com os espectros obtidos pelo modo Gota 2. Foi construído com cinco variáveis latentes e foi aplicado ao conjunto teste. A probabilidade de cada amostra pertencer a cada classe é mostrada na **Figura 5.8**. A linha horizontal pontilhada em vermelho refere-se ao limite de decisão para a amostra pertencer ou não a determinada classe calculada pelo teorema de Bayes.

Na **Figura 5.8.a** aparentemente uma amostra do grupo teste do Petróleo A não foi classificada como pertencente, enquanto uma amostra de treinamento do Petróleo B é classificada como pertencente. Porém, o modelo não conseguiu prever essas amostras dentro de nenhuma das classes, de acordo com a matriz de confusão. O mesmo acontece na **Figura 5.8.b e c**.

De acordo com o modelo apenas uma amostra de treinamento do Petróleo B é classificada incorretamente como Petróleo C, a qual pode ser vista na **Figura 5.8.b**, amostra abaixo da linha de decisão. Isso se traduz como uma taxa de erro na calibração de 0,01 e exatidão de 0,98.

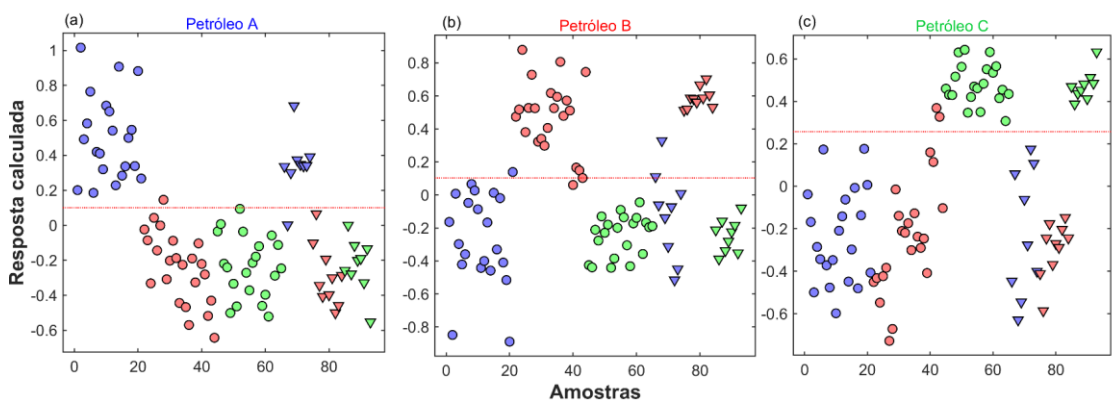


Figura 5.8.Gráfico dos scores do modelo PLS-DA construído com dados NIR obtidos no equipamento portátil para a probabilidade das amostras pertencerem à classe do **(a)** Petróleo A (amostras em azul); **(b)** B (amostras em vermelho); **(c)** C (amostras em verde). Amostras de treinamento (o) e teste (Δ).

A **Tabela 5.1** traz os valores para os parâmetros de desempenho avaliados no modelo. A amostra do Petróleo B classificada erroneamente como Petróleo C se traduz na sensibilidade de 95% para o modelo classificar a classe do Petróleo B e na especificidade de 97% da classe do Petróleo C, o que prejudica a exatidão, 95%, do modelo em classificar a classe Petróleo C. Porém, não prejudicou os valores de predição com sensibilidade, especificidade e exatidão igual a 100% para todas as classes.

Tabela 5.1.Parâmetros de desempenho do modelo PLS-DA com 5 variáveis latentes.

	Parâmetros	Petróleo A	Petróleo B	Petróleo C
Treinamento	Especificidade	100%	100%	97%
	Sensibilidade	100%	95%	100%
	Exatidão	100%	100%	95%
Validação Cruzada	Especificidade	100%	100%	97%
	Sensibilidade	100%	95%	100%
	Exatidão	100%	100%	95%
Teste	Especificidade	100%	100%	100%
	Sensibilidade	100%	100%	100%
	Exatidão	100%	100%	100%

Assim, o modelo PLS-DA se mostrou eficiente na resolução do problema em comparação com a PCA como já havia sido avaliado em relatórios anteriores.

5.3.2. Comparação com instrumento de bancada para a classificação das amostras de petróleos

Outra problemática levantada foi se a sensibilidade do instrumento portátil era responsável pela dificuldade na resolução de problemas com a matriz petróleo, se comparado com o instrumento de bancada. Isso pois são adquiridos menos pontos por corrida (125 *versus* 4151) e a faixa de análise é mais estreita (11.013 – 5.967 *versus* 13.300 – 4000 cm^{-1}).

Assim, os espectros adquiridos no equipamento de bancada tiveram modelos PCA e PLS-DA construídos de três formas diferentes:

Modo 1) Alta resolução com a faixa completa (4151 pontos entre 13.300 – 4.000 cm^{-1}): **Figura 5.9.a;**

Modo 2) Espectros cortados na faixa correspondente ao MicroNIR - Alta resolução com faixa estreita (2524 pontos entre 11.013 – 5.967 cm^{-1}): **Figura 5.9.b;**

Modo 3) Espectros cortados na faixa correspondente ao MicroNIR com interpolação aplicada para que o número de variáveis fosse também correspondente ao MicroNIR - Baixa resolução (125 pontos entre 11.013 – 5.967 cm^{-1}): **Figura 5.9.c.**

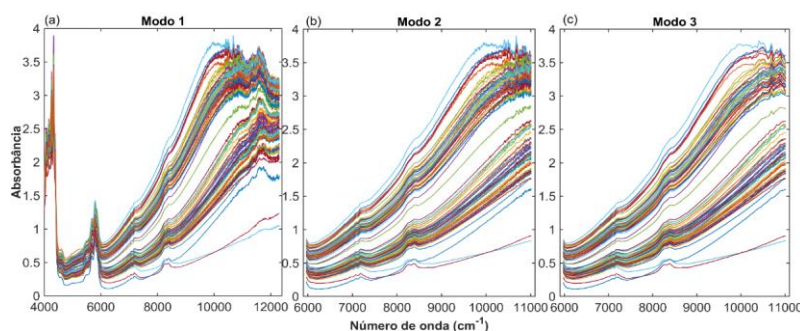


Figura 5.9.a -c. Espectros de infravermelho na região do infravermelho próximo adquiridos pelo instrumento de bancada para os Modos 1, 2 e 3, respectivamente.

É possível observar que (**Figura 5.9.a-c**) acima de 10.000 cm^{-1} há muito ruído, isso porque está acima da faixa ótima de trabalho do equipamento. Também há grande variação na linha de base espectral, a qual foi corrigida com primeira derivada.

Ao fazer o corte dos espectros para coincidir com a faixa de trabalho do MicroNIR são eliminadas as bandas de estiramento CH (em torno de 4.608 cm^{-1}), primeiro sobretudo do estiramento CH_2 e das combinações dos estiramentos CH_2 e CH_3 (entre $6.024 - 5.586\text{ cm}^{-1}$) para os Modos 2 e 3 (**Figura 5.9.b e c**). Porém isso não interferiu de forma significativa na classificação pela PCA entre os três modos, como podemos ver nos gráficos de *scores* e *loadings* (**Figura 4.6.a-f**). As duas primeiras componentes, as quais apresentam a maior variância explicada, foram as que mostraram melhor a separação dos grupos (**Figura 4.6.a-c**). Porém, em nenhum dos três casos foi possível classificar de acordo com as três origens. Apenas foi possível separar Petróleo B e C.

É interessante observar que para o Modo 3, o gráfico dos *scores* e *loadings* (**Figura 4.6 c e f**) é muito semelhante ao observado nos dados de MicroNIR, o que confirma que dentro das condições do equipamento portátil a faixa que promove a separação está concentrada abaixo de 1100 nm .

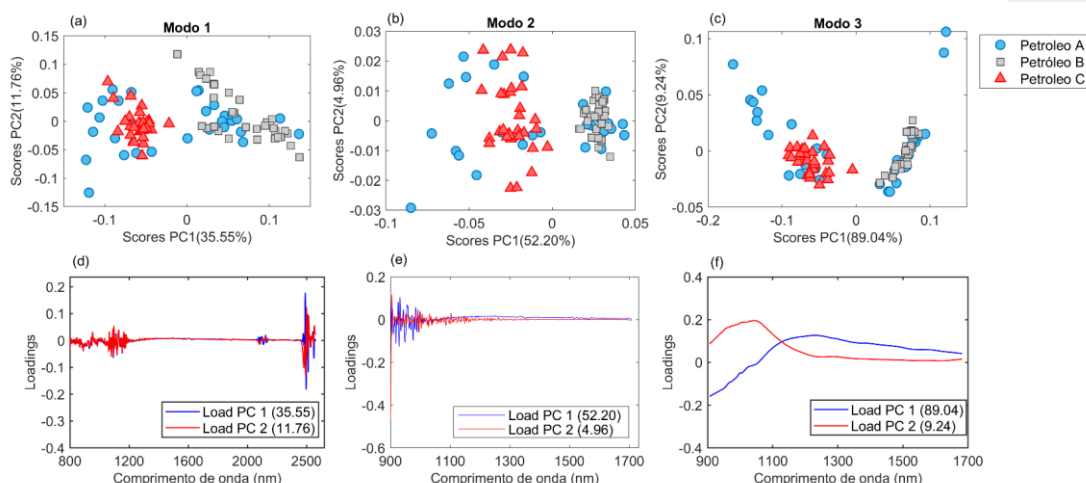


Figura 5.10. Gráfico dos *scores* e *loadings*, respectivamente, dos modelos PCA construídos para o (a,d) Modo 1, (b,e) Modo 2 e (c, f) Modo 3. (○) Petróleo A; (□) Petróleo B e; (▲) Petróleo C.

Os modelos PLS-DA foram construídos com 7 variáveis latentes. Esse foi o número ótimo, de acordo com a menor taxa de erro e menor quantidade de amostras não classificadas. De forma a facilitar a compreensão, para os modelos PLS-DA dos Modos 1, 2 e 3 é apresentada apenas a **Tabela 4.2** que traz os valores dos parâmetros de desempenho. Porém, observa-se que os resultados não apresentaram exatidão comparável aos obtidos no equipamento portátil. Os parâmetros de calibração para os Modos 1 e 2 foram bem inferiores a 100%. Isso pode ser estar relacionado ao espalhamento da informação por todo espectro da região do infravermelho. Com a “seleção de variáveis” que ocorre no Modo 3, regiões que tem menor intensidade, passam a carregar maior volume de variância explicada e a ter importância na separação das classes. Além disso, a interpolação pode ter suavizado regiões de ruído.

Tabela 5.2. Parâmetros de desempenho dos modelos PLS-DA para os Modos 1, 2, 3 para espectros NIR obtidos com equipamento de bancada.

		Parâmetros	Petróleo A	Petróleo B	Petróleo C
Modo 1	Treinamento	Especificidade	100%	100%	100%
		Sensibilidade	100%	100%	100%
		Exatidão	100%	100%	100%
	Validação Cruzada	Especificidade	92%	94%	79%
		Sensibilidade	48%	95%	87%
		Exatidão	77%	91%	59%
	Teste	Especificidade	90%	88%	80%
		Sensibilidade	80%	100%	88%
		Exatidão	80%	88%	100%
Modo 2	Treinamento	Especificidade	100%	100%	100%
		Sensibilidade	100%	100%	100%
		Exatidão	100%	100%	100%
	Validação Cruzada	Especificidade	88%	83%	78%

Modo 3		Sensibilidade	18%	100%	67%
		Exatidão	38%	79%	57%
		Especificidade	73%	40%	100%
	Teste	Sensibilidade	100%	100%	25%
		Exatidão	100%	54%	100%
		Especificidade	100%	100%	97%
	Treinamento	Sensibilidade	100%	95%	100%
		Exatidão	100%	100%	95%
		Especificidade	95%	97%	85%
	Validação Cruzada	Sensibilidade	65%	100%	89%
		Exatidão	87%	95%	74%
		Especificidade	100%	82%	100%
	Teste	Sensibilidade	67%	100%	100%
		Exatidão	100%	75%	100%

5.4. Conclusão

Com base nos modelos PCA foi possível entender que a discriminação entre Petróleo A, B e C não estavam ligadas a uma menor eficiência do equipamento portátil, visto que os modelos PCA construídos com equipamento de bancada também não conseguiram discriminar as amostras em três distintos agrupamentos. Além de evidenciar que o melhor modo de análise para petróleos realmente foi a metodologia Gota 2 que obteve espectros com menos ruídos e modelos com maior variância explicada.

Já os modelos PLS-DA demonstraram ser um método mais apropriado para a resolução da problemática. Para os três grupos, na etapa de predição, as amostras conseguiram ser 100% corretamente classificadas. O modelo PLS-DA também evidencia a eficiência do equipamento portátil, pois apesar de ter menor resolução espectral e faixa espectral obteve melhor desempenho quando comparados aos modelos PLS-DA em três condições distintas com o equipamento de bancada. Para o

modo 1, 2 e 3 houveram erros de classificação de amostras na etapa de predição. Dessa forma, fechamos uma metodologia eficiente para utilização de um espectrômetro NIR portátil para análise de petróleos.

6. APLICAÇÃO 2

Caracterização de petróleos com espectrômetro NIR portátil

6.1. Introdução

Petróleos são caracterizados por suas propriedades físico-químicas, determinadas por normas internacionais. Em substituição aos métodos tradicionais de análise, muitos estudos têm proposto métodos alternativos utilizando técnicas espectroscópicas associadas à análise multivariada de dados. (CORREA PABÓN; SOUZA FILHO, 2019) Propriedades que já foram estimadas com dados de infravermelho próximo (NIR): densidade, gravidade API ($^{\circ}$ API), teor de saturado (Sat), aromático (Aro), resinas (Res) e asfaltenos (Asf), ponto de ebulição verdadeiro curva, teor de nitrogênio total, teor de nitrogênio básico (TNB), teor de enxofre total (TET) e número de acidez total. (MORO *et al.*, 2021) A espectroscopia na região do infravermelho próximo é uma das mais utilizadas para esse fim, não apenas para petróleos, mas em diversas outras matrizes. (PASQUINI, 2018) Espectrômetros NIR portáteis já foram utilizados, por exemplo em estudos com alimentos, (TEIXEIRA DOS SANTOS *et al.*, 2013) controle de qualidade de combustíveis, (BROUILLETTE *et al.*, 2016; CORREIA *et al.*, 2017; DA SILVA, Neirivaldo Cavalcante *et al.*, 2017a; PAIVA *et al.*, 2015; SANTOS *et al.*, 2021) análise de medicamentos (DEIDDA *et al.*, 2019) e investigações forenses. (SILVA, Carolina S.; BRAZ; PIMENTEL, 2019) Esse tipo de equipamento proporciona análises rápidas e não destrutivas que dispensam o preparo de amostras e o uso de reagentes. No entanto, possui uma resolução menor (número de pontos adquiridos por espectro) do que um instrumento de bancada. Em geral, os dispositivos portáteis têm resolução de 10 nm, SNR abaixo de 10.000:1 e funcionam em apenas parte da faixa NIR (geralmente, operam na faixa entre 900 e 1700 nm). (PASQUINI, 2018) Isso pode afetar o desempenho dos modelos construídos para estimar as propriedades físico-químicas em matrizes complexas como petróleos.

Avanços tecnológicos como os filtros de interferência variável (do inglês, linear variable filter – VLF) possibilitaram o desenvolvimento de componentes ópticos miniaturizados e, conseqüentemente, a dispositivos portáteis com alta robustez. (CROCOMBE, 2018) Os espectrômetros NIR portáteis podem ser implementados *in*

situ, permitindo a flexibilidade de medições no local e fornecendo instrumentos de dados em tempo real. Os primeiros espectrômetros NIR ficaram disponíveis nos anos 2000,(CAPITÁN-VALLVEY; PALMA, 2011; CROCOMBE, 2018) Nessa época, a estimativa das propriedades físico-químicas de petróleos em espectrômetros de bancada também se iniciava.(HIDAJAT; CHONG, 2000) Porém, em equipamento portátil, o primeiro trabalho encontrado é de 2010, Além disso, não trata exatamente da estimativa de propriedades físico-químicas em petróleos, mas sim da quantificação de contaminantes hidrocarbonetos proveniente de solos impregnados com petróleo.(CHAKRABORTY *et al.*, 2010) Os autores conseguem um modelo PLS com R^2 equivalente a 0,64. Nosso estudo se propõe a ser o primeiro trabalho tratando da estimativa de propriedades físico-químicas em amostras de petróleos. O objetivo principal é avaliar a aplicabilidade de um espectrômetro NIR portátil para prever parâmetros físico-químicos de caracterização de petróleos brutos. Para isso, foram utilizados os métodos PLS e de regressão de vetores-suporte (SVR). Estimamos sete propriedades (densidade API, SARA, teor de enxofre total e teor de nitrogênio total) com um espectrômetro NIR portátil, bem como com um espectrômetro de bancada para comparar as respostas.

6.2. Metodologia

Figura 6.1 resume a metodologia contruída nesse trabalho. As amostras de petróleo foram primeiramente caracterizadas pelos métodos de referência e, em seguida, adquiridos os espectros nos equipamentos de bancada e portátil. Com esses dois conjuntos de dados espectrais construímos cinco diferentes matrizes, alterando entre elas a resolução e faixa espectral: FT-NIR I, FT-NIR II, FT-NIR III, Micronir I e Micronir II). Com essas cinco matrizes foram construídos os modelos PLS. Além disso, também construiu-se modelos SVR para a matriz Micronir I.

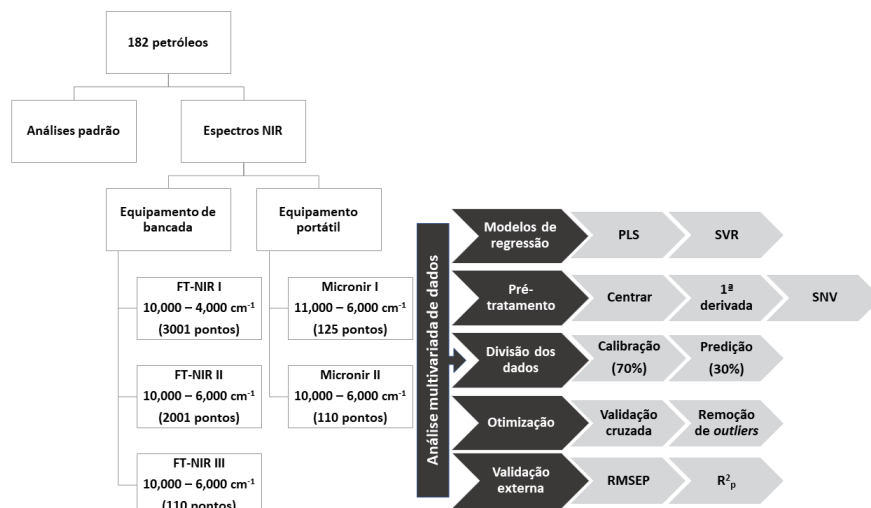


Figura 6.1. Resumo das principais etapas da metodologia construída

6.2.1. Amostras

O conjunto amostral refere-se a 182 óleos brutos da bacia sedimentar brasileira fornecidos pelo Centro de Pesquisa e Desenvolvimento da Petrobras (CENPES). O procedimento de amostragem e manuseio seguiu as diretrizes do ASTM D5854.(ASTM INTERNATIONAL, 2019) A amostragem ocorreu em triplicata. Os métodos para a caracterização das amostras (ou seja, as análises de referência) estão na **Tabela 6.1**, além da média dos resultados e as medianas. Esses resultados ressaltam a variabilidade do conjunto amostral.

Tabela 6.1. Informações sobre as análises de referência utilizadas para caracterização das amostras de petróleo

Prop	Unidade	Norma	Valor médio	Mediana
°API	Grau API	(ISO 12185, 1996)	11.4 – 57.5	29.0
TET	% m/m	ASTM D4294 (ASTM INTERNATIONAL, 2008)	0,0044 – 1.97	0,33
TNB	% m/m	UOP 269(UOP LLC, 2010)	0 – 0,47	0,10
Sat	% m/m		29.2 – 90,6	57.1
Aro	% m/m	ASTM D2549(ASTM INTERNATIONAL, 2002)	7.00 – 37.2	23.4
Res	% m/m		0 – 42.8	18.0
Asf	% m/m		0 – 10,3	0,65

6.2.2. Espectroscopia NIR

O espectrômetro portátil utilizado foi um MicroNIR® Pro 1700 com software para análise versão 3.0 da Viavi Solutions (Micronir). As condições de operação: modo reflectância difusa, faixa espectral de 11.013 a 5.967 cm^{-1} com resolução de 40 cm^{-1} e 100 varreduras. O espectro de referência para 100% de reflectância utilizou um Spectralon™. Mais informação pode ser encontrada em um trabalho anterior.(SANTOS *et al.*, 2021)

O instrumento de bancada utilizado foi um FT-IR Spectrum 400 com acessório externo de reflectância para NIR (FT-NIR) com software para análise da Perkin Elmer. As condições de operação: modo reflectância difusa, faixa espectral de 10.000 a 4.000 cm^{-1} com resolução de 2 cm^{-1} e 32 varreduras.

6.2.3. Análise de dados

Como mostrado pela **Figura 6.1** organizamos os dados espectrais obtidos em cada um dos equipamentos (FT-NIR e Micronir) em cinco diferentes matrizes (FT-NIR I, FT-NIR II, FT-NIR III, Micronir I e Micronir II). A diferença entre elas está na quantidade de pontos e números de onda por espectro. FT-NIR I contém 3.001 variáveis e faixa espectral entre 10.000 e 4.000 cm^{-1} ($\mathbf{X}_{182 \times 3001}$); FT-NIR II contém 2.001 variáveis e faixa entre 10.000 e 5.967 cm^{-1} ($\mathbf{X}_{182 \times 2001}$); FT-NIR III contém 110 variáveis e faixa entre 10.000 e 5.967 cm^{-1} ($\mathbf{X}_{182 \times 110}$); Micronir I contém 125 variáveis e faixa entre 11.013 e 5.967 cm^{-1} ($\mathbf{X}_{182 \times 125}$) e Micronir II contém 110 variáveis e faixa entre 10.000 e 5.967 cm^{-1} ($\mathbf{X}_{182 \times 110}$). Organizamos os dados nestes cinco conjuntos para observar o desempenho dos modelos para a configuração padrão de cada instrumento (FT-NIR I e Micronir I), na faixa espectral comum aos dois instrumentos (FT-NIR II e Micronir II) e igualeo a resolução espectral (FT-NIR III). Usamos um algoritmo de interpolação para reduzir a matriz de 3.001 pontos obtidos em FT-NIR para os mesmos 110 números de onda do conjunto de dados Micronir II.

Os espectros NIR foram pré-tratados (pretrat) com diferentes algoritmos (um por um e combinados). Apresentamos na discussão apenas os modelos com os melhores parâmetros de desempenho. Os algoritmos utilizados foram o algoritmo de primeira derivada com um polinômio de segunda ordem e uma janela de 7 pontos (Der),(SAVITZKY; GOLAY, 1964) centralização (Center) e variação normal padrão (do inglês, steard normal variate – SNV). Posteriormente, modelos de regressão foram

construídos uso algoritmos PLS e SVR. Todos os cálculos foram realizados no software Mathworks MATLAB 7.0 (R2013a) (Eigenvector Research, Inc.).

As amostras foram separadas por ranqueamento pelo método Kennard-Stone (LIANG *et al.*, 2016) nos subconjuntos calibração (70% das amostras) e predição (30% restantes). Na construção dos modelos PLS, o número de variáveis lantes (VL) foi otimizado. A otimização foi executada com validação cruzada pelo método Monte Carlo e as amostras anômalas (*outliers*) retiradas pela análise do gráfico de resíduos.(LI; XU; LIANG, 2014) Os modelos SVR foram construídos conforme metodologia de Folli e colaboradores (2020).(FOLLI *et al.*, 2020)

O desempenho dos modelos foi determinada pelos parâmetros média da raiz do erro quadrático de predição (do inglês, root mean square error of prediction – RMSEP) e coeficiente de determinação de predição (R_p^2) calculados pelas $RMSEP =$

$$\sqrt{\frac{1}{n_{pred}} \sum_{i=1}^{n_{pred}} (y_i - \hat{y}_i)^2} \quad \text{Equação 6.1 e}$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad \text{Equação}$$

6.2, respectivamente:

$$RMSEP = \sqrt{\frac{1}{n_{pred}} \sum_{i=1}^{n_{pred}} (y_i - \hat{y}_i)^2} \quad \text{Equação 6.1}$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad \text{Equação 6.2}$$

onde n_{pred} é o número de amostras no subconjunto de predição, y_i , $\hat{y}_i$ e $\bar{y}$ são, respectivamente, os valores de referência, predito, e média observados pelo modelo.

6.3. Resultados e discussão

Os espectros para as amostras de petróleo podem ser vistos na **Figura 6.2** para o equipamento portátil (**Figura 6.2.a**) e para o equipamento de bancada (**Figura 6.2.b**). Na faixa espectral que compreende os dois instrumentos, de 11.013 cm^{-1} a 4.000 cm^{-1} , estão localizadas as seguintes bandas de absorção: em torno de 11.000 cm^{-1} , segundo sobretom OH e terceiro sobretom CH; Em torno de 10.000 cm^{-1} , segundo sobretom NH; Entre 9.000 e 8.300 cm^{-1} , segundo sobretom CH; Entre 7.700 e 7.100 cm^{-1} , primeiro sobretom OH e combinações dos modos vibracionais CH; em torno de 6.600 cm^{-1} , primeiro sobretom NH; entre $6,200$ e $5,500 \text{ cm}^{-1}$, primeiro

sobretudo SH, segundo sobretudo C=O e primeiro sobretudo CH; entre 5.000 e 4.000 cm^{-1} , combinações dos modos vibracionais das ligações OH, NH, C-C e CH.(XIAOBO *et al.*, 2010) Acima de 6.000 cm^{-1} há intenso espalhamento de radiação causado pelos asfaltos,(PANTOJA *et al.*, 2017) por isso na **Figura 6.2.a** aparentemente há apenas uma banda larga. Todas as bandas de absorção descritas nessa região estão sobrepostas. Na **Figura 6.2.b** essa larga banda também é vista, porém também é possível observar entre 4.500 e 4.000 cm^{-1} as combinações dos modos vibracionais das ligações CH e o primeiro sobretudo CH entre 5.500 e 6.250 cm^{-1} .

Devido ao intenso espalhamento da radiação, para todos os modelos apresentados nas tabelas abaixo, informamos quais pré-tratamentos induzem o melhor desempenho. O pré-tratamento reduz os efeitos não informativos nos espectros.(PANTOJA *et al.*, 2017)

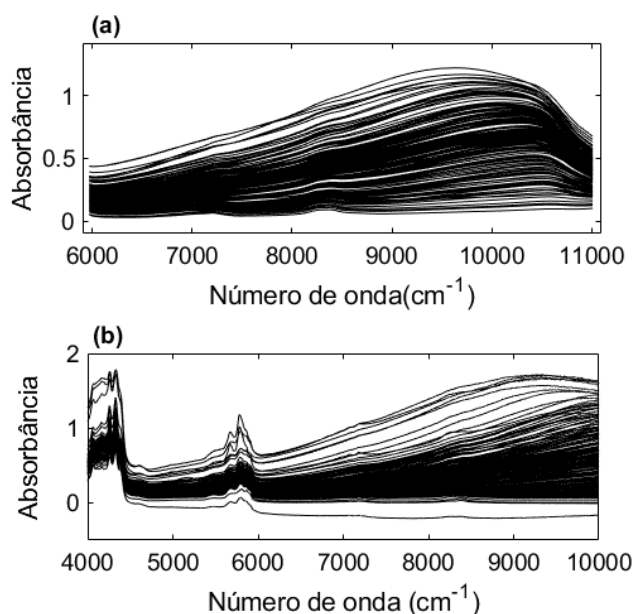


Figura 6.2. Espectros brutos das amostras de petróleos adquiridos com os espectrômetros (a) Micronir e (b) FT-NIR

Primeiramente, mostramos a comparação entre o instrumento de bancada e o portátil. Na **Tabela 6.2** estão os parâmetros de desempenho dos modelos PLS para as sete propriedades dos conjuntos de dados FT-NIR I e Micronir I. Considerando que

R^2 acima de 0,5 indica uma correlação moderada e acima de 0,7 uma correlação forte,(MOORE; NOTZ; FLIGNER, 2017) os modelos °API, TNB e TET foram satisfatórios em ambos instrumentos. Sat e Res foram satisfatórios apenas em FT-NIR I e apenas Asf teve melhor desempenho com Micronir I do que com FT-NIR I (R^2_p 0,72; RMSEP 0,34 % m/m *versus* R^2_p 0,60; RMSEP 0,42 % m/m, respectivamente). Os modelos Aro não foram satisfatórios para ambos instrumentos ($R^2_{\text{Micronir I}}$ 0,34; $R^2_{\text{FT-NIR I}}$ 0,44).

Tabela 6.2. Desempenho dos modelos PLS para as sete propriedades construídos com os dados NIR dos instrumentos portátil e de bancada, respectivamente.

Micronir I					FT-NIR I			
Prop	Pré -trat	RMSEP	R ² _p	VL	Pré -trat	RMSEP	R ² _p	VL
°API	Center	2,45	0,80	19	SNV	1,44	0,94	10
TET	Der	0,11 % m/m	0,62	5	Der	0,07 % m/m	0,82	8
TNB	Der	0,02 % m/m	0,73	5	Der	0,01 % m/m	0,86	7
Sat	SNV+Der	7,37 % m/m	0,34	10	Der	5,77 % m/m	0,58	7
Aro	Der	3,69 % m/m	0,34	9	SNV+Der	3,37 % m/m	0,44	5
Res	Center	5,80 % m/m	0,33	4	Der	4,84 % m/m	0,55	6
Asf	Center	0,34 % m/m	0,72	5	Der	0,42 % m/m	0,60	6

Densidade e °API são as propriedades que mais foram estimadas para petróleo e seus derivados.(BAIRD; OJA, 2016; MORO *et al.*, 2021) É comum que sejam a propriedade usada inicialmente para o teste de um novo modelo com dados de infravermelho, por causa do comportamento linear dessas propriedades.(BAIRD; OJA, 2016) A correlação forte acontece pois a informação está distribuída por todo o espectro.(DE PAULO *et al.*, 2022; RAINHA *et al.*, 2019) Assim, não é surpreendente que os modelos °API tenham apresentado o melhor desempenho. O modelo °API para FT-NIR I apresentou R^2_p 0,94 e RMSEP 1,44 °API. Esse não é o melhor resultado na literature, há trabalhos com R^2_p acima de 0,98 e RMSEP abaixo de 1,00 °API. (PASQUINI; BUENO, 2007; RAINHA *et al.*, 2019) Porém, eles não utilizam a mesma faixa spectral, 10.000 cm^{-1} a 4000 cm^{-1} . Eles utilizam faixas mais estreitas por remover a faixa acima de 9000 cm^{-1} , por conta do espalhamento de radiação nessa região. Apenas Pábon e colaboradores (2019) usam a mesma faixa spectral.(CORREA PABÓN; SOUZA FILHO, 2019) Eles obtiveram R^2 0,99 e RMSEP 1,11 °API. Porém, utilizam um conjunto amostral muito menor, apenas 37 amostras de petróleo.

Não obstante, o modelo $^{\circ}\text{API}$ para o conjunto Micronir I obteve R^2_p 0,80 e RMSEP 2,45 $^{\circ}\text{API}$. Apesar de não obter o mesmo desempenho dos trabalhos previamente discutidos, é o único que trabalha com um espectrômetro portátil. A condição mais próxima talvez esteja no trabalho de Lammoglia e Souza Filho (2011). (LAMMOGLIA; FILHO, 2011) Os autores utilizaram um espectrorradiômetro VIS-NIR e obtiveram para o modelo PLS para $^{\circ}\text{API}$, R^2_p 0,997 e RMSEP 0,597 $^{\circ}\text{API}$. Entretanto, a faixa espectral difere, 28.571 a 4.000 cm^{-1} , e o conjunto amostral é de apenas 17 petróleos na faixa $^{\circ}\text{API}$ entre 13 e 40,2. Ou seja, o espectrômetro NIR utilizado em nosso estudo coleta uma região menor, logo menos bandas de absorção, além de nosso conjunto amostral apresentar maior variabilidade, o que significa resultados mais robustos. A menor quantidade de informação vibracional coletada pode ter influenciado o número de VL (19). Um número alto, mas que está dentro do permitido de acordo com a ASTM E1655. (ASTM E1655-05, 2012) Além disso, pelas métricas não há superajuste em nosso modelo e por isso consideramos o modelo estatisticamente satisfatório. (MOORE; NOTZ; FLIGNER, 2017)

Para TET, FT-NIR I apresentou R^2_p 0,82 e RMSEP 0,07 % m/m. Enquanto, Micronir I apresentou R^2_p 0,62 e RMSEP 0,11 % m/m. O primeiro sobretudo da ligação SH está localizado na região entre 6,200 e 5,500 cm^{-1} . Quase toda essa região não é adquirida pelo instrument portátil. Três trabalhos estimaram o teor de enxofre em petróleos. O primeiro, Long e colaboradores (2019) usaram 153 petróleos com teores de enxofre entre 0,3 e 3,8 % m/m e um FT-NIR opereio na faixa entre 4.599–7.500 cm^{-1} . (LONG *et al.*, 2019) Eles não apresentam os parâmetros de desempenho da etapa de validação externa, apenas da calibração, R^2_c 0,93 e RMSEC 0,37 % m/m. O segundo estudo, Folli e colaboradores (2020) usaram 93 amostras de petróleos com TET entre 0,0044 e 1, 97 % m/m e um espectrômetro FT-NIR na faixa 4.000–10.000 cm^{-1} . (FOLLI *et al.*, 2020) O melhor resultado foi obtido com um modelo PLS combinado a função kernel (k-PLS), R^2_p 0,60 e RMSEP 0,08 % m/m. Finalmente, Šašić e colaboradores (2021) usaram 1.300 petróleos com TET entre 0,7 e 3,0 % m/m e também um equipamento FT-NIR. Os autores apresentam seis modelos PLS, cada um construído numa faixa espectral diferente contida na região entre 8.000 e 4.600 cm^{-1} . Não é apresentado o valor de R^2_p , mas os valores para RMSEP estavam todos em torno de 0,1 % m/m. Assim, os resultados de nossos modelos ficaram similares aos resultados prévios destes autores.

Para TNB, FT-NIR I apresentou R^2_p 0,86 e RMSEP 0,01 % m/m. Enquanto Micronir I apresentou R^2_p 0,73 e RMSEP 0,02 % m/m. Rainha e colaboradores (2019) usaram 194 amostras de petróleo com TNB entre 0,0006 e 0,26 % m/m e um espectrômetro FT-NIR. O melhor resultado obtido pelos autores utilizou o modelo PLS combinado a seleção de variáveis amostragem reponderada adaptativa competitiva (do inglês, *competitive adaptive reweighted sampling* - CARS), o qual selecionou 197 variáveis na faixa 8.460–4.006 cm^{-1} e obteve R^2_p 0,92 e RMSEP 0,01 % m/m. Enquanro, Folli e colaboradores (2020) usou 92 amostras de petróleo com TNB entre 0,0005 e 0,26 % m/m e obtiveram o melhor resultado com modelo k-PLS,. R^2_p 0,83 e RMSEP 0,02 % m/m (FOLLI *et al.*, 2020) Apesar do R^2_p para os nossos modelos ser menor do que desses estudos, os erros são similares. Assim, consideramos que os resultados são competitivos.

Finalmente, os modelos Sat, Aro, Res e Asf, respectivamente. Para FT-NIR I: R^2_p 0,58 e RMSEP 5,77 % m/m; R^2_p 0,44 e RMSEP 3,37 % m/m; R^2_p 0,55 e RMSEP 4,84 % m/m; R^2_p 0,60 e RMSEP 0,42 % m/m. Já, para Micronir: R^2_p 0,34 e RMSEP 7,37 % m/m; R^2_p 0,34 e RMSEP 3,69 % m/m; R^2_p 0,33 e RMSEP 5,80 % m/m; R^2_p 0,72 e RMSEP 0,34 % m/m. O único outro estudo que estimou essas propriedades em petróleos com dados NIR foi Aske e colaboradores (2001)(ASKE; KALLEVIK; SJÖBLOM, 2001). Eles utilizaram 18 petróleos e adquiriram os espectros com um equipamento de bancada na região entre 9.009 a 4.545 cm^{-1} . Todos modelos apresentaram R^2 de 0,99 e RMSEP, para Sat 2,8 % m/m; Para Aro, RMSEP 2,4 % m/m; Para Res, RMSEP 1,4 % m/m e para Asf, RMSEP 1,0 % m/m. Novamente, apesar de os resultados serem melhores, a variabilidade das amostras usada pelos autores é menor com isso é mais fácil encontrar uma solução linear que se ajuste aos dados.

É mais notável, porém, o déficit de estudos para o estudo do SARA com espectroscopia NIR. Isso ocorre, pois, a região de impressão digital está contida no MIR (entre 1750 a 650 cm^{-1}) e é a mais apropriada para estudo dessas propriedades, dado que contém bandas de absorção mais intensas e com menos sobreposição referentes as cadeias alifáticas e aromáticas (RILEY *et al.*, 2016). Assim, compreende-se os modelos Aro – tanto para FT-NIR I, quanto para Micronir I – terem baixa correlação ($R^2_p < 0,5$), pois a região de maior importância não está contida nos espectros e a informação presente dentro da faixa NIR é sobreposta por outros sinais. Também por isso a resposta de Sat e Res apresentam correlação moderada para FT-

NIR I e insatisfatória para Micronir I – que ainda por cima não contém boa parte da região do primeiro sobretom e combinações das ligações CH ($6.200\text{--}4.000\text{ cm}^{-1}$).

Os altos valores de absorvância na região entre 11.013 a 5.967 cm^{-1} estão relacionados a presença de compostos poliaromáticos, especialmente asfaltenos – principalmente acima de 10.000 cm^{-1} . (MULLINS, 1990; PASQUINI; BUENO, 2007) Assim, de forma oposta a estimativa das outras frações (Sat, Aro e Res), o modelo Asf para Micronir I (R^2_p 0,72 e RMSEP 0,34 % m/m) apresentou por esse motivo o melhor desempenho em comparação com FT-NIR-I (R^2_p 0,60 e RMSEP 0,42 % m/m). Para todos outros seis modelos PLS, os modelos Micronir I obtiveram R^2_p entre 0,33 a 0,80, enquanto FT-NIR I apresentou R^2_p entre 0,44 e 0,94. A **Figura 6.3** e **Figura 9.1** demonstram esse ponto com as curvas de calibração para esses modelos PLS. Na **Figura 6.3.(a-c)** e **Figura 9.1.(a-d)** observa-se o ajuste dos dados a linha de referência, há maior variação entre os dados dos modelos Micronir I do que nos dados FT-NIR I (**Figura 6.3.(d-f)** e **Figura 9.1.(e-h)**). (DA SILVA, Neirivaldo Cavalcante *et al.*, 2017b) apresenta resultados similares. Os autores utilizaram um espectrômetro NIR portátil e de bancada para estimar propriedades físico-químicas em gasolinas e obtiveram para os modelos com instrumento portátil R^2_p entre 0,65 e 0,97, enquanto para o equipamento de bancada R^2_p 0,80 – 0,99. Ainda assim, os autores pontuam a relevância da portabilidade e que os resultados são, então, satisfatórios.

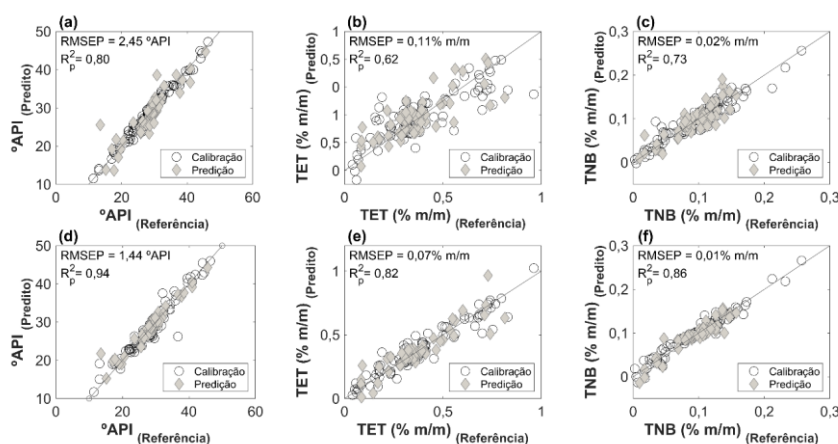


Figura 6.3. Curvas de calibração dos modelos PLS °API, TET e TNB, respectivamente, para (a-c) Micronir I e (d-f) FT-NIR I

A diferença no desempenho poderia estar relacionada às diferenças entre faixa e resolução espectral dos equipamentos. Então, conforme descrito na metodologia, construiu-se os conjuntos Micronir II, FT-NIR II e FT-NIR III que apresentam a mesma faixa espectral 10.000 – 6.000 cm^{-1} . Além disso, Micronir II e FT-NIR III têm os mesmos 110 pontos por espectro – FT-NIR II tem 2.001 pontos. O desempenho dos modelos PLS para esses três conjuntos de dados estão apresentados na **Tabela 6.3**.

Tabela 6.3. Parâmetros de desempenho dos modelos PLS para estimar sete propriedades físico-químicas em petróleos com os conjuntos de dados NIR na faixa espectral de 10.000 cm^{-1} a 6.000 cm^{-1} (Micronir II, FT-NIR II e FT-NIR III).

Micronir II					FT-NIR II				FT-NIR III			
Prop	Pre	RMSEP	R^2_p	VL	Pre	RMSEP	R^2_p	VL	Pre	RMSEP	R^2_p	VL
°API	Center	2,01	0,86	20	SNV+Der	3,10	0,68	7	Der+SNV	2,49	0,80	20
TET	Der+SNV	0,11	0,60	6	Center	0,11	0,64	12	Center	0,10	0,66	8
TNB	SNV	0,02	0,79	10	Center	0,02	0,63	12	Der+SNV	0,02	0,72	9
Sat	SNV	6,48	0,49	10	Der+SNV	6,60	0,47	3	Center	7,04	0,40	8
Aro	Der+SNV	3,61	0,36	9	SNV+Der	4,06	0,19	7	Center	3,88	0,27	8
Res	Center	5,59	0,36	13	Der	6,01	0,28	6	Center	5,52	0,37	13
Asf	SNV+Der	0,40	0,59	2	Der	0,46	0,49	5	Der	0,43	0,51	5

Como já mencionado, a região entre 11.000 e 10.000 cm^{-1} está relacionada aos grupos poliaromáticos. A remoção dessa região afetou fortemente o modelo PLS para a propriedade Asf. Para os três conjuntos (Micronir II: RMSEP 0,40% m/m e R^2_p 0,59; FT-NIR II: RMSEP 0,46% m/m e R^2_p 0,49; FT-NIR III RMSEP 0,43% m/m e R^2_p 0,51) observa-se desempenho menor do que os discutidos anteriormente com os conjuntos Micronir I (RMSEP 0,34 % m/m e R^2_p 0,72) e FT-NIR I (RMSEP 0,42 % m/m e R^2_p 0,60). Além disso, observa-se que o desempenho foi semelhante para esses três conjuntos. Ou seja, a resolução spectral não afetou os modelos como o corte de parte da faixa espectral.

O oposto é observado em Micronir II para estimar °API e Sat. O modelo °API apresentou R^2_p 0,86 e RMSEP 2.01% m/m, enquanto o modelo Sat apresentou R^2_p 0,49 e RMSEP 6.48 % m/m. Como os grupos poliaromáticos causam espalhamento da luz com um declive/deslocamento na linha de base, removendo parte dessa região há a diminuição da interferência causada aos sobretons das ligações CH ao longo da faixa espectral (PASQUINI; BUENO, 2007). Para os outros modelos Micronir II: TET (R^2_p 0,60 e RMSEP 0,11 % m/m), TNB (R^2_p 0,79 e RMSEP 0,02 % m/m), Aro (R^2_p 0,36

e RMSEP 3.61 % m/m) e Res (R^2_p 0,36 e RMSEP 5.59 % m/m) não houve diferenças expressivas pela remoção dessa região. O mesmo não pode ser dito para os outros seis modelos, exceto Asf, obtidos em FT-NIR II e FT-NIR III. A remoção da faixa entre 6.000 cm^{-1} e 4.000 cm^{-1} piorou todos os modelos quando comparados aos resultados de FT-NIR I. Isso ocorre pois nessa região estão o primeiro sobretom SH, segundo sobretom C=O, primeiro sobretom CH e combinações dos modos vibracionais das ligações OH, NH, C-C e CH (XIAOBO *et al.*, 2010).

A **Figura 6.4** mostra os modelos PLS para a propriedade TET (% m/m) para cada um dos conjuntos de dados (Micronir I, Micronir II, FT-NIR I, FT-NIR II, FT-NIR III, respectivamente). **Figura 6.4.(a,b)** mostra uma distribuição dos dados similar, Micronir I (RMSEP 0,11 % m/m e R^2_p 0,62) e Micronir II (RMSEP 0,11 % m/m e R^2_p 0,60), respectivamente. O melhor ajuste está em **Figura 6.4.(c)** para FT-NIR I, RMSEP 0,07 % m/m e R^2_p 0,82. Com a remoção da faixa espectral entre 6.000 cm^{-1} e 4.000 cm^{-1} os modelos para FT-NIR II (RMSEP 0,11 % m/m e R^2_p 0,64) e FT-NIR III (RMSEP 0,10 % m/m e R^2_p 0,66), **Figura 6.4.(d)** e **Figura 6.4.(e)**, tem exatidão similar a Micronir I e Micronir II. Assim, novamente, isso demonstra que a piora no desempenho dos modelos está relacionada a faixa espectral e não à resolução. Isso é coerente pois o primeiro sobretom SH, como já mencionado, está localizado na faixa 6.200 a 5.500 cm^{-1} .

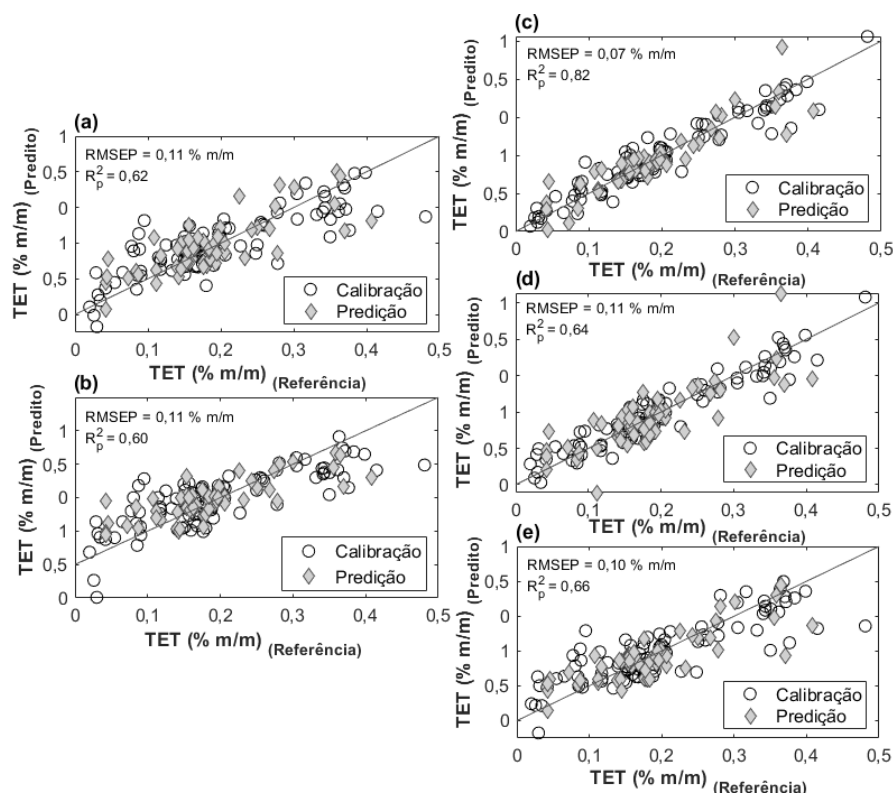


Figura 6.4. Curvas de calibração dos modelos PLS para estimar TET (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.

Figura 9.2 a Figura 9.7 contêm a comparação entre todos conjuntos para cada uma das outras seis propriedades físico-químicas (°API, TNB, Sat, Aro, Res e Asf, respectivamente). Em todos os casos, os piores resultados foram obtidos por FT-NIR II. Essa piora nos modelos informa que a região entre 6.000 a 4.000 cm^{-1} contém variáveis importantes sendo perdidas. Ainda mais interessante, porém, é que os modelos FT-NIR III têm melhor desempenho do que FT-NIR II. Ou seja, uma resolução spectral não necessariamente contribui para a exatidão do modelo. Esse é o princípio dos algoritmos de seleção de variáveis, um conjunto de variáveis menor que aprimora o desempenho dos modelos por remover variáveis não informativas e minimizar ruídos. (XIAOBO *et al.*, 2010) Além disso, as figuras também mostram que os modelos FT-NIR III têm desempenho semelhante aos modelos Micronir II. Isso demonstra que a construção do dispositivo portátil é realmente robusta e competitiva.

Por fim, como vimos, os resultados dos modelos PLS para os dados Micronir I não apresentaram forte correlação para TET, Sat, Aro e Res. Assim, construiu-se modelos SVR com esse mesmo conjunto de dados, pois em trabalhos anteriores o algoritmo SVR foi indicado capturar melhor a variabilidade de amostras de petróleos. (BALABIN; LOMAKINA, 2011; FILGUEIRAS *et al.*, 2016; SANTOS *et al.*, 2021) Os resultados desses modelos estão na **Tabela 6.4**. Todos os modelos SVR (R^2_p 0,46 – 0,90) aprimoraram o desempenho da estimativa das propriedades físico-químicas em comparação com os modelos PLS, R^2_p 0,33 – 0,80.

Tabela 6.4. Parâmetros de desempenho dos modelos de regressão (PLS e SVR) para a estimativa das propriedades físico-químicas construídos com os dados Micronir I.

Prop	PLS			SVR		
	Pré-trat	RMSEP	R^2_p	Pré-trat	RMSEP	R^2_p
°API	Center	2,45 °API	0,80	SNV + Der	1,69 °API	0,90
TET	Der	0,11% m/m	0,62	SNV + Der	0,06% m/m	0,75
TNB	Der	0,02% m/m	0,73	SNV + Der	0,01% m/m	0,81
Sat	SNV+Der	7,37% m/m	0,34	Der	5,07% m/m	0,62
Aro	Der	3,69% m/m	0,34	Der	3,18% m/m	0,50
Res	Center	5,80% m/m	0,33	SNV + Der	3,57% m/m	0,46
Asf	Center	0,34% m/m	0,72	SNV	0,32	0,76

Individualmente comparando cada modelo obtido com estudos da literatura, começando com TET e TNB temos o estudo de (FOLLI *et al.*, 2020). Os autores usaram modelos SVR e SVR combinado com algoritmo de busca angular e fator de inflação de variância (ASA-VIF-SVR) e obtiveram para TET com SVR R^2_p 0,43 e RMSEP 0,11 % m/m, enquanto para ASA-VIF-SVR R^2_p 0,55 e RMSEP 0,09 % m/m. O nosso modelo SVR para TET obteve melhor desempenho, R^2_p 0,75 e RMSEP 0,06 % m/m. Para TNB, os modelos SVR e ASA-VIF-SVR obtiveram, respectivamente, R^2_p 0,77, R^2_p 0,71 e RMSEP 0,02 % m/m para ambos modelos. Enquanto o nosso modelo SVR para TNB também é melhor, R^2_p 0,81 e RMSEP 0,01 % m/m.

Para as outras propriedades (°API, Sat, Aro, Res e Asf - SARA) não há estudos com petróleos e dados NIR, assim a comparação será feita com estudos considerados mais próximos que utilizam a faixa MIR. Em (FILGUEIRAS *et al.*, 2014), os autores

estimaram °API de 68 petróleos usando um espectrômetro de bancada e modelo SVR obteve R^2_p 0,98 e RMSEP 0,25, enquanto nosso modelo SVR apresentou R^2_p 0,90 e RMSEP 1,7. Para SARA, (MOHAMMADI *et al.*, 2021) construíram modelos SVR combinados com algoritmo genético com 12 amostras de petróleo. Os modelos obtiveram R^2_p 0,8 – 0,94 e RMSEP 0,78 – 2.8 % m/m, enquanto nossos modelos SARA R^2_p 0,46 – 0,76 e RMSEP 0,32 – 5.1 % m/m. Como já mencionado, a região de impressão digital (1.750 a 650 cm^{-1}) contém bandas de absorção de cadeias alifáticas e aromáticas mais intensas e com menos sobreposição, por isso não é surpreendente que os resultados sejam melhores que os nossos. Porém, apesar de nossos resultados não serem superiores tanto para °API, quanto para SARA, eles são estatisticamente satisfatórios ($R^2_p > 0,5$). Além de enfatizar a robustez dos resultados pela variabilidade do conjunto amostral utilizado e pela inovação da utilização de um instrumento portátil – uma tecnologia que ainda está mais disponível em espectrômetros na faixa NIR, do que na faixa MIR.(CROCOMBE, 2018)

E com a utilização do algoritmo SVR a estimativa das propriedades com equipamento portátil (dados Micronir I) obtêm desempenhos muito similares aos obtidos com os modelos PLS para os dados do equipamento de bancada (FT-NIR I) - **Tabela 6.5.** Para Sat, Aro e Asf, respectivamente, os resultados Micronir I (RMSEP 5,07% m/m e R^2_p 0,62; RMSEP 3,18% m/m e R^2_p 0,50; RMSEP 0,32% m/m e R^2_p 0,76) foram melhores que FT-NIR I (RMSEP 5,77% m/m e R^2_p 0,58; RMSEP 3,37% m/m e R^2_p 0,44; RMSEP 0,42% m/m e R^2_p 0,60). Enquanto para °API, TET, TNB e Res, FT-NIR I (RMSEP 1,44 °API e R^2_p 0,94; RMSEP 0,07% m/m e R^2_p 0,82; RMSEP 0,01% m/m e R^2_p 0,86; RMSEP 4,84% m/m e R^2_p 0,55) ainda supera Micronir I (RMSEP 1,69 °API e R^2_p 0,90; RMSEP 0,06% m/m e R^2_p 0,75; RMSEP 0,01% m/m e R^2_p 0,81; RMSEP 3,57% m/m e R^2_p 0,46). Afinal, como já discutido, a região entre 6.000 a 4.000 cm^{-1} contém variáveis importantes que não estão disponíveis para a construção dos modelos com equipamento portátil. Mas esse problema é contornado com SVR pois sendo uma técnica de regressão não-linear consegue captar melhor a variabilidade de amostras complexas como petróleos (BALABIN; LOMAKINA, 2011a) .

Tabela 6.5. Parâmetros de desempenho dos modelos de regressão SVR (Micronir I) e PLS (FT-NIR I) para a estimativa das sete propriedades físico-químicas.

Micronir I (SVR)	FT-NIR I (PLS)
------------------	----------------

Prop	Pré-trat	RMSEP	R ² _p	Pré-trat	RMSEP	R ² _p
°API	SNV + Der	1,69 °API	0,90	SNV	1,44 °API	0,94
TET	SNV + Der	0,06% m/m	0,75	Der	0,07 % m/m	0,82
TNB	SNV + Der	0,01% m/m	0,81	Der	0,01 % m/m	0,86
Sat	Der	5,07% m/m	0,62	Der	5,77 % m/m	0,58
Aro	Der	3,18% m/m	0,50	SNV+Der	3,37 % m/m	0,44
Res	SNV + Der	3,57% m/m	0,46	Der	4,84 % m/m	0,55
Asf	SNV	0,32% m/m	0,76	Der	0,42 % m/m	0,60

6.4. Conclusão

O espectrômetro NIR portátil não adquire a faixa espectral entre 6.000 e 4.000 cm⁻¹, aonde estão localizadas as combinações e primeiros sobretons das ligações SH, NH, CH, OH importantes para construção dos modelos PLS para estimar as propriedades investigadas nesse estudo. Porém, com a utilização do modelo SVR esse problema é superado. Com isso o desempenho do equipamento se torna equivalente ao do equipamento de bancada. Deveras, o modelo para estimar Asf teve o melhor desempenho no equipamento portátil (RMSEP 0,32 % m/m e R²_p 0,76). Além disso, quando igualamos a faixa espectral (10.000 cm⁻¹ a 6.000 cm⁻¹) e resolução (110 pontos por espectro) o equipamento portátil supera o equipamento de bancada. Assim, a combinação do instrumento portátil com algoritmo SVR é a metodologia acertada para trabalhar com petróleos.

7. APLICAÇÃO 3

Discriminação de petróleos e combustíveis utilizando um espectrômetro NIR portátil

7.1. Introdução

A indústria de energia recebe um grande volume de dinheiro e transações em todo o mundo. Isso a torna alvo de grupos criminosos. (RAMALHO, 2019) Assim, investigações forenses de furtos, adulterações e vazamentos são necessárias para mitigar os problemas e tornar o processo seguro. A grande dificuldade está na complexidade das matrizes. Combustíveis e óleos fósseis são predominantemente hidrocarbonetos. Os óleos brutos são misturas complexas que podem conter centenas a milhares de compostos, enquanto os combustíveis fósseis são subprodutos obtidos por craqueamento térmico em diferentes faixas de temperatura (KHANMOHAMMADI; GARMARUDI; DE LA GUARDIA, 2012). No entanto, como as diferenças físico-químicas influenciam na qualidade final dos produtos, é muito comum misturar óleos brutos para obter produtos com as melhores qualidades. Assim, embora os óleos brutos de diferentes localidades tenham certas características distinguíveis (RILEY *et al.*, 2016b), esta informação pode tornar-se mais difícil de ler em misturas. Além disso, em casos de acidentes com vazamento, a resposta à ação deve ser rápida para que os danos causados não coloquem em risco a sociedade e o meio ambiente.

No Brasil, a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP, regulamenta a atividade das empresas autorizadas e os limites para cada propriedade importante nos produtos, a fim de manter um padrão de qualidade e segurança no processo e nos produtos (ANP, 2001). No entanto, as refinarias de reprocessamento de óleos usados compram óleos roubados e derivados, com o objetivo de obter produtos de melhor qualidade e vender combustíveis clandestinamente para obter maiores lucros. Esse mercado ilegal, nos últimos quatro anos, desviou 42 milhões de litros de petróleo bruto e derivados da rede de dutos no Brasil (RAMALHO, 2019).

Assim, a classificação de petróleos e combustíveis tem grande importância. Métodos cromatográficos como em Camada Delgada (em inglês, *thin layer* (TLC)), Líquida de Alta Performance (em inglês, *high-performance liquid* (HPLC)), gasosa (GC), e permeação a gel (em inglês, *gel permeation* (GPC)) são os métodos mais

utilizados para análise composicional de petróleo e combustíveis (VANINI *et al.*, 2018). A resposta dessas técnicas traz informação detalhada sobre a composição química de cada classe de hidrocarbonetos presente, a chamada impressão digital (COUTINHO *et al.*, 2018; RILEY *et al.*, 2016b; VANINI *et al.*, 2018). Entretanto, esses métodos são destrutivos, laboriosos e geram resíduos pela utilização de solventes químicos. Outras técnicas analíticas podem ser utilizadas para o mesmo propósito mantendo a confiabilidade e exatidão, como no caso das técnicas espectroscópicas Raman (FORTUNATO *et al.*, 2017), NIR (CARNEIRO *et al.*, 2008; DUPUY *et al.*, 2015; FERNEES *et al.*, 2008) e MIR (BAIRD; OJA, 2016; DALY *et al.*, 2016; DOS SANTOS, Victor Hugo J.M. *et al.*, 2017).

Entre essas, a espectroscopia NIR tem sido aplicada para o controle de qualidade de combustíveis como quantificação de água e metanol em amostras de etanol e gasolina, (CARNEIRO *et al.*, 2008; CHO *et al.*, 2005; FERNEES *et al.*, 2008; GUCHARDI; PASQUINI, 2001; SIVLA *et al.*, 2012) teor de biodiesel e teor de enxofre total em amostras de diesel, (ALVES; POPPI, 2013; CORREIA *et al.*, 2017; DA SIVLA *et al.*, 2017b; PAIVA *et al.*, 2015; PONTES *et al.*, 2011) e caracterização de propriedades físico-químicas como viscosidade, teor de enxofre, densidade, entre outros para diferentes combustíveis (BROUILLETTE *et al.*, 2016; DA SILVA, Vitor Hugo *et al.*, 2015; KHANMOHAMMADI; GARMARUDI; DE LA GUARDIA, 2012; LAXALDE *et al.*, 2014; ROCHA *et al.*, 2016). A espectroscopia MIR também já obteve sucesso ao ser aplicada (CROCOMBE, 2018; DALY *et al.*, 2016; KHANMOHAMMADI; GARMARUDI; DE LA GUARDIA, 2012; MALDONADO *et al.*, 2018; MAZIVILA, 2018; MOURA *et al.*, 2019). A região MIR apresenta bandas com maior razão sinal-ruído do que a NIR. Porém, espectrômetros MIR para análises em campo e portáteis ainda não são tão comuns pois apresentam elementos ópticos sensíveis. Além de que a interação no sistema NIR entre o feixe de luz e a amostra é mais efetiva (PASQUINI, 2018).

A portabilidade nos espectrômetros NIR é uma recente inovação. Algumas vantagens são alta frequência analítica, design ergonômico, baixo custo de operação, baixo consumo de energia e fácil manuseio com pouca ou nenhuma preparação de amostra (CROCOMBE, 2018; GAŁUSZKA; MIGASZEWSKI; NAMIEŚNIK, 2015; PASQUINI, 2018). Em menos de 10 segundos, os espectros NIR podem ser coletados e enviados pela internet, uma vantagem fascinante em casos de vazamentos e acidentes, por exemplo. Espectrômetros portáteis combinam portabilidade com uma

variedade de acessórios que permitem a análise de diferentes produtos (CORREIA *et al.*, 2018; CROCOMBE, 2018; PASQUINI, 2018).

Modelos quimiométricos de classificação e quantificação são úteis no estudo de matrizes complexas, pois permitem a afiliação de uma amostra desconhecida a um conjunto conhecido de amostras (GELADI, 2003). A técnica de espectroscopia portátil combinada com ferramentas quimiométricas tem sido uma poderosa ferramenta analítica para a indústria e no controle de qualidade de alimentos e produtos químicos (CRAIG; FRANCA; IRUDAYARAJ, 2015).

Diferentes métodos qualitativos multivariados têm sido desenvolvidos para a classificação de amostras ou controle de qualidade. Dentre eles, destacam-se os modelos supervisionados como análise discriminante por mínimos quadrados parciais (PLS-DA) e análise discriminante linear (LDA) (BALABIN; SAFIEVA, 2011; BALABIN; SAFIEVA; LOMAKINA, 2011; DA SIVLA *et al.*, 2014; GALTIER *et al.*, 2011; MILANEZ *et al.*, 2016; PONTES *et al.*, 2011) e não supervisionados, como a análise por componentes principais (PCA) e análise para agrupamentos hierárquicos (HCA) (BALABIN; SAFIEVA, 2011; BALABIN; SAFIEVA; LOMAKINA, 2010; BARBOSA *et al.*, 2013). Enquanto em estudos quantitativos, PLS é o método mais aplicado (KHANMOHAMMADI; GARMARUDI; DE LA GUARDIA, 2012; PANTOJA *et al.*, 2017), mas SVR, um método não linear, pode prover uma resposta mais adequada para matrizes complexas como petróleo (BALABIN; LOMAKINA, 2011b; DA SIVLA *et al.*, 2014; DEVOS *et al.*, 2009; FILGUEIRAS *et al.*, 2014; THISSEN *et al.*, 2004).

Modelos PLS construídos com dados obtidos em equipamentos portáteis mostraram resultados competitivos ou melhores aos obtidos em equipamentos de bancada (DA SIVLA *et al.*, 2017b; KIRCHLER *et al.*, 2017; MALEGORI *et al.*, 2017). Há publicações utilizando espectrômetros NIR portáteis para analisar diferentes combustíveis e parâmetros de caracterização (BROUILLETTE *et al.*, 2016; COOPER *et al.*, 2011; CORREIA *et al.*, 2017; DA SIVLA *et al.*, 2017b; DE ARAUJO *et al.*, 2018; DOUGLAS *et al.*, 2018; LUTZ *et al.*, 2014; MÁQUINA *et al.*, 2019; PAIVA *et al.*, 2019). Assim, um espectrômetro NIR portátil foi utilizado neste estudo para resolver problemas de classificação e quantificação em óleos brutos e combustíveis. Para a classificação foi utilizado o modelo PLS-DA e os parâmetros de desempenho avaliados foram sensibilidade, especificidade e precisão. As amostras de petróleo foram discriminadas de óleo de motor usado (visualmente semelhante a um petróleo e por isso um potencial adulterante). Os combustíveis foram classificados de acordo

com seu tipo (nafta, gasolina, diesel e querosene). Na seção de quantificação, foram construídos modelos PLS e SVR e comparados para quantificar misturas ternárias (variando de 0 a 80 % m/m) de óleo leve (densidade API 56,0 °API), óleo pesado (densidade API 19,9 °API) e óleo de motor (densidade API de 29,1 °API). Os parâmetros utilizados para esta avaliação foram o erro quadrático médio de calibração (RMSEC), erro quadrático médio de predição (RMSEP), coeficiente de determinação dos conjuntos de calibração (R^2_c) e coeficiente de determinação dos conjuntos de predição (R^2_p). Por fim, também foram construídos modelos PLS para quantificar misturas binárias de combustíveis (nafta e gasolina, diesel e querosene) variando de 0 a 100 % m/m o volume de cada componente da mistura.

7.2. Metodologia

Todas amostras de petróleos e combustíveis foram cedidas pelo Cenpes e Transpetro. Portanto, referem-se a amostras dos campos de produção do território brasileiro. O óleo de motor usado é um resíduo retirado da troca de óleo dos carros e armazenado em barris de metal para descarte ambiental correto. As amostras de óleo de motor usado foram doadas por postos de gasolina locais.

7.2.1. Espectroscopia NIR

Os espectros NIR foram adquiridos com um espectrômetro portátil modelo MicroNIR™ Pro 1700 com software 3.0 da Viavi Solutions. Alguns parâmetros de análise: faixa espectral de 908 a 1676 nm, 200 varreduras, tempo de integração de 8 ms e aquisição de 125 pontos por espectro. Mais informações sobre o funcionamento desse equipamento podem ser encontradas em (O'BRIEN *et al.*, 2012). Para a análise de petróleos foi utilizado o acessório do equipamento *Gota* e para combustíveis *Vial Holder*. Ambos acessórios são conectados sobre a fonte de luz, de forma ao feixe incidir sob a amostra para coleta dos dados. O espectro de referência para 100% de reflectância foi obtido utilizando Spectralon® e para 0% de reflectância utilizou-se a cela vazia e as luzes auxiliares apagadas. Esses espectros de referência foram coletados a cada dez minutos corridos de análise.

7.2.2. Modelo qualitativo

Dois conjuntos de amostras foram estudados: o primeiro contendo 275 amostras de petróleos e 70 amostras de óleo de motor usado (denominado Óleo de motor). As amostras foram preparadas em triplicata e os espectros NIR foram adquiridos em ordem aleatória, amostrando uma gota na janela de safira do acessório. Usamos a média da triplicata para construção dos modelos. Portanto, um conjunto de dados $X_{345 \times 125}$. O segundo conjunto de dados foi formado com amostras de nafta, gasolina, diesel e querosene transferidas para frascos de vidro transparente com capacidade volumétrica de 4 mL. Cada conjunto de combustível continha 40 frascos cheios. Logo, essa segunda matriz tinha dimensão $X_{160 \times 125}$.

Todos os espectros foram pré-processados usando a primeira derivada pelo algoritmo de Savitzky-Golay (SAVITZKY; GOLAY, 1964) com polinômio de segunda ordem e janela de sete pontos. Para ambos os conjuntos de amostras, os espectros NIR foram divididos em conjuntos de calibração e teste para construir o modelo e avaliar a capacidade discriminatória. O conjunto de calibração continha 70% de todas as amostras e os 30% restantes formaram o conjunto de dados de teste. O modelo PLS-DA foi avaliado quanto à sensibilidade, especificidade e precisão, respectivamente (BALLABIO; CONSONNI, 2013a, b).

7.2.3. Modelo quantitativo

Para avaliar a capacidade dos em quantificar o óleo de motor usado adicionado em amostras de petróleo foi construído uma curva de calibração contendo uma mistura de três componentes: uma amostra do conjunto Óleo de motor e duas amostras de petróleo, a amostra com densidade API de 56,0 °API foi denominada Óleo leve e outra com densidade API de 19,9 °API denominada de Óleo pesado. A curva analítica elaborada teve as proporções, em massa, variadas nas misturas de 0 a 100 % m/m, gerando 53 amostras. As medidas foram realizadas em triplicata, totalizando 159 espectros NIR adquiridos.

Para a quantificação do combustível, dois modelos de calibração multivariada foram construídos a partir de misturas binárias variando o teor de 0 a 100 % m/m em

um intervalo de 5-5 % m/m. O primeiro modelo era composto por nafta e gasolina e o segundo por diesel e querosene.

Diferentes pré-tratamentos foram testados nos espectros NIR: primeira derivada com um polinômio de segunda ordem e uma janela de 7 pontos (SAVITZKY; GOLAY, 1964), airPLS (ZHANG; CHEN; LIANG, 2010), centralização, SNV e MSC. Nos resultados mostramos os melhores modelos e sinalizamos quais pré-tratamentos foram usados para tais modelos. Em seguida foram construídos os modelos PLS e SVR. Todos cálculos foram executados com o software Matlab R2013a (Eigenvector Research).

As amostras foram separadas pelo método k-fold em conjuntos de dados de calibração e teste para cobrir todas as faixas percentuais das misturas. Na construção dos modelos PLS, o número de variáveis latentes foi otimizado. E nos modelos SVR, os parâmetros custo (C), margem (ϵ) e gama (γ) da função kernel foram otimizados. A otimização dos modelos foi realizada pelo procedimento de validação cruzada de 5-fold. Os modelos foram avaliados pelas métricas RMSEC, RMSEP, R^2_c e R^2_p . (MARTENS; NAES, 1990)

7.3. Resultados e discussão

7.3.1. Conjunto de dados 1: Petróleos e Óleo de motor

7.3.1.1. Modelo qualitativo

O espectro médio para o conjunto de petróleos e óleo de motor é mostrado na

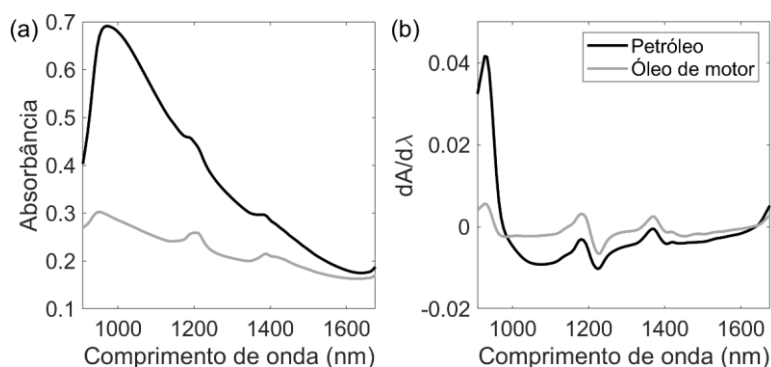


Figura 7.1. É possível observar uma grande variação da linha de base espectral nos dados não tratados (

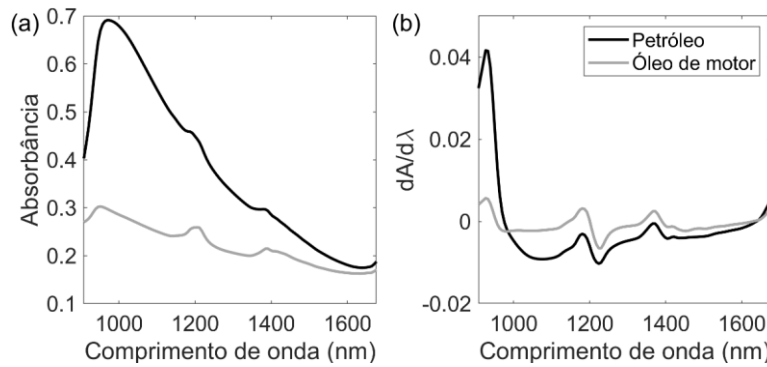


Figura 7.1.a) que ocorre devido ao efeito de espalhamento da radiação (PANTOJA *et al.*, 2017). Ele é corrigido após aplicação da primeira derivada (

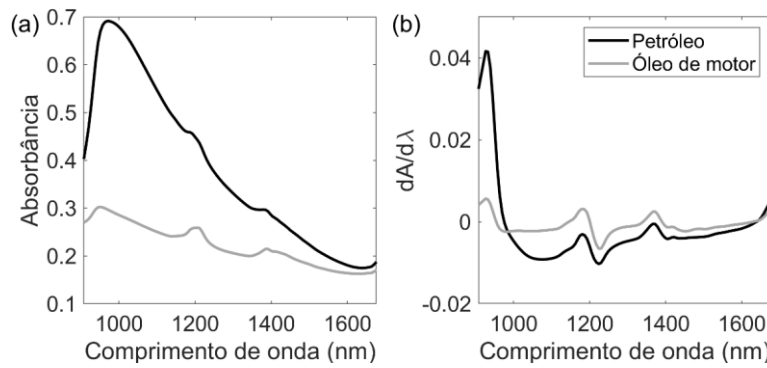


Figura 7.1.b).

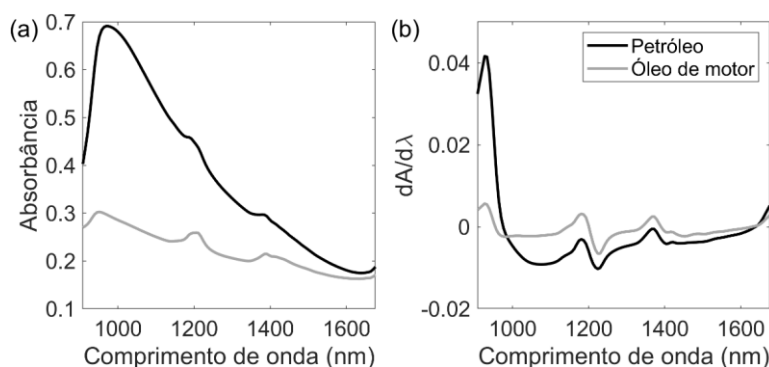


Figura 7.1. Espectro NIR médio dos petróleos e óleo de motor **(a)** bruto e **(b)** pré-tratado com primeira derivada

Na **Figura 7.1.b**, é possível ver bandas de absorção em torno de 900 nm (segundo sobretom OH), 1200 nm (Segundo sobretom de CH₂) e 1400 nm (combinações e primeiro sobretom CH₂) que se relacionam a natureza majoritariamente composta por hidrocarbonetos para essas matrizes (CORREA PABÓN; SOUZA FILHO, 2019). A região com maior distinção entre petróleo e óleo de motor é a da faixa entre 900 e 1000 nm. Essa região é influenciada por asfaltenos (LAXALDE *et al.*, 2014; PASQUINI; BUENO, 2007). Entretanto, quando muitos espectros de petróleos e óleos de motor são sobrepostos essa clara distinção se perde pela variação entre espectros.

O PLS-DA identifica as variáveis latentes que permitem a maior discriminação entre dois grupos distintos de amostras com base em seus espectros (matriz **X**) e de acordo com sua covariância máxima para uma classe alvo estabelecida na matriz **Y** (BALLABIO; CONSONNI, 2013b). Para minimizar os erros foi selecionado o número de VL, de acordo com a taxa de erro de classificação, prevista pelo processo de validação cruzada, neste caso, nove VL.

O modelo PLS-DA foi aplicado ao conjunto de teste e a probabilidade de cada amostra pertencer a cada uma das classes é mostrada na **Figura 7.2.a,b**. A linha pontilhada horizontal refere-se ao limite de decisão para a amostra pertencer a determinada classe. As amostras dos conjuntos de calibração e teste são representadas pelos símbolos (o) e (Δ), respectivamente.

Os números presentes em algumas amostras (**Figura 7.2**) referem-se às amostras classificadas erroneamente, 53 e 115. Essas amostras foram classificadas fora da classe Petróleo (**Figura 7.2.a**) e dentro da classe Óleo de motor (**Figura 7.2.b**). Eles não podem ser considerados *outliers*, pois não possuem altos valores de resíduos. Estas amostras têm densidade API acima de 30 °API e, portanto, baixo teor de asfaltenos. No entanto, o conjunto de petróleos apresenta muitas outras amostras com características semelhantes. Assim, deve haver outras características de composição dessas amostras que se assemelhem à sua resposta aos óleos de motor usados.

Em contraste, nenhuma das amostras de óleo de motor usado foi classificada incorretamente, **Figura 7.2**. Dessa forma, o espectrômetro portátil é capaz de discriminar amostras de petróleo e de óleo de motor usado.

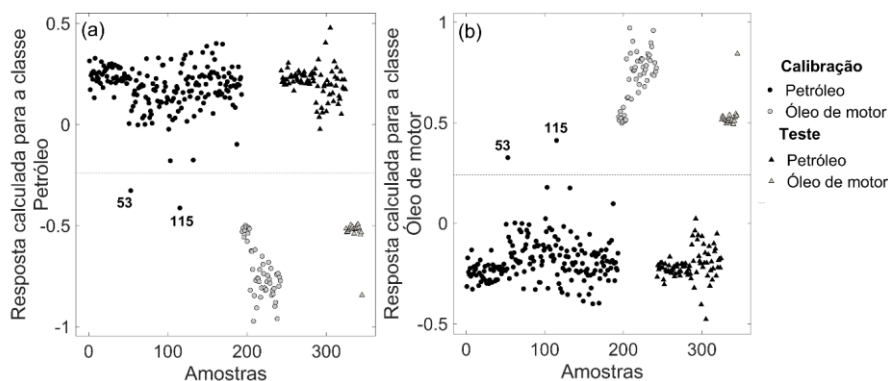


Figura 7.2. Resposta calculada para a probabilidade das amostras pertencerem a classe (a) Petróleo, (b) Óleo de Motor dos modelos PLS-DA.

Os parâmetros de desempenho do modelo PLS-DA são mostrados na **Tabela 7.1**. Não há referências usando um espectrômetro portátil na região do NIR para discriminar a origem de amostras de petróleo bruto ou para diferenciar produtos que podem ser usados para confundir a composição química do petróleo bruto óleo (como no caso do óleo de motor usado no estudo). Mas usando um instrumento de bancada e analisando a impressão digital da região MIR-FTIR (BAGHERI GARMARUDI *et al.*, 2019) classificou 251 amostras de sete campos iranianos e três tipos de misturas de petróleo bruto com modelagem independente suave de analogia de classe (SIMCA)

Comentado [FD1]: Ver se tenho o gráfico de resíduos

com precisão de 96,7%, RMSEP de 1,41% m/m e R^2 de 0,957.(RILEY *et al.*, 2016a) obtiveram espectros MIR de asfaltenos de 11 amostras (oito de petróleos de diferentes localizações e duas de óleos combustíveis pesados). Foi possível diferenciar todas as amostras com comparações de razão de altura de pico com desvio relativo de 20% como limite de decisão para avaliar a semelhança amostral. (GALTIER *et al.*, 2011) classificaram a origem de 36 amostras de petróleo de quatro regiões diferentes usando espectros MIR combinados aos modelos SIMCA, PLS-DA e PLS-DA-SIMCA. Todos os modelos obtiveram 100% de classificação correta.

Em comparação com os resultados das publicações mencionadas e ressaltando que nosso estudo foi realizado com um espectrômetro portátil, a sensibilidade do modelo PLS-DA do petróleo teve capacidade discriminatória semelhante. Devido às duas amostras de petróleo classificadas erroneamente no conjunto de calibração, a sensibilidade do petróleo bruto está abaixo de 100% na calibração e validação cruzada. Isso leva à perda de especificidade e precisão para óleo de motor. No entanto, cumpre o objetivo de discriminar óleos de motor usados de petróleos com o modelo apresentando sensibilidade de 100%.

Tabela 7.1. Parâmetros de desempenho do modelo PLS-DA para as classes Petróleo e Óleo de motor.

	Parâmetros	Petróleo	Óleo de motor
Calibração	Especificidade	100%	99%
	Sensibilidade	99%	100%
	Precisão	100%	96%
Validação cruzada	Especificidade	100%	98%
	Sensibilidade	98%	100%
	Precisão	100%	94%
Teste	Especificidade	100%	100%
	Sensibilidade	100%	100%
	Precisão	100%	100%

7.3.1.2. Modelo quantitativo

Os espectros para as 53 amostras, compostas por uma mistura ternária de óleo leve, óleo pesado e óleo de motor variando o teor de 0 a 100% m/m, são mostrados na **Figura 7.3.a**. Para o modelo PLS, foram utilizadas seis variáveis latentes e a centralização na média provou ser a melhor metodologia para correção de espectros (**Figura 7.3.b**), enquanto para o modelo SVR, a melhor foi a primeira derivada seguida

de normalização (**Figura 7.3.c**) . Como pode ser visto na **Figura 7.3** todas as amostras têm o mesmo perfil e não é possível distingui-las apenas observando os espectros.

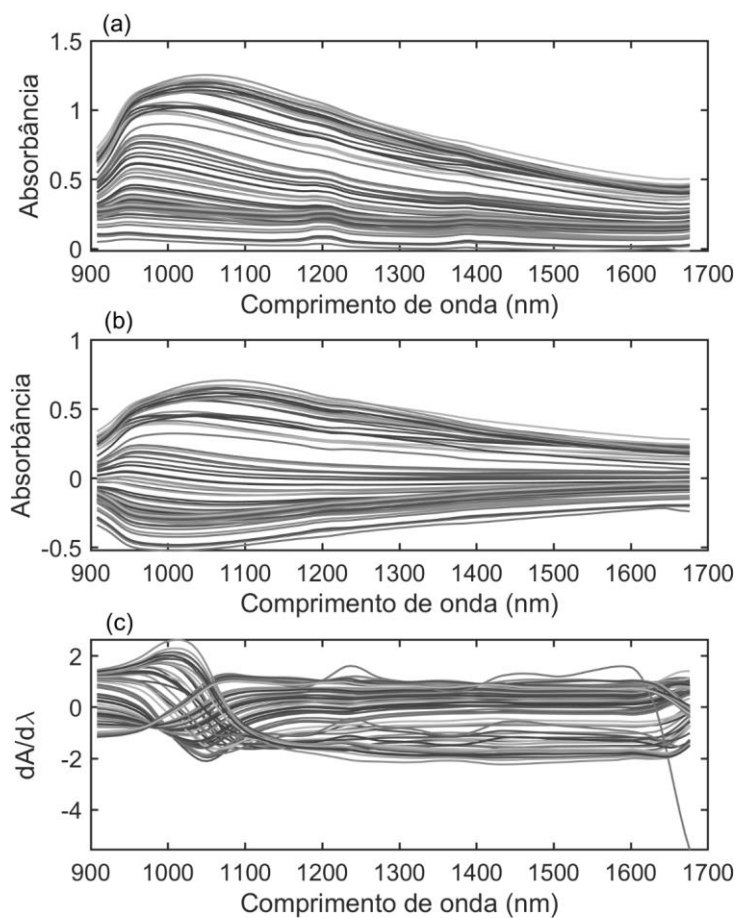


Figura 7.3. Espectros NIR dos 53 *blends* (a) sem tratamento e pré-tratados com (b) centralização e (c) primeira derivada e normalização.

Os resultados dos modelos com os melhores valores médios de precisão são apresentados na **Tabela 7.2**. Para Óleo Leve e Pesado, os melhores resultados foram obtidos com o modelo SVR. No caso do Óleo Leve, os valores de RMSEC, RMSEP,

R^2_c e R^2_p foram 5,1 % m/m, 4,0 % m/m, 0,97 e 0,99, respectivamente. Por outro lado, melhores valores foram obtidos para o Óleo Pesado (RMSEC = 2,7 % m/m; RMSEP = 0,9 % m/m, $R^2_c = 0,99$ e $R^2_p = 0,99$). Por fim, para Óleo de Motor, pode-se dizer que o PLS é o melhor modelo, pois o desempenho foi muito similar entre PLS e SVR, mas o modelo PLS é uma ferramenta muito mais simples de utilizar com mais trabalhos na literatura pra comparação (BALABIN; LOMAKINA, 2011b). Os modelos PLS apresentaram erros inferiores a 10% m/m (valor do incremento a cada amostra) com coeficiente de correlação superior a 0,96, portanto, os dados estão bem ajustados e por isso recomenda-se utilizar este modelo para atender este tipo de aplicação.

(CIPULLO *et al.*, 2019) avaliaram 21 regiões de cinco solos contaminados com petróleo, um total de 105 amostras, e estimaram a quantidade total e biodisponível (mg/kg) de compostos hidrocarbonetos de petróleo. Os espectros de infravermelho próximo-visível (VIS-NIR) foram obtidos na faixa de 350-2500 nm e o modelo aplicado foi RF, os resultados para R^2 ficaram entre 0,40 e 0,75. (DOUGLAS *et al.*, 2018) também utilizou equipamento VIS-NIR para prever a quantidade total de hidrocarbonetos de petróleo (mg/kg) em 85 amostras de solo. Foram construídos os modelos RF e PLS, os resultados de R^2 encontrados foram 0,68 e 0,54, respectivamente. Um valor de RMSEP abaixo de 15% m/m indica que o modelo é preciso (CIPULLO *et al.*, 2019). Dessa forma, nossos modelos são precisos e ainda que não seja exatamente a mesma problemática construída por esses estudos citados e o nosso, nossos resultados são superiores.

Tabela 7.2. Parâmetros de desempenho para os modelos SVR e PLS para quantificar óleo leve, óleo pesado e óleo de motor nas misturas ternárias.

Componente	Pré-tratamento	Modelo	RMSEC (% m/m)	RMSEP (% m/m)	R^2_c	R^2_p
Óleo leve	Der + Norm	SVR	5,1	4,0	0,97	0,99
Óleo leve	Centralização	PLS (6)*	6,1	6,2	0,96	0,98
Óleo pesado	Der + Norm	SVR	2,7	0,9	0,99	0,99
Óleo pesado	Centralização	PLS (6)	4,6	2,9	0,98	0,98
Óleo de motor	Der + Norm	SVR	5,3	6,2	0,97	0,99
Óleo de motor	Centralização	PLS (6)	5,9	4,8	0,97	0,99

*Valor entre parênteses se refere ao número de VL.

A **Figura 7.4** mostra os valores de referência e estimados dos modelos PLS (**Figura 7.4.a,c,e**) e SVR (**Figura 7.4.b,d,f**) para o teor de Óleo leve (**Figura 7.4.a,b**),

Óleo pesado (**Figura 7.4.c,d**) e Óleo de motor (**Figura 7.4.e,f**). Resumidamente, nenhuma diferença pode ser observada entre os resultados dos modelos PLS (à esquerda) e SVR (à direita), mas observa-se bom ajuste dos dados para ambos modelos. Porém, como visto na **Tabela 7.2** os parâmetros de desempenho (RMSE e R^2) foram melhores para os modelos SVR, quando comparados aos modelos PLS. (ALVES; POPPI, 2013) avaliaram amostras de diesel, onde os modelos SVR tiveram valores de RMSEP variando de 21 a 54% abaixo dos valores obtidos nos modelos PLS. De forma semelhante, em nosso estudo o modelo SVR foi superior ao PLS. No entanto, quando os parâmetros de desempenho obtidos entre o PLS e SVR são próximos, é melhor optar por utilizar o PLS.

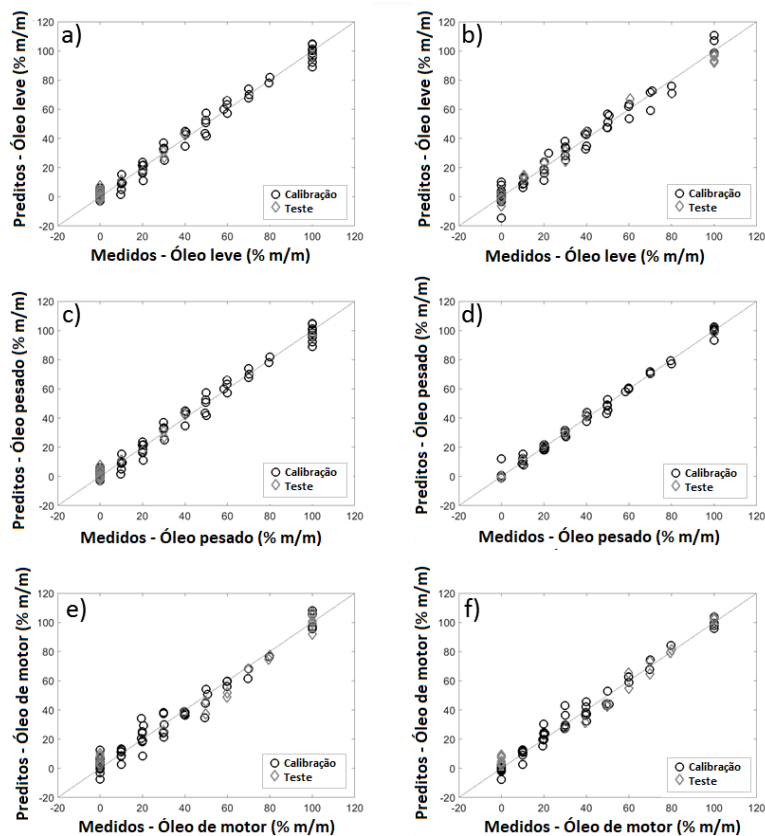


Figura 7.4. Curvas de calibração dos modelos (a,c,e) PLS e (b,c,f) SVR para estimar os teores de (a,b) Óleo leve, (c,d) Óleo pesado, (e,f) Óleo de motor.

7.3.2. Conjunto de dados 2: Combustíveis

7.3.2.1. Modelo qualitativo

Os espectros para os quatro combustíveis (nafta, gasolina, diesel e querosene) apresentam grande similaridades (**Figura 7.5**). Como a diferença entre esses produtos está na faixa de temperatura de ebulição no processo de destilação, há similaridade composicional, o que explica a similaridade nos espectros (BROUILLETTE *et al.*, 2016; SPEIGHT, 2002). Depois de pré-tratados com a primeira derivada, algumas regiões espectrais apresentam diferenças sutis: em torno de 1250 e 1510 nm. Estas bandas estão relacionadas às combinações e segundo sobretom CH, respectivamente (BROUILLETTE *et al.*, 2016).

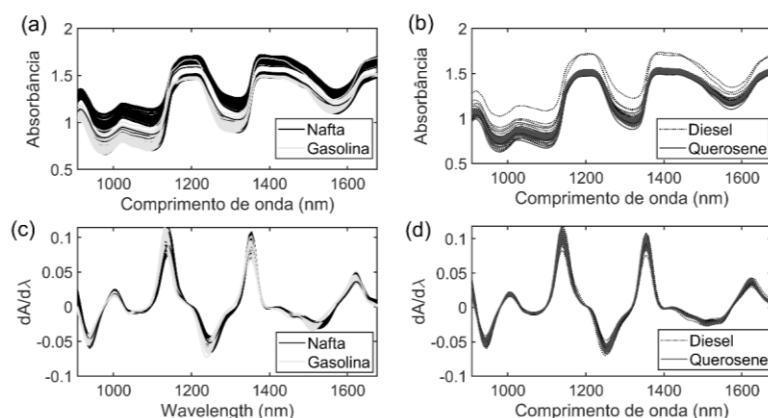


Figura 7.5. Espectros (a, b) brutos e (c, d) pré-tratados com primeira derivada para (a, c) nafta e gasolina e (b, d) diesel e querosene.

A especificidade, sensibilidade e precisão foi 100% para o modelo PLS-DA, não houve nenhuma amostra sendo erroneamente classificada para nenhuma das quatro classes (

Tabela 7.3).

Sobre estudos na literatura, para a classificação com PLS-DA de combustíveis. Para gasolinas, (BALABIN; SAFIEVA; LOMAKINA, 2010) usaram um espectrômetro NIR de bancada para três problemas de classificação de gasolinas: processadas em três diferentes refinarias, obtidas por seis diferentes modos de processamento e tipos

de gasolina. Os erros de classificação de cada um desses três modelos utilizando o PLS-DA foram 9%, 20% e 10%, respectivamente. Enquanto, (DA SILVA *et al.*, 2014) usou PLS-DA para classificar 125 amostras de gasolina em dois grupos: com ou sem aditivos e apenas 2 amostras com aditivos foram classificadas errado.

Para amostras de lubrificante, (BALABIN; SAFIEVA; LOMAKINA, 2011) usaram um espectrômetro NIR de bancada para três problemas de classificação tipos de lubrificantes, e para dois parâmetros de caracterização. Os erros de classificação para cada um desses modelos foi 36%, 59% e 52%.

Para amostras de diesel, (PONTES *et al.*, 2011) utilizou o modelo PLS-DA para o controle de qualidade desse combustível com adição de biodiesel, ou seja, classificando-as como conformes ou não-conformes. Eles utilizaram 140 amostras e o erro de classificação foi de 3%.

Por fim, apesar de não utilizar espectroscopia NIR, pontuamos o trabalho de (MALDONADO *et al.*, 2018) por ser o único utilizando um equipamento portátil, no caso utiliza um espectrômetro MIR. Outra semelhança está no problema que também era a classificação de diferentes combustíveis (gasolina, diesel e diferentes tipos de álcool). Eles não utilizam PLS-DA, utilizam a análise de variância multivariada e obtiveram 100% de classificação correta.

Com base nesses estudos, observa-se que nossos resultados foram superiores e a consistência da técnica NIR para resolução desse tipo de problema, a qual tem sido cada vez mais implementada em plantas industriais (BALABIN; LOMAKINA, 2011b; PANTOJA *et al.*, 2017). Os trabalhos de Balabin, Safieva e Lomakina (2010 e 2011) apesar de apresentarem os maiores erros têm importância por terem trabalhado em escala industrial, seus conjuntos amostrais apresentam maior variabilidade e por isso os modelos lineares como PLS-DA não conseguem um bom ajuste (BALABIN; LOMAKINA, 2011b). Para o trabalho de Maldonado e colaboradores (2018) é principalmente interessante a utilização de um equipamento portátil, porém, a faixa MIR apresenta bandas de absorção intensas e com pouca sobreposição, por isso podem ser mais facilmente correlacionados aos grupos funcionais. Mas os espectrômetros NIR ainda são mais comuns e robustos para utilizações in situ e em linha (CROCOMBE, 2018).

Tabela 7.3. Parâmetros de desempenho para o modelo PLS-DA para as classes Nafta, Gasolina, Diesel e Querosene.

	Parâmetros	Nafta	Gasolina	Diesel	Querosene
Calibração	Especificidade	100%	100%	100%	100%
	Sensibilidade	100%	100%	100%	100%
	Precisão	100%	100%	100%	100%
Validação cruzada	Especificidade	100%	100%	100%	100%
	Sensibilidade	100%	100%	100%	100%
	Precisão	100%	100%	100%	100%
Teste	Especificidade	100%	100%	100%	100%
	Sensibilidade	100%	100%	100%	100%
	Precisão	100%	100%	100%	100%

A **Figura 7.6** mostra visualmente que nenhuma das amostras foi classificada incorretamente. Mas além disso, observa-se a distância entre as naftas (**Figura 7.6.a**) e gasolinas (**Figura 7.6.b**) para os combustíveis que não pertencem à classe referente. Apesar de os espectros desses quatro combustíveis, como vimos, não poderem ser distintos, o algoritmo PLS-DA maximiza as diferenças sem dificuldade para, pois o perfil entre amostras da mesma classe é bem definido. Já para diesel (**Figura 7.6.c**) e querosene (**Figura 7.6.d**) observa-se maior dispersão entre as amostras, ou seja, há maior variabilidade nesses conjuntos amostrais. Dessa forma, se aumentássemos o conjunto amostral provavelmente começaríamos a ver erros de classificação (BALABIN; LOMAKINA, 2011b).

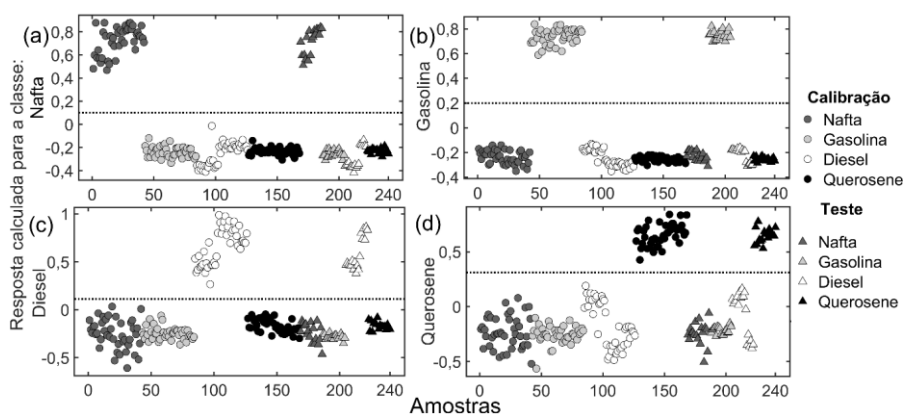


Figura 7.6. Resposta calculada para a probabilidade das amostras pertecerem a cada classe **(a)** nafta, **(b)** gasolina, **(c)** diesel, **(d)** querosene para o modelo PLS-DA com 5 VL.

7.3.2.2. Modelo quantitativo

Foram preparadas misturas de nafta e gasolina e diesel e querosene. 20 blends para cada conjunto, variando o teor entre amostras em 5 % m/m, atingindo uma faixa de 0 a 100 % m/m. O problema de quantificação foi construído definindo nafta e querosene como adulterantes aos outros dois combustíveis, uma extrapolação coerente com a realidade pois adulterantes em combustíveis, em geral, são frações mais baratas (VEMPATAPU; KANAUIA, 2017). Os valores de RMSEC e RMSEP foram semelhantes para ambos os modelos, com valores de RMSEP inferiores a três vezes maiores que o RMSEC. Esses resultados indicam que não há sobreajuste dos modelos PLS (**Tabela 7.4**). Os modelos apresentaram linearidade acima de 90% na estimativa das amostras (**Figura 7.7**). As amostras com 0% de nafta e querosene foram previstas com teor abaixo dos valores medidos. Esse erro pode estar relacionado à similaridade química desses combustíveis, intrínseca ao processo de destilação.

Tabela 7.4. Parâmetros de desempenho para os modelos PLS para estimar teor de nafta em gasolina e querosene em diesel.

Parâmetros	Nafta	Querosene
RMSEC (% m/m)	2,9	31,7
RMSEP (% m/m)	1,4	11,4
R ² _c	0,99	0,92
R ² _p	0,99	0,94
LD (% m/m)	1,3	2,8
LQ (% m/m)	4,4	9,3

Esses resultados mostram que é possível detectar nafta na gasolina acima de 1,3% m/m e quantificar acima de 4,4% m/m com RMSEP de 1,4% m/m. Para as misturas de diesel com querosene, é possível detectar e quantificar acima de 2,8 e 9,3 % m/m, respectivamente, com RMSEP de 11,4 % m/m.

Embora nenhum estudo quantifique as mesmas misturas de combustível, é importante observar os estudos que utilizaram espectrômetros NIR portáteis para quantificar o teor de combustível. Para amostras de diesel, (PAIVA *et al.*, 2015) obteve LD de 0,31% v/v para teor de biodiesel e RMSEP 0,22 % v/v LD de 0,40% v/v para teor de óleos vegetais no diesel com RMSEP de 0,34 % v/v. No entanto, eles avaliaram uma faixa estreita de concentração (0-15% v/v). Já (CORREIA *et al.*, 2017) avaliaram 133 amostras de biodiesel/diesel na faixa de 0-100 % m/m e valores obtidos para prever a concentração de biodiesel de LD = 0,5 % m/m, LQ = 1,8 % m/m e

RMSEP = 1,8 % m/m. Para amostras de gasolina, (LUTZ *et al.*, 2014) obtiveram LD e LQ para concentração de etanol na gasolina de 0,68 e 2,04 % m/m, respectivamente, com RMSEP de 0,21 % m/m em misturas contendo 0-10 % m/m. Enquanto, (CORREIA *et al.*, 2017) resultados para gasolina, etanol e concentrações de metanol em amostras de gasolina (com faixa de concentração de 0,0–75,0 % m/m gasolina, 25,0–75,0 % m/m etanol e 0,0–50,0 % m/m metanol), respectivamente foram LD de 0,55% m/m, 0,75% m/m, 0,85% m/m; LQ de 1,84% m/m, 2,5% m/m, 2,84% m/m; e RMSEP de 0,81% m/m, 3,81% m/m e 1,80% m/m. Os resultados para o modelo construído para prever o teor de nafta estão próximos de outras publicações. O modelo de previsão do diesel obteve resultados superiores, mas as misturas aqui estudadas apresentam maior composição química. Ainda são valores interessantes. Caso estes modelos sejam utilizados em casos de adulteração de combustível, são aplicáveis. Porque a adulteração acontece com a adição de grandes volumes de adulterante.

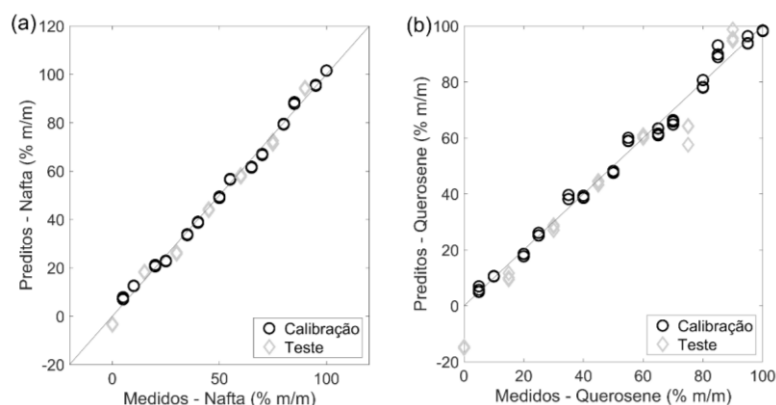


Figura 7.7. Curvas de calibração dos modelos PLS para estimativa do teor de **(a)** nafta e **(b)** querosene construídos com 4 VL em um intervalo de confiança de 95%.

7.4. Conclusão

O espectrômetro NIR portátil é eficiente para classificar e quantificar amostras de petróleo e seus derivados. O modelo PLS-DA conseguiu discriminar petróleos de óleos de motor usados com sensibilidade, especificidade e precisão acima de 94%. O modelo PLS também foi eficiente em quantificar o teor de óleo de motor de uma mistura ternária com dois petróleos e óleo de motor usado (Óleo leve, Óleo pesado e

Óleo de motor) com precisão média de 4,8 % m/m. Sobre os combustíveis, nafta, gasolina, diesel e querosene, o PLS-DA permite a identificação destes produtos com taxa de classificação de 100%. Para as misturas de gasolina e nafta, LD é 1,3% m/m e LQ 4,4% m/m com RMSEP 1,4% m/m. Da mesma forma, para as misturas de diesel e querosene, o modelo PLS permite a quantificação do querosene com LD 2,8 % m/m, LQ 9,3 % m/m com RMSEP de 11,4 % m/m.

8. CONCLUSÕES

Reportagens

- [1] Disponível em: <https://www.valor.com.br/brasil/6310595/furto-de-combustiveis-alimenta-ganhos-de-refinarias-clandestinas> . Acessado em 14/07/2020
- [2] Disponível em: <https://g1.globo.com/mg/sul-de-minas/noticia/2018/08/30/operacao-do-mp-apura-furto-e-venda-de-gasolina-no-sul-de-mg-prejuizo-chega-a-r-25-milhoes.ghtml> . . Acessado em 14/07/2020
- [3] Disponível em: https://www.em.com.br/app/noticia/gerais/2018/08/31/interna_gerais.984863/prejuizo-com-desvio-de-gasolina-da-regap-e-estimado-em-r-4-milhoes.shtml . Acessado em 14/07/2020
- [4] Disponível em: <https://www.metrojornal.com.br/foco/2019/06/08/quadrilha-combustivel-tunel-van.html> . Acessado em 14/07/2020,
- [5] Disponível em: <https://recordtv.r7.com/jornal-da-record/videos/furtos-de-combustivel-em-dutos-da-petrobras-triplicam-em-tres-anos-03062019>. Acessado em 14/07/2020,
- [6] Disponível em: <https://odia.ig.com.br/rio-de-janeiro/2019/06/5652123-policia-prende-envoVLidos-em-esquema-de-roubo-de-combustivel-na-baixada.html#foto=1> . Acessado em 14/07/2020
- [7] Disponível em: <https://www.metropoles.com/distrito-federal/apos-furto-de-combustivel-transpetro-descarta-desabastecimento-no-df> . Acessado em 14/07/2020,
- [8] Disponível em: <https://g1.globo.com/mundo/noticia/2019/01/19/oleoduto-explode-deixeio-mortos-e-feridos-no-mexico.ghtml> . Acessado em 14/07/2020,
- [9] Disponível em: <https://g1.globo.com/rj/rio-de-janeiro/noticia/2018/12/09/tentativa-de-furto-de-oleo-em-duto-em-mage-provoca-vazamento-em-rio-e-chega-a-manguezal.ghtml> . Acessado em 14/07/2020,
- [10] Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2019-11/operacao-prende-cinco-pessoas-por-furto-de-petroleo> . Acessado em 14/07/2020,

- [11] Disponível em: <https://g1.globo.com/sp/piracicaba-regiao/noticia/2020/06/26/operacao-cumpre-meados-contradulteracao-e-furto-de-combustiveis-na-regiao-de-limeira.ghtml> . Acessado em 14/07/2020,
- [12] Disponível em: <https://www.istoedinheiro.com.br/policia-desarticula-esquema-de-furto-de-combustivel-na-baixada-fluminense/> . Acessado em 14/07/2020,
- [13] Disponível em: <https://www.folhavoria.com.br/policia/noticia/03/2020/lider-de-quadrilha-especializada-em-furto-de-dutos-de-petroleo-e-presos-em-vila-velha>. Acessado em 14/07/2020
- [14] Disponível em: <https://g1.globo.com/sp/mogi-das-cruzes-suzano/noticia/2019/12/26/cetesb-vistoria-area-em-guararema-onde-houve-vazamento-de-petroleo-durante-furto.ghtml> . Acessado em 14/07/2020
- [15] Disponível em: <https://www.jornalcontabil.com.br/petrobras-reduz-em-31-perda-de-combustivel-por-furto-em-dutos/> . Acessado em 14/07/2020
- [16] Disponível em: <https://clickpetroleoegas.com.br/petrobras-descobre-tunel-de-14-metros-para-furto-de-combustivel-em-campinas-sp/>. Acessado em 14/07/2020
- [17] Disponível em: <https://noticias.r7.com/economia/transpetro-estanca-vazamento-de-duto-a-vl-o-de-furto-de-petroleo-em-rio-das-ostras-23092019>. Acessado em 14/07/2020
- [18] Disponível em: <https://epbr.com.br/christino-aureo-quer-tipificar-roubo-de-dutos-como-crime-contrordem-economica/> . Acessado em 14/07/2020
- [19] Disponível em: <https://economia.estadao.com.br/noticias/negocios,petrobras-cria-grupo-interno-para-enfrentar-os-ladros-de-combustivel,70003018551>. Acessado em 14/07/2020
- [20] Disponível em: <https://economia.estadao.com.br/noticias/negocios,petrobras-cria-grupo-interno-para-enfrentar-os-ladros-de-combustivel,70003018551>. Acessado em 14/07/2020
- [21] Disponível em: <https://noticias.r7.com/sao-paulo/policia-descobre-esquema-de-furto-de-combustivel-em-cacapava-sp-29112019> . Acessado em 14/07/2020

[22] Disponível em: <https://g1.globo.com/pr/norte-noroeste/noticia/2019/11/05/policia-investiga-se-empresa-de-roleia-recebeu-parte-do-petroleo-furtado-no-rio-de-janeiro.ghtml>. Acessado em 14/07/2020

[23] Disponível em: <https://www.reporterdiario.com.br/noticia/2733057/tentativa-de-furto-em-dutos-da-petrobras-deixa-parque-sao-bernardo-em-panico/> . Acessado em 14/07/2020

[24] Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2019/08/28/policia-descobre-quadrilha-que-planejava-furtar-produto-derivado-do-petroleo-na-zona-leste-de-sp.ghtml> . Acessado em 14/07/2020

[25] Disponível em: <https://g1.globo.com/rj/rio-de-janeiro/noticia/2020/12/04/operacao-mira-furto-de-combustivel-em-dutos-da-transpetro-e-da-petrobras.ghtml>

[26] Disponível em: <https://g1.globo.com/rj/rio-de-janeiro/noticia/2021/02/02/operacao-mira-quadrilha-de-furto-de-combustiveis-de-dutos-no-rj.ghtml>

[27] Disponível em: <https://g1.globo.com/rj/rio-de-janeiro/noticia/2021/03/02/policia-e-mpri-realizam-operacao-contra-quadrilha-especializada-em-furto-de-petroleo.ghtml>

Comentado [FD2]: Apagar

9. REFERÊNCIAS BIBLIOGRÁFICAS

ABNEY, W. W.; FESTING, E. R. On the influence of the atomic grouping in the molecules of organic bodies on their absorption in the infra-red region of the spectrum. **Royal Society**, vol. 172, p. 887–918, 31 Dec. 1881. DOI 10.1098/rstl.1881.0020. Available at: <https://royalsocietypublishing.org/doi/10.1098/rstl.1881.0020>.

ADETUNJI, J.; DRONSFIELD, A.; MORRIS, P. Infrared Spectroscopy – Why did it take so long ? **Education in Chemistry**, vol. 42, no. 2, p. 50–53, 2015. .

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**, vol. 104, p. 155–161, Jan. 2013. DOI 10.1016/j.talanta.2012.11.033. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0039914012009770>.

ANDERSSON, M. A comparison of nine PLS1 algorithms. **Journal of Chemometrics**, vol. 23, no. 10, p. 518–529, 2009. <https://doi.org/10.1002/cem.1248>.

ANP. **PANP 316 - 2001**. Brazil: [s. n.], 2001. Available at: <http://legislacao.anp.gov.br/>.

ARIA, M.; CUCCURULLO, C. bibliometrix: An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, vol. 11, no. 4, p. 959–975, 2017. DOI 10.1016/j.joi.2017.08.007. Available at: <https://doi.org/10.1016/j.joi.2017.08.007>.

ASKE, N.; KALLEVIK, H.; SJÖBLOM, J. Determination of saturate, aromatic, resin, and asphaltenic (SARA) components in crude oils by means of infrared and near-infrared spectroscopy. **Energy and Fuels**, vol. 15, no. 5, p. 1304–1312, 2001. <https://doi.org/10.1021/ef010088h>.

ASTM. **Symposium on Composition of Petroleum Oils, Determination and Evaluation**. West Conshohocken: ASTM International, 1958. DOI 10.1520/STP224-EB. Available at: <http://www.astm.org/doiLink.cgi?STP224-EB>.

ASTM E1655-05. Standard Practices for Infrared Multivariate Quantitative Analysis. **ASTM International**, vol. 05, no. Reapproved 2012, p. 29, 2012. DOI 10.1520/E1655-17.Copyright. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Standard+Practices+for+Infrared+Multivariate+Quantitative+Analysis#0>.

ASTM INTERNATIONAL. **ASTM 1898-1998: a century of progress**. [S. l.: s. n.], 1998. vol. 26, . Available at: https://www.astm.org/ABOUT/history_book.html.

ASTM INTERNATIONAL. ASTM D2549. Standard Test Method for Separation of Representative Aromatics and Nonaromatics Fractions of High-Boiling Oils by Elution Chromatography. **Annual Book of ASTM Standards**, 2002. .

ASTM INTERNATIONAL. ASTM D4294. Standard Test Method for Sulfur in Petroleum and Petroleum Products by Energy Dispersive X-Ray Fluorescence Spectrometry. **Annual Book of ASTM Standards**, 2008. .

ASTM INTERNATIONAL. ASTM D5854-19a. Standard Practice for Mixing and Handling of Liquid Samples of Petroleum and Petroleum Products. **Annual Book of ASTM Standards**, vol. 96, p. 1–22, 2019. <https://doi.org/10.1520/D5854-19A>.

BAGHERI GARMARUDI, A.; KHANMOHAMMADI, M.; GHAFORI FARD, H.; DE LA GUARDIA, M. Origin based classification of crude oils by infrared spectrometry and chemometrics. **Fuel**, vol. 236, no. November 2016, p. 1093–1099, 2019. DOI 10.1016/j.fuel.2018.09.013. Available at: <https://doi.org/10.1016/j.fuel.2018.09.013>.

BAIRD, Z. S.; OJA, V. Predicting fuel properties using chemometrics: a review and an extension to temperature dependent physical properties by using infrared spectroscopy to predict density. **Chemometrics and Intelligent Laboratory Systems**, vol. 158, p. 41–47, 2016. DOI 10.1016/j.chemolab.2016.08.004. Available at: <http://dx.doi.org/10.1016/j.chemolab.2016.08.004>.

BALABIN, R. M.; LOMAKINA, E. I. Support vector machine regression (SVR/LS-SVM)-an alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. **Analyst (Cambridge, United Kingdom)**, vol. 136, no. 8, p. 1703–1712, 2011a. <https://doi.org/10.1039/c0an00387e>.

BALABIN, R. M.; LOMAKINA, E. I. Support vector machine regression (SVR/LS-SVM)

- An alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. **Analyst**, vol. 136, no. 8, p. 1703–1712, 2011b. <https://doi.org/10.1039/c0an00387e>.

BALABIN, R. M.; LOMAKINA, E. I.; SAFIEVA, R. Z. Neural network (ANN) approach to biodiesel analysis: Analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy. **Fuel**, vol. 90, no. 5, p. 2007–2015, 2011. <https://doi.org/10.1016/j.fuel.2010.11.038>.

BALABIN, R. M.; SAFIEVA, R. Z. Biodiesel classification by base stock type (vegetable oil) using near infrared spectroscopy data. **Analytica Chimica Acta**, vol. 689, no. 2, p. 190–197, 2011. DOI 10.1016/j.aca.2011.01.041. Available at: <http://dx.doi.org/10.1016/j.aca.2011.01.041>.

BALABIN, R. M.; SAFIEVA, R. Z.; LOMAKINA, E. I. Gasoline classification using near infrared (NIR) spectroscopy data: Comparison of multivariate techniques. **Analytica Chimica Acta**, vol. 671, no. 1–2, p. 27–35, 2010. DOI 10.1016/j.aca.2010.05.013. Available at: <http://dx.doi.org/10.1016/j.aca.2010.05.013>.

BALABIN, R. M.; SAFIEVA, R. Z.; LOMAKINA, E. I. Near-infrared (NIR) spectroscopy for motor oil classification: From discriminant analysis to support vector machines. **Microchemical Journal**, vol. 98, no. 1, p. 121–128, 2011. <https://doi.org/10.1016/j.microc.2010.12.007>.

BALABIN, R. M.; SMIRNOV, S. V. Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data. **Analytica Chimica Acta**, vol. 692, no. 1–2, p. 63–72, 2011. <https://doi.org/10.1016/j.aca.2011.03.006>.

BALLABIO, D.; CONSONNI, V. Classification toolbox for MATLAB. 2013a. Available at: <http://www.michem.unimib.it/download/matlab-toolboxes/classification-toolbox-for->

matlab.

BALLABIO, D.; CONSONNI, V. Classification tools in chemistry. Part 1: linear models. PLS-DA. **Analytical Methods**, vol. 5, no. 16, p. 3790, 2013b. DOI 10.1039/c3ay40582f. Available at: <http://xlink.rsc.org/?DOI=c3ay40582f>.

BALLABIO, D.; CONSONNI, V. Classification tools in chemistry. Part 1: Linear models. PLS-DA. **Analytical Methods**, vol. 5, no. 16, p. 3790–3798, 2013c. <https://doi.org/10.1039/c3ay40582f>.

BARBOSA, L. L.; SAD, C. M. S.; MORGAN, V. G.; SANTOS, M. F. P.; CASTRO, E. V. R. Time-domain proton nuclear magnetic resonance and chemometrics for identification and classification of Brazilian petroleum. **Energy and Fuels**, vol. 27, no. 11, p. 6560–6566, 2013. <https://doi.org/10.1021/ef4015313>.

BRATTAIN, R. R.; RASMUSSEN, R. S.; CRAVATH, A. M. A spectrophotometric method for the analysis of multicomponent mixtures and its infra-red application. **Journal of Applied Physics**, vol. 14, no. 8, p. 418–428, 1943. <https://doi.org/10.1063/1.1715009>.

BROUILLETTE, C.; SMITH, W.; SHENDE, C.; GLADDING, Z.; FARQUHARSON, S.; MORRIS, R. E.; CRAMER, J. A.; SCHMITIGAL, J. Analysis of Twenty-Two Performance Properties of Diesel, Gasoline, and Jet Fuels Using a Field-Portable Near-Infrared (NIR) Analyzer. **Applied Spectroscopy**, vol. 70, no. 5, p. 746–755, 2016. <https://doi.org/10.1177/0003702816638279>.

BROWN, C. W.; LYNCH, P. F.; AHMADJIAN, M. Applications of Infrared Spectroscopy in Petroleum Analysis and Oil Spill Identification. **Applied Spectroscopy Reviews**, vol. 9, no. 1, p. 223–248, 1975. <https://doi.org/10.1080/05704927508081491>.

BUERCK, J.; ROTH, S.; KRAEMER, K.; SCHOLZ, M.; KLAAS, N. Application of a fiber-optic NIR-EFA sensor system for in situ monitoring of aromatic hydrocarbons in contaminated groundwater. **Journal of Hazardous Materials**, vol. 83, no. 1–2, p. 11–28, 2001. [https://doi.org/10.1016/S0304-3894\(00\)00335-6](https://doi.org/10.1016/S0304-3894(00)00335-6).

BURNS, D. A.; CIURCZAK, E. W. Handbook of Near-Infrared Analysis, Third Edition. , p. 836, 2007. Available at: <https://books.google.ru/books?id=6EEd1a0uka0C>.

CAÓ, V. M.; ROMÃO, W.; CORREIA, R. M.; FILGUEIRAS, P. R.; DOMINGOS, E.; TOSATO, F.; FONTES, A. M.; AQUINO, L. F. M. Banknote analysis by portable near infrared spectroscopy. **Forensic Chemistry**, vol. 8, p. 57–63, 2018. DOI 10.1016/j.forc.2018.02.003. Available at: <https://doi.org/10.1016/j.forc.2018.02.003>.

CAPITÁN-VALLVEY, L. F.; PALMA, A. J. Recent developments in handheld and portable optosensing-A review. **Analytica Chimica Acta**, vol. 696, no. 1–2, p. 27–46, 2011. <https://doi.org/10.1016/j.aca.2011.04.005>.

CARNEIRO, H. S. P.; MEDEIROS, A. R. B.; OLIVEIRA, F. C. C.; AGUIAR, G. H. M.; RUBIM, J. C.; SUAREZ, P. A. Z. Determination of ethanol fuel adulteration by methanol using partial least-squares models based on Fourier transform techniques. **Energy and Fuels**, vol. 22, no. 4, p. 2767–2770, 2008. <https://doi.org/10.1021/ef8000218>.

CAVALCANTE DA SILVA, N.; CARIBÉ DE GÓES MASSA, A. R.; DOMINGOS, D.; AMIGO, J. M.; DAS VIRGENS REBOUÇAS, M.; PASQUINI, C.; PIMENTEL, M. F. NIR-based octane rating simulator for use in gasoline compounding processes. **Fuel**, vol. 243, no. July 2018, p. 381–389, 2019. <https://doi.org/10.1016/j.fuel.2019.01.145>.

CHAKRABORTY, S.; WEINDORF, D. C.; MORGAN, C. L. S.; GE, Y.; GALBRAITH, J. M.; LI, B.; KAHN, C. S. Rapid Identification of Oil-Contaminated Soils Using Visible Near-Infrared Diffuse Reflectance Spectroscopy. **Journal of Environmental Quality**,

vol. 39, no. 4, p. 1378–1387, 2010. <https://doi.org/10.2134/jeq2010.0183>.

CHO, S.; CHUNG, H.; WOO, Y. A.; KIM, H. J. Determination of water content in ethanol by miniaturized near-infrared (NIR) system. **Bulletin of the Korean Chemical Society**, vol. 26, no. 1, p. 115–118, 2005. <https://doi.org/10.5012/bkcs.2005.26.1.115>.

CHRISTISON, R. IX. Chemical Examination of the Petroleum of Rangoon. **Transactions of the Royal Society of Edinburgh**, vol. 13, no. 1, p. 118–123, 17 Jan. 1835. DOI 10.1017/S0080456800022225. Available at: https://www.cambridge.org/core/product/identifier/S0080456800022225/type/journal_article.

CHUNG, H.; CHOI, H. J.; KU, M. S. Rapid identification of petroleum products by near-infrared spectroscopy. **Bulletin of the Korean Chemical Society**, vol. 20, no. 9, p. 1021–1025, 1999. .

CIPULLO, S.; NAWAR, S.; MOUAZEN, A. M.; CAMPO-MORENO, P.; COULON, F. Predicting bioavailability change of complex chemical mixtures in contaminated soils using visible and near-infrared spectroscopy and random forest regression. **Scientific Reports**, vol. 9, no. 1, p. 1–12, 2019. DOI 10.1038/s41598-019-41161-w. Available at: <http://dx.doi.org/10.1038/s41598-019-41161-w>.

COOPER, J. B.; LARKIN, C. M.; SCHMITIGAL, J.; MORRIS, R. E.; ABDELKADER, M. F. Rapid analysis of jet fuel using a handheld near-infrared (NIR) analyzer. **Applied Spectroscopy**, vol. 65, no. 2, p. 187–192, 2011. <https://doi.org/10.1366/10-06076>.

CORREA PABÓN, R. E.; SOUZA FILHO, C. R. de. Crude oil spectral signatures and empirical models to derive API gravity. **Fuel**, vol. 237, no. October 2018, p. 1119–1131, 2019. DOI 10.1016/j.fuel.2018.09.098. Available at: <https://doi.org/10.1016/j.fuel.2018.09.098>.

CORREA PABÓN, R. E.; SOUZA FILHO, C. R. de; OLIVEIRA, W. J. de. Reflectance and imaging spectroscopy applied to detection of petroleum hydrocarbon pollution in bare soils. **Science of The Total Environment**, vol. 649, p. 1224–1236, Feb. 2019.

DOI 10.1016/j.scitotenv.2018.08.231. Available at:
<https://linkinghub.elsevier.com/retrieve/pii/S0048969718331954>.

CORREIA, R. M.; ROMÃO, W.; FERREIRA, E. C.; AQUINO, L. F. M.; SENA, S.; ARAUJO, B. R. F.; DOMINGOS, E.; FILGUEIRAS, P. R.; CÁO, V. M.; FONTES, A. M.; PINHEIRO, L. U. Portable near infrared spectroscopy applied to fuel quality control.

Talanta, vol. 176, no. August 2017, p. 26–33, 2017. DOI 10.1016/j.talanta.2017.07.094. Available at:
<http://dx.doi.org/10.1016/j.talanta.2017.07.094>.

CORREIA, R. M.; TOSATO, F.; DOMINGOS, E.; RODRIGUES, R. R. T.; AQUINO, L. F. M.; FILGUEIRAS, P. R.; LACERDA, V.; ROMÃO, W. Portable near infrared spectroscopy applied to quality control of Brazilian coffee. **Talanta**, vol. 176, no. April 2017, p. 59–68, 2018. <https://doi.org/10.1016/j.talanta.2017.08.009>.

COUTINHO, D. M.; FRANÇA, D.; VANINI, G.; MENDES, L. A. N.; GOMES, A. O.; PEREIRA, V. B.; ÁVILA, B. M. F.; AZEVEDO, D. A. Rapid hydrocarbon group-type semi-quantification in crude oils by comprehensive two-dimensional gas chromatography. **Fuel**, vol. 220, no. October 2017, p. 379–388, 2018. <https://doi.org/10.1016/j.fuel.2018.02.009>.

CRAIG, A. P.; FRANCA, A. S.; IRUDAYARAJ, J. Vibrational spectroscopy for food quality and safety screening. **High Throughput Screening for Food Safety Assessment: Biosensor Technologies, Hyperspectral Imaging and Practical Applications**. 1st ed. [S. l.]: Elsevier Ltd., 2015. p. 165–194. DOI 10.1016/B978-0-

85709-801-6.00007-1. Available at: <http://dx.doi.org/10.1016/B978-0-85709-801-6.00007-1>.

CROCOMBE, R. A. Portable Spectroscopy. **Applied Spectroscopy**, vol. 72, no. 12, p. 1701–1751, 2018. <https://doi.org/10.1177/0003702818809719>.

DA SILVA, M. P. F.; BRITO, L. R. E.; HONORATO, F. A.; PAIM, A. P. S.; PASQUINI, C.; PIMENTEL, M. F. Classification of gasoline as with or without dispersant and detergent additives using infrared spectroscopy and multivariate classification. **Fuel**, vol. 116, p. 151–157, 2014. <https://doi.org/10.1016/j.fuel.2013.07.110>.

DA SILVA, N. C.; CAVALCANTI, C. J.; HONORATO, F. A.; AMIGO, J. M.; PIMENTEL, M. F. Standardization from a benchtop to a handheld NIR spectrometer using mathematically mixed NIR spectra to determine fuel quality parameters. **Analytica Chimica Acta**, vol. 954, p. 32–42, 2017a. <https://doi.org/10.1016/j.aca.2016.12.018>.

DA SILVA, N. C.; CAVALCANTI, C. J.; HONORATO, F. A.; AMIGO, J. M.; PIMENTEL, M. F. Standardization from a benchtop to a handheld NIR spectrometer using mathematically mixed NIR spectra to determine fuel quality parameters. **Analytica Chimica Acta**, vol. 954, p. 32–42, Feb. 2017b. DOI 10.1016/j.aca.2016.12.018. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0003267016314921>.

DA SILVA, V. H.; REBOUCAS, M. V.; SALLES, A. R.; PIMENTEL, M. F.; PONTES, M. J. C.; PASQUINI, C. Determination of naphtha composition by near infrared spectroscopy and multivariate regression to control steam cracker processes. **Fuel Processing Technology**, vol. 131, p. 230–237, 2015. DOI 10.1016/j.fuproc.2014.10.035. Available at: <http://dx.doi.org/10.1016/j.fuproc.2014.10.035>.

DALY, S. R.; NIEMEYER, K. E.; CANNELLA, W. J.; HAGEN, C. L. Predicting fuel

research octane number using Fourier-transform infrared absorption spectra of neat hydrocarbons. **Fuel**, vol. 183, p. 359–365, Nov. 2016. DOI 10.1016/j.fuel.2016.06.097. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0016236116305567>.

DE ARAUJO, W. R.; CARDOSO, T. M. G.; DA ROCHA, R. G.; SANTANA, M. H. P.; MUÑOZ, R. A. A.; RICHTER, E. M.; PAIXÃO, T. R. L. C.; COLTRO, W. K. T. Portable analytical platforms for forensic chemistry: A review. **Analytica Chimica Acta**, vol. 1034, p. 1–21, 2018. <https://doi.org/10.1016/j.aca.2018.06.014>.

DE PAULO, E. H.; DOS SANTOS, F. D.; FOLLI, G. S.; SANTOS, L. P.; NASCIMENTO, M. H. C.; MORO, M. K.; DA CUNHA, P. H. P.; CASTRO, E. V. R.; CUNHA NETO, A.; FILGUEIRAS, P. R. Determination of gross calorific value in crude oil by variable selection methods applied to ¹³C NMR spectroscopy. **Fuel**, vol. 311, no. November 2021, p. 0–3, 2022. <https://doi.org/10.1016/j.fuel.2021.122527>.

DEIDDA, R.; SACRE, P. Y.; CLAVAUD, M.; COÏC, L.; AVOHOU, H.; HUBERT, P.; ZIEMONS, E. Vibrational spectroscopy in analysis of pharmaceuticals: Critical review of innovative portable and handheld NIR and Raman spectrophotometers. **TrAC - Trends in Analytical Chemistry**, vol. 114, p. 251–259, 2019. DOI 10.1016/j.trac.2019.02.035. Available at: <https://doi.org/10.1016/j.trac.2019.02.035>.

DEVOS, O.; RUCKEBUSCH, C.; DURAND, A.; DUPONCHEL, L.; HUVENNE, J. P. Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. **Chemometrics and Intelligent Laboratory Systems**, vol. 96, no. 1, p. 27–33, 2009. DOI 10.1016/j.chemolab.2008.11.005. Available at: <http://dx.doi.org/10.1016/j.chemolab.2008.11.005>.

DOS SANTOS, V. H. J. M.; RAMOS, A. S.; PIRES, J. P.; ENGELMANN, P. de M.;

LOUREGA, R. V.; KETZER, J. M. M.; RODRIGUES, L. F. Discriminant analysis of biodiesel fuel blends based on combined data from Fourier Transform Infrared Spectroscopy and stable carbon isotope analysis. **Chemometrics and Intelligent Laboratory Systems**, vol. 161, p. 70–78, Feb. 2017. DOI 10.1016/j.chemolab.2016.12.004. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0169743916302787>.

DOUGLAS, R. K.; NAWAR, S.; ALAMAR, M. C.; MOUAZEN, A. M.; COULON, F. Rapid prediction of total petroleum hydrocarbons concentration in contaminated soil using vis-NIR spectroscopy and regression techniques. **Science of the Total Environment**, vol. 616–617, p. 147–155, 2018. DOI 10.1016/j.scitotenv.2017.10.323. Available at: <https://doi.org/10.1016/j.scitotenv.2017.10.323>.

DUPUY, N.; BRAHEM, Z.; AMAT, S.; KISTER, J. Near-Infrared Spectroscopy Analysis of Heavy Fuel Oils Using a New Diffusing Support. **Applied Spectroscopy**, vol. 69, no. 10, p. 1137–1143, Oct. 2015. DOI 10.1366/14-07725. Available at: <http://journals.sagepub.com/doi/10.1366/14-07725>.

FALLA, F. S.; LARINI, C.; LE ROUX, G. A. C.; QUINA, F. H.; MORO, L. F. L.; NASCIMENTO, C. A. O. Characterization of crude petroleum by NIR. **Journal of Petroleum Science and Engineering**, vol. 51, no. 1–2, p. 127–137, 2006. <https://doi.org/10.1016/j.petrol.2005.11.014>.

FERNANDES, H. L.; RAIMUNDO JR, I. M.; PASQUINI, C.; ROHWEDDER, J. J. R. Simultaneous determination of methanol and ethanol in gasoline using NIR spectroscopy: Effect of gasoline composition. **Talanta**, vol. 75, no. 3, p. 804–810, 15 May 2008. DOI 10.1016/j.talanta.2007.12.025. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S003991400700879X>.

FERREIRA, M. M. C. **Quimiometria: conceitos, métodos e aplicações**. 1st ed. Campinas: Editora da Unicamp, 2015.

FILGUEIRAS, P. R.; PORTELA, N. A.; SILVA, S. R. C.; CASTRO, E. V. R.; OLIVEIRA, L. M. S. L.; DIAS, J. C. M.; NETO, A. C.; ROMÃO, W.; POPPI, R. J. Determination of Saturates, Aromatics, and Polars in Crude Oil by ^{13}C NMR and Support Vector Regression with Variable Selection by Genetic Algorithm. **Energy and Fuels**, vol. 30, no. 3, p. 1972–1978, 2016. <https://doi.org/10.1021/acs.energyfuels.5b02377>.

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**, vol. 116, p. 123–130, 2014. DOI 10.1016/j.fuel.2013.07.122. Available at: <http://dx.doi.org/10.1016/j.fuel.2013.07.122>.

FOLLI, G. S.; NASCIMENTO, M. H. C.; DE PAULO, E. H.; DA CUNHA, P. H. P.; ROMÃO, W.; FILGUEIRAS, P. R. Variable selection in support vector regression using angular search algorithm and variance inflation factor. **Journal of Chemometrics**, no. January, p. 1–16, 2020. <https://doi.org/10.1002/cem.3282>.

FORTUNATO, F. M.; VIEIRA, A. L.; GOMES NETO, J. A.; DONATI, G. L.; JONES, B. T. Expanding the potentialities of standard dilution analysis: Determination of ethanol in gasoline by Raman spectroscopy. **Microchemical Journal**, vol. 133, p. 76–80, Jul. 2017. DOI 10.1016/j.microc.2017.03.015. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0026265X16304799>.

FRIEDRICH, D. M.; HULSE, C. A.; VON GUNTEN, M.; WILLIAMSON, E. P.; PEDERSON, C. G.; O'BRIEN, N. A. Miniature near-infrared spectrometer for point-of-use chemical analysis. **Photonic Instrumentation Engineering**, vol. 8992, no. March

2014, p. 899203, 2014. <https://doi.org/10.1117/12.2040669>.

GALTIER, O.; ABBAS, O.; LE DRÉAU, Y.; REBUFA, C.; KISTER, J.; ARTAUD, J.; DUPUY, N. Comparison of PLS1-DA, PLS2-DA and SIMCA for classification by origin of crude petroleum oils by MIR and virgin olive oils by NIR for different spectral regions. **Vibrational Spectroscopy**, vol. 55, no. 1, p. 132–140, 2011. DOI 10.1016/j.vibspec.2010.09.012. Available at: <http://dx.doi.org/10.1016/j.vibspec.2010.09.012>.

GALUSZKA, A.; MIGASZEWSKI, Z. M.; NAMIEŚNIK, J. Moving your laboratories to the field – Advantages and limitations of the use of field portable instruments in environmental sample analysis. **Environmental Research**, vol. 140, p. 593–603, Jul. 2015. DOI 10.1016/j.envres.2015.05.017. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0013935115001632>.

GARCÍA-MENCÍA, M. .; ANDRADE, J. .; LÓPEZ-MAHÍA, P.; PRADA, D. An empirical approach to update multivariate regression models intended for routine industrial use. **Fuel**, vol. 79, no. 14, p. 1823–1832, Nov. 2000. DOI 10.1016/S0016-2361(00)00046-6. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0016236100000466>.

GELADI, P. Chemometrics in spectroscopy. Part 1. Classical chemometrics. **Spectrochimica Acta - Part B Atomic Spectroscopy**, vol. 58, no. 5, p. 767–782, 2003. [https://doi.org/10.1016/S0584-8547\(03\)00037-5](https://doi.org/10.1016/S0584-8547(03)00037-5).

GLOBONEWS. Tentativa de furto em duto da Petrobras em Magé provoca vazamento de 60 mil litros de óleo em rio da Baía de Guanabara. 2018. **G1**. Available at: <https://g1.globo.com/rj/rio-de-janeiro/noticia/2018/12/09/tentativa-de-furto-de-oleo-em-duto-em-mage-provoca-vazamento-em-rio-e-chega-a-manguezal.ghtml>.

Accessed on: 17 Feb. 2022.

GORDY, W.; STANFORD, S. C. Spectroscopic Evidence for Hydrogen Bonds: SH, NH and NH₂ Compounds. **Journal of the American Chemical Society**, vol. 62, no. 3, p. 497–505, 1 Mar. 1940. DOI 10.1021/ja01860a011. Available at: <https://pubs.acs.org/doi/abs/10.1021/ja01860a011>.

GUCHARDI, R.; PASQUINI, C. Evaluation of a dual-beam near-infrared spectrometer based on acousto-optic tunable filters. **Applied Spectroscopy**, vol. 55, no. 4, p. 454–457, 2001. <https://doi.org/10.1366/0003702011951966>.

GUZMÁN-OSORIO, F.; DOMÍNGUEZ-RODRÍGUEZ, V.; ADAMS, R.; LOBATO-GARCÍA, C.; GUERRERO-PEÑA, A.; BARAJAS-HERNÁNDEZ, J. Classification of petroleum origin and integrity by FTIR. **Egyptian Journal of Petroleum**, vol. 30, no. 2, p. 63–67, Jun. 2021. DOI 10.1016/j.ejpe.2021.03.003. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S111006212100012X>.

HEMMATEENEJAD, B.; DOROSTKAR, S. Identification of the Source of Geographical Origin of Iranian Crude Oil by Chemometrics Analysis of Fourier Transform Infrared Spectra. **Energy & Fuels**, vol. 28, no. 2, p. 761–765, 20 Feb. 2014. DOI 10.1021/ef4017202. Available at: <https://pubs.acs.org/doi/10.1021/ef4017202>.

HIDAJAT, K.; CHONG, S. M. Quality characterisation of crude oils by partial least square calibration of NIR spectral profiles. **Journal of Near Infrared Spectroscopy**, vol. 8, no. 1, p. 53–59, 2000. <https://doi.org/10.1255/jnirs.264>.

HOLLAND-MORITZ, K.; STACH, W.; HOLLAND-MORITZ, I. Possibilities of FTIR instrumentation for the investigation of fast processes: elongation and relaxation phenomena in polymers. **Journal of Molecular Structure**, vol. 60, p. 1–6, 1980. DOI 10.1016/0022-2860(80)80025-1. Available at: <https://linkinghub.elsevier.com/retrieve/pii/0022286080800251>.

ISO 12185. CRUDE PETROLEUM AND PETROLEUM PRODUCTS – DETERMINATION OF DENSITY – OSCILLATING U-TUBE METHOD. Geneva, Switzerland, 1996. .

KHANMOHAMMADI, M.; GARMARUDI, A. B.; DE LA GUARDIA, M. Characterization of petroleum-based products by infrared spectroscopy and chemometrics. **TrAC - Trends in Analytical Chemistry**, vol. 35, p. 135–149, 2012. <https://doi.org/10.1016/j.trac.2011.12.006>.

KIRCHLER, C. G.; PEZZEI, C. K.; BEĆ, K. B.; MAYR, S.; ISHIGAKI, M.; OZAKI, Y.; HUCK, C. W. Critical evaluation of spectral information of benchtop vs. portable near-infrared spectrometers: Quantum chemistry and two-dimensional correlation spectroscopy for a better understanding of PLS regression models of the rosmarinic acid content in Rosmarin. **Analyst**, vol. 142, no. 3, p. 455–464, 2017. <https://doi.org/10.1039/c6an02439d>.

LAMMOGLIA, T.; FILHO, C. R. de S. Spectroscopic characterization of oils yielded from Brazilian offshore basins: Potential applications of remote sensing. **Remote Sensing of Environment**, vol. 115, no. 10, p. 2525–2535, 2011. DOI 10.1016/j.rse.2011.04.038. Available at: <http://dx.doi.org/10.1016/j.rse.2011.04.038>.

LARKIN, P. J. **Infrared and raman spectroscopy: principles and spectral information**. 1st ed. New York: Elsevier, 2011.

LAXALDE, J.; CAILLOL, N.; WAHL, F.; RUCKEBUSCH, C.; DUPONCHEL, L. Combining near and mid infrared spectroscopy for heavy oil characterisation. **Fuel**, vol. 133, p. 310–316, 2014. <https://doi.org/10.1016/j.fuel.2014.05.041>.

LAXALDE, J.; RUCKEBUSCH, C.; DEVOS, O.; CAILLOL, N.; WAHL, F.; DUPONCHEL, L. Characterisation of heavy oils using near-infrared spectroscopy:

Optimisation of pre-processing methods and variable selection. **Analytica Chimica Acta**, vol. 705, no. 1–2, p. 227–234, 2011. <https://doi.org/10.1016/j.aca.2011.05.048>.

LI, H.; XU, Q.; LIANG, Y. libPLS: An Integrated Library for Partial Least Squares Regression and Discriminant Analysis. **PeerJ Preprints**, vol. 0, p. 410083, 2014. <https://doi.org/10.7287/peerj.preprints.190>.

LIANG, C.; YUAN, H. F.; ZHAO, Z.; SONG, C. F.; WANG, J. J. A new multivariate calibration model transfer method of near-infrared spectral analysis. **Chemometrics and Intelligent Laboratory Systems**, vol. 153, p. 51–57, 2016. DOI 10.1016/j.chemolab.2016.01.017. Available at: <http://dx.doi.org/10.1016/j.chemolab.2016.01.017>.

LONG, J.; WANG, K.; YANG, M.; ZHONG, W. Rapid crude oil analysis using near-infrared reflectance spectroscopy. **Petroleum Science and Technology**, vol. 37, no. 3, p. 354–360, 2019. DOI 10.1080/10916466.2018.1547754. Available at: <https://doi.org/10.1080/10916466.2018.1547754>.

LOVATTI, B. P. O.; NASCIMENTO, M. H. C.; RAINHA, K. P.; OLIVEIRA, E. C. S.; NETO, Á. C.; CASTRO, E. V. R.; FILGUEIRAS, P. R. Different strategies for the use of random forest in NMR spectra. **Journal of Chemometrics**, no. February, p. 1–10, 2 Mar. 2020. DOI 10.1002/cem.3231. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.3231>.

LOVATTI, B. P. O.; SILVA, S. R. C.; PORTELA, N. de A.; SAD, C. M. S.; RAINHA, K. P.; ROCHA, J. T. C.; ROMÃO, W.; CASTRO, E. V. R.; FILGUEIRAS, P. R. Identification of petroleum profiles by infrared spectroscopy and chemometrics. **Fuel**, vol. 254, no. June, 2019. <https://doi.org/10.1016/j.fuel.2019.115670>.

LUTZ, O. M. D.; BONN, G. K.; RODE, B. M.; HUCK, C. W. Reproducible quantification

of ethanol in gasoline via a customized mobile near-infrared spectrometer. **Analytica Chimica Acta**, vol. 826, no. 1, p. 61–68, 2014. <https://doi.org/10.1016/j.aca.2014.04.002>.

MALDONADO, M.; BARREIRO, P.; GUTIÉRREZ, R.; VERGARA, G. Mid- infrared uncooled sensor for the identification of pure fuel, additives and adulterants in gasoline. **Fuel Processing Technology**, vol. 171, no. September 2017, p. 287–292, 2018. DOI 10.1016/j.fuproc.2017.09.018. Available at: <https://doi.org/10.1016/j.fuproc.2017.09.018>.

MALEGORI, C.; NASCIMENTO MARQUES, E. J.; DE FREITAS, S. T.; PIMENTEL, M. F.; PASQUINI, C.; CASIRAGHI, E. Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the evaluation of acerola fruit quality, using PLS and SVM regression algorithms. **Talanta**, vol. 165, no. November 2016, p. 112–116, 2017. DOI 10.1016/j.talanta.2016.12.035. Available at: <http://dx.doi.org/10.1016/j.talanta.2016.12.035>.

MÁQUINA, A. D. V.; SITOIE, 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**, vol. 237, no. June 2018, p. 373–379, 2019. DOI 10.1016/j.fuel.2018.10.011. Available at: <https://doi.org/10.1016/j.fuel.2018.10.011>.

MARTENS, H.; NAES, T. **Multivariate calibration**. [S. l.]: John Wiley & Sons, INC, 1990. vol. 4, . DOI 10.1002/cem.1180040607. Available at: <http://doi.wiley.com/10.1002/cem.1180040607>.

MAZIVILA, S. J. Trends of non-destructive analytical methods for identification of biodiesel feedstock in diesel-biodiesel blend according to European Commission

Directive 2012 / 0288 / EC and detecting diesel-biodiesel blend adulteration : A brief review. **Talanta**, vol. 180, no. October 2017, p. 239–247, 2018. DOI 10.1016/j.talanta.2017.12.057. Available at: <https://doi.org/10.1016/j.talanta.2017.12.057>.

MILANEZ, K. D. T. M.; SILVA, A. C.; PAZ, J. E. M.; MEDEIROS, E. P.; PONTES, M. J. C. Standardization of NIR data to identify adulteration in ethanol fuel. **Microchemical Journal**, vol. 124, p. 121–126, 2016. DOI 10.1016/j.microc.2015.08.013. Available at: <http://dx.doi.org/10.1016/j.microc.2015.08.013>.

MOHAMMADI, M.; KHANMOHAMMADI KHORRAMI, M.; VATANI, A.; GHASEMZADEH, H.; VATANPARAST, H.; BAHRAMIAN, A.; FALLAH, A. Genetic algorithm based support vector machine regression for prediction of SARA analysis in crude oil samples using ATR-FTIR spectroscopy. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, vol. 245, p. 118945, 2021. DOI 10.1016/j.saa.2020.118945. Available at: <https://doi.org/10.1016/j.saa.2020.118945>.

MOHAMMADI, M.; KHANMOHAMMADI KHORRAMI, M.; VATANI, A.; GHASEMZADEH, H.; VATANPARAST, H.; BAHRAMIAN, A.; FALLAH, A. Rapid determination and classification of crude oils by ATR-FTIR spectroscopy and chemometric methods. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, vol. 232, p. 118157, May 2020. DOI 10.1016/j.saa.2020.118157. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1386142520301359>.

MOORE, D. S.; NOTZ, W. I.; FLIGNER, M. A. **The Basic Practice of Statistics**. 8th ed. New York: W. H. Freeman, 2017.

MORO, M. K.; DOS SANTOS, F. D.; FOLLI, G. S.; ROMÃO, W.; FILGUEIRAS, P. R.

A review of chemometrics models to predict crude oil properties from nuclear magnetic resonance and infrared spectroscopy. **Fuel**, vol. 303, no. June, 2021. <https://doi.org/10.1016/j.fuel.2021.121283>.

MOURA, H. O. M. A.; CÂMARA, A. B. F.; SANTOS, M. C. D.; MORAIS, C. L. M.; DE LIMA, L. A. S.; LIMA, K. M. G.; DE CARVALHO, L. S. Advances in chemometric control of commercial diesel adulteration by kerosene using IR spectroscopy. **Analytical and Bioanalytical Chemistry**, vol. 411, no. 11, p. 2301–2315, 2019. <https://doi.org/10.1007/s00216-019-01671-y>.

MULLINS, O. C. Asphaltenes in Crude Oil: Absorbers and/or Scatterers in the Near-Infrared Region? **Analytical Chemistry**, vol. 62, no. 5, p. 508–514, 1990. <https://doi.org/10.1021/ac00204a016>.

NG, W.; MALONE, B. P.; MINASNY, B. Rapid assessment of petroleum-contaminated soils with infrared spectroscopy. **Geoderma**, vol. 289, p. 150–160, 2017. DOI 10.1016/j.geoderma.2016.11.030. Available at: <http://dx.doi.org/10.1016/j.geoderma.2016.11.030>.

NIKI, H.; MAKER, P. D.; BREITENBACH, L. P.; SAVAGE, C. M. FTIR studies of the kinetics and mechanism for the reaction of Cl atom with formaldehyde. **Chemical Physics Letters**, vol. 57, no. 4, p. 596–599, Aug. 1978. DOI 10.1016/0009-2614(78)85328-7. Available at: <https://linkinghub.elsevier.com/retrieve/pii/0009261478853287>.

NØRGAARD, L.; SAUDLAND, A.; WAGNER, J.; NIELSEN, J. P.; MUNCK, L.; ENGELSEN, S. B. Interval Partial Least-Squares Regression (i PLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. **Applied Spectroscopy**, vol. 54, no. 3, p. 413–419, 31 Mar. 2000. DOI

10.1366/0003702001949500. Available at:
<http://journals.sagepub.com/doi/10.1366/0003702001949500>.

NORRIS, K. H.; BUTLER, W. L. Techniques for Obtaining Absorption Spectra on Intact Biological Samples. **IRE Transactions on Bio-Medical Electronics**, vol. 8, no. 3, p. 153–157, Jul. 1961. DOI 10.1109/TBMEL.1961.4322890. Available at:
<http://ieeexplore.ieee.org/document/4322890/>.

NORRIS, Karl H.; HART, J. R. 4. Direct spectrophotometric determination of moisture content of grain and seeds. **Journal of Near Infrared Spectroscopy**, vol. 4, no. 1–4, p. 23–30, 1996. .

O'BRIEN, N. A.; HULSE, C. A.; FRIEDRICH, D. M.; VAN MILLIGEN, F. J.; VON GUNTEN, M. K.; PFEIFER, F.; SIESLER, H. W. Miniature near-infrared (NIR) spectrometer engine for handheld applications. 1 May 2012. [S. l.: s. n.], 1 May 2012. p. 837404. DOI 10.1117/12.917983. Available at:
<http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.917983>.

O'GRADY, M. R.; CONROY, C. M.; TAYLOR, L. T.; KNIGHT, C. L.; BODNAR, R. J. FTIR Microspectrometry of Individual Petroleum Fluid Inclusions In Geological Samples. 1 Dec. 1989. [S. l.: s. n.], 1 Dec. 1989. p. 613. DOI 10.1117/12.969629. Available at:
<http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.969629>.

PAIVA, E. M.; ROHWEDDER, J. J. R.; PASQUINI, C.; PEREIRA, C. F. Method for building a portable near infrared photometer based on LEDs and interference filters chosen by a spectral variable selection algorithm. **Microchemical Journal**, vol. 146, no. November 2018, p. 842–849, 2019. DOI 10.1016/j.microc.2019.01.074. Available at: <https://doi.org/10.1016/j.microc.2019.01.074>.

PAIVA, E. M.; ROHWEDDER, J. J. R.; PASQUINI, C.; PIMENTEL, M. F.; PEREIRA, C. F. Quantification of biodiesel and adulteration with vegetable oils in diesel/biodiesel blends using portable near-infrared spectrometer. **Fuel**, vol. 160, p. 57–63, 2015. DOI 10.1016/j.fuel.2015.07.067. Available at: <http://dx.doi.org/10.1016/j.fuel.2015.07.067>.

PANTOJA, P. A.; LÓPEZ-GEJO, J.; NASCIMENTO, C. A. O. do; ROUX, G. A. C. Le. Application of Near-Infrared Spectroscopy to the Characterization of Petroleum. **Analytical Characterization Methods for Crude Oil and Related Products**, , p. 221–243, 2017. <https://doi.org/10.1002/9781119286325.ch8>.

PARISOTTO, G.; FERRÃO, M. F.; MÜLLER, A. L. H.; MÜLLER, E. I.; SANTOS, M. F. P.; GUIMARÃES, R. C. L.; DIAS, J. C. M.; FLORES, È. M. M. Total acid number determination in residues of crude oil distillation using ATR-FTIR and variable selection by chemometric methods. **Energy and Fuels**, vol. 24, no. 10, p. 5474–5478, 2010. <https://doi.org/10.1021/ef1002974>.

PASQUINI, C. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. **Analytica Chimica Acta**, vol. 1026, p. 8–36, 2018. <https://doi.org/10.1016/j.aca.2018.04.004>.

PASQUINI, C. Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, vol. 14, no. 2, p. 198–219, 2003. <https://doi.org/10.1590/S0103-50532003000200006>.

PASQUINI, C.; BUENO, A. F. Characterization of petroleum using near-infrared spectroscopy: Quantitative modeling for the true boiling point curve and specific gravity. **Fuel**, vol. 86, no. 12–13, p. 1927–1934, 2007. <https://doi.org/10.1016/j.fuel.2006.12.026>.

PONTES, M. J. C.; PEREIRA, C. F.; PIMENTEL, M. F.; VASCONCELOS, F. V. C.;

SILVA, A. G. B. Screening analysis to detect adulteration in diesel/biodiesel blends using near infrared spectrometry and multivariate classification. **Talanta**, vol. 85, no. 4, p. 2159–2165, 2011. <https://doi.org/10.1016/j.talanta.2011.07.064>.

R CORE TEAM. R: A language and environment for statistical computing. Viena, 2021. Available at: <https://www.r-project.org/>.

RAINHA, K. P.; ROCHA, J. T. C.; RODRIGUES, R. R. T.; LOVATTI, B. P. O.; DE CASTRO, E. V. R.; FILGUEIRAS, P. R. Determination of API Gravity and Total and Basic Nitrogen Content by Mid- and Near-Infrared Spectroscopy in Crude Oil with Multivariate Regression and Variable Selection Tools. **Analytical Letters**, vol. 52, no. 18, p. 2914–2930, 2019. DOI 10.1080/00032719.2019.1628246. Available at: <https://doi.org/10.1080/00032719.2019.1628246>.

RAMALHO, A. Furto de combustíveis alimenta ganhos de refinarias clandestinas. 18 Jun. 2019. **Valor Econômico**. Available at: <https://valor.globo.com/brasil/noticia/2019/06/18/furto-de-combustiveis-alimenta-ganhos-de-refinarias-clandestinas.ghtml>. Accessed on: 17 Feb. 2022.

REUTERS. Petrobrás cria grupo interno para enfrentar ladrões de combustível. 2019. **Estadão**. Available at: <https://economia.estadao.com.br/noticias/negocios,petrobras-cria-grupo-interno-para-enfrentar-os-ladros-de-combustivel,70003018551>. Accessed on: 17 Feb. 2022.

RIAZI, M. **Characterization and Properties of Petroleum Fractions**. 1 st. West Conshohocken, PA: ASTM International, 2007. <https://doi.org/https://doi.org/10.1520/MNL50-EB>.

RIAZI, M. R. **Characterization and Properties of Petroleum Fractions**. 1st ed. Philadelphia: American Society for Testing and Materials, 2005.

RICCI, L. Prejuízo com desvio de gasolina da Regap é estimado em R\$ 4 milhões. 2018. **Estado de Minas.** Available at: https://www.em.com.br/app/noticia/gerais/2018/08/31/interna_gerais,984863/prejuizo-com-desvio-de-gasolina-da-regap-e-estimado-em-r-4-milhoes.shtml. Accessed on: 17 Feb. 2022.

RICH, S. History of the American Petroleum Institute (API). 2019. Available at: <https://drillers.com/history-of-the-american-petroleum-institute-api/>. Accessed on: 9 Feb. 2021.

RILEY, B. J.; LENNARD, C.; FULLER, S.; SPIKMANS, V. An FTIR method for the analysis of crude and heavy fuel oil asphaltenes to assist in oil fingerprinting. **Forensic Science International**, vol. 266, p. 555–564, Sep. 2016a. DOI 10.1016/j.forsciint.2016.07.018. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S037907381630322X>.

RILEY, B. J.; LENNARD, C.; FULLER, S.; SPIKMANS, V. An FTIR method for the analysis of crude and heavy fuel oil asphaltenes to assist in oil fingerprinting. **Forensic Science International**, vol. 266, p. 555–564, Sep. 2016b. DOI 10.1016/j.forsciint.2016.07.018. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S037907381630322X>.

ROCHA, J. T. C.; OLIVEIRA, L. M. S. L.; DIAS, J. C. M.; PINTO, U. B.; MARQUES, M. D. L. S. P.; OLIVEIRA, B. P.; FILGUEIRAS, P. R.; CASTRO, E. V. R.; DE OLIVEIRA, M. A. L. Sulfur Determination in Brazilian Petroleum Fractions by Mid-infrared and Near-infrared Spectroscopy and Partial Least Squares Associated with Variable Selection Methods. **Energy and Fuels**, vol. 30, no. 1, p. 698–705, 2016. <https://doi.org/10.1021/acs.energyfuels.5b02463>.

ROSSINI, F. D. American Petroleum Institute Spectral Data and Standard Samples.

Applied Spectroscopy, vol. 6, no. 1, p. 3–13, 1951.

<https://doi.org/10.1366/000370251774652759>.

SANTOS, F. D.; SANTOS, L. P.; CUNHA, P. H. P.; BORGHI, F. T.; ROMÃO, W.; DE CASTRO, E. V. R.; DE OLIVEIRA, E. C.; FILGUEIRAS, P. R. Discrimination of oils and fuels using a portable NIR spectrometer. **Fuel**, vol. 283, no. January 2020, p. 1–10, 2021. <https://doi.org/10.1016/j.fuel.2020.118854>.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, vol. 36, no. 8, p. 1627–1639, 1964. <https://doi.org/10.1021/ac60214a047>.

SENICH, G. A.; MACKNIGHT, W. J. Fourier Transform Infrared Thermal Analysis of a Segmented Polyurethane. **Macromolecules**, vol. 13, no. 1, p. 106–110, 1 Jan. 1980. DOI 10.1021/ma60073a021. Available at: <https://pubs.acs.org/doi/abs/10.1021/ma60073a021>.

SILVA, A. C.; LIRA PONTES, L. F. B.; PIMENTEL, M. F.; PONTES, M. J. C. Detection of adulteration in hydrated ethyl alcohol fuel using infrared spectroscopy and supervised pattern recognition methods. **Talanta**, vol. 93, p. 129–134, 2012. DOI 10.1016/j.talanta.2012.01.060. Available at: <http://dx.doi.org/10.1016/j.talanta.2012.01.060>.

SILVA, C. S.; BRAZ, A.; PIMENTEL, M. F. Vibrational spectroscopy and chemometrics in forensic chemistry: Critical review, current trends and challenges. **Journal of the Brazilian Chemical Society**, vol. 30, no. 11, p. 2259–2290, 2019. <https://doi.org/10.21577/0103-5053.20190140>.

SILVERSTEIN, R. M.; WEBSTER, F. X.; KIEMLE, D. J.; BRYCE, D. L. **Spectrometric**

identification of organic compounds (2015, John Wiley & Sons, Inc.) -.pdf. 8th ed.

[S. l.]: John Wiley & Sons, 2014.

SKOOG, D. A.; HOLLER, J. F.; CROUCH, S. R. **Principles of Instrumental Analysis**.

7th ed. Boston: Cengage Learning, 2018.

SKOOG, D. A.; WEST, D. M.; HOLLER, F. J.; CROUCH, S. R. **Fundamentals of analytical chemistry**. 9th ed. Belmonte: Cengage Learning, 2014.

SOLOMON, P. R.; CARANGELO, R. M.; GRAVEL, D.; BAILLARGEON, M.; BAUDAIS, F.; VAIL, G. **Characterization of petroleum feedstocks by TG-FTIR**. [S. l.: s. n.], 1988.

SPEIGHT, J. G. **Handbook of Petroleum Product Analysis**. Hoboken: Wiley-Interscience, 2002. vol. 140, . DOI 10.1001/jama.1949.02900410073036. Available at: <http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.1949.02900410073036>.

SPEIGHT, J. G. **Handbook of petroleum product analysis (2nd edition)**. [S. l.: s. n.], 2015. vol. 53, . <https://doi.org/10.1017/CBO9781107415324.004>.

SPRINGER. **Near-Infrared Spectroscopy**. Singapore: Springer Singapore, 2021. DOI 10.1007/978-981-15-8648-4. Available at: <http://link.springer.com/10.1007/978-981-15-8648-4>.

SUN, L.; HSIUNG, C.; PEDERSON, C. G.; ZOU, P.; SMITH, V.; VON GUNTEN, M.; O'BRIEN, N. A. Pharmaceutical Raw Material Identification Using Miniature Near-Infrared (MicroNIR) Spectroscopy and Supervised Pattern Recognition Using Support Vector Machine. **Applied Spectroscopy**, vol. 70, no. 5, p. 816–825, 2016. <https://doi.org/10.1177/0003702816638281>.

TEIXEIRA DOS SANTOS, C. A.; LOPO, M.; PÁSCOA, R. N. M. J.; LOPES, J. A. A

review on the applications of portable near-infrared spectrometers in the agro-food industry. **Applied Spectroscopy**, vol. 67, no. 11, p. 1215–1233, 2013. <https://doi.org/10.1366/13-07228>.

THISSEN, U.; PEPERS, M.; ÜSTÜN, B.; MELSEN, W. J.; BUYDENS, L. M. C. Comparing support vector machines to PLS for spectral regression applications. **Chemometrics and Intelligent Laboratory Systems**, vol. 73, no. 2, p. 169–179, Oct. 2004. DOI 10.1016/j.chemolab.2004.01.002. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0169743904000103>.

THORNE, A. P. Fourier Transform Spectrometry in the Ultraviolet. **Analytical Chemistry**, vol. 63, no. 2, p. 57A-65A, 15 Jan. 1991. DOI 10.1021/ac00002a712. Available at: <https://pubs.acs.org/doi/abs/10.1021/ac00002a712>.

TOSATO, F.; BARROS, E.; CUNHA, D.; SANTOS, F.; CORRÊA, T.; NUNES, A.; JASTROW, I.; SILVEROL, M.; PINHEIRO, L.; SEABRA, A.; FILGUEIRAS, P.; NETO, Á.; WANG, L.; FERREIRA, E.; ROMÃO, W. Análise de amostras de combustíveis por fotometria, NIR portátil e RMN de 1 H – uma comparação com os resultados encontrados por técnicas normatizadas. **Química Nova**, vol. 43, no. 2, p. 155–167, 2020. DOI 10.21577/0100-4042.20170473. Available at: http://quimicanova.sbq.org.br/audiencia_pdf.asp?aid2=8029&nomeArquivo=AR20190271.pdf.

UOP LLC. UOP269-10 Nitrogen Bases in Hydrocarbons by Potentiometric Titration. 2010. .

VANINI, G.; PEREIRA, V. B.; ROMÃO, W.; GOMES, A. O.; OLIVEIRA, L. M. S. L.; DIAS, J. C. M.; AZEVEDO, D. A. Analytical advanced techniques in the molecular-level characterization of Brazilian crude oils. **Microchemical Journal**, vol. 137, p. 111–118,

2018. <https://doi.org/10.1016/j.microc.2017.10.003>.

VEMPATAPU, B. P.; KANAUIA, P. K. Monitoring petroleum fuel adulteration: A review of analytical methods. **TrAC - Trends in Analytical Chemistry**, vol. 92, p. 1–11, 2017. <https://doi.org/10.1016/j.trac.2017.04.011>.

WHEELER, O. H. Near Infrared Spectra Of Organic Compounds. **Chemical Reviews**, vol. 59, no. 4, p. 629–666, 1959. <https://doi.org/10.1021/cr50028a004>.

WILLIAMS, D. An Application of Infrared Spectroscopy to Rubber Chemistry. **Physics**, vol. 7, no. 11, p. 399–402, Nov. 1936. DOI 10.1063/1.1745348. Available at: <http://aip.scitation.org/doi/10.1063/1.1745348>.

WILT, B. K.; WELCH, W. T.; RANKIN, J. G. Determination of Asphaltenes in Petroleum Crude Oils by Fourier Transform Infrared Spectroscopy. **Energy & Fuels**, vol. 12, no. 5, p. 1008–1012, 1 Sep. 1998. DOI 10.1021/ef980078p. Available at: <https://pubs.acs.org/doi/10.1021/ef980078p>.

XIAOBO, Z.; JIEWEN, Z.; POVEY, M. J. W.; HOLMES, M.; HANPIN, M. Variables selection methods in near-infrared spectroscopy. **Analytica Chimica Acta**, vol. 667, no. 1–2, p. 14–32, 2010. DOI 10.1016/j.aca.2010.03.048. Available at: <http://dx.doi.org/10.1016/j.aca.2010.03.048>.

ZHANG, Z.-M.; CHEN, S.; LIANG, Y.-Z. Baseline correction using adaptive iteratively reweighted penalized least squares. **The Analyst**, vol. 135, no. 5, p. 1138, 2010. DOI 10.1039/b922045c. Available at: <http://xlink.rsc.org/?DOI=b922045c>.

ZONTOV, Y. V.; BALKLOVA, K. S.; TITOVA, A. V.; RODIONOVA, O. Y.; POMERANTSEV, A. L. Journal of Pharmaceutical and Biomedical Analysis Chemometric aided NIR portable instrument for rapid assessment of medicine quality.

Journal of Pharmaceutical and Biomedical Analysis, vol. 131, p. 87–93, 2016. DOI
10.1016/j.jpba.2016.08.008. Available at: <http://dx.doi.org/10.1016/j.jpba.2016.08.008>.

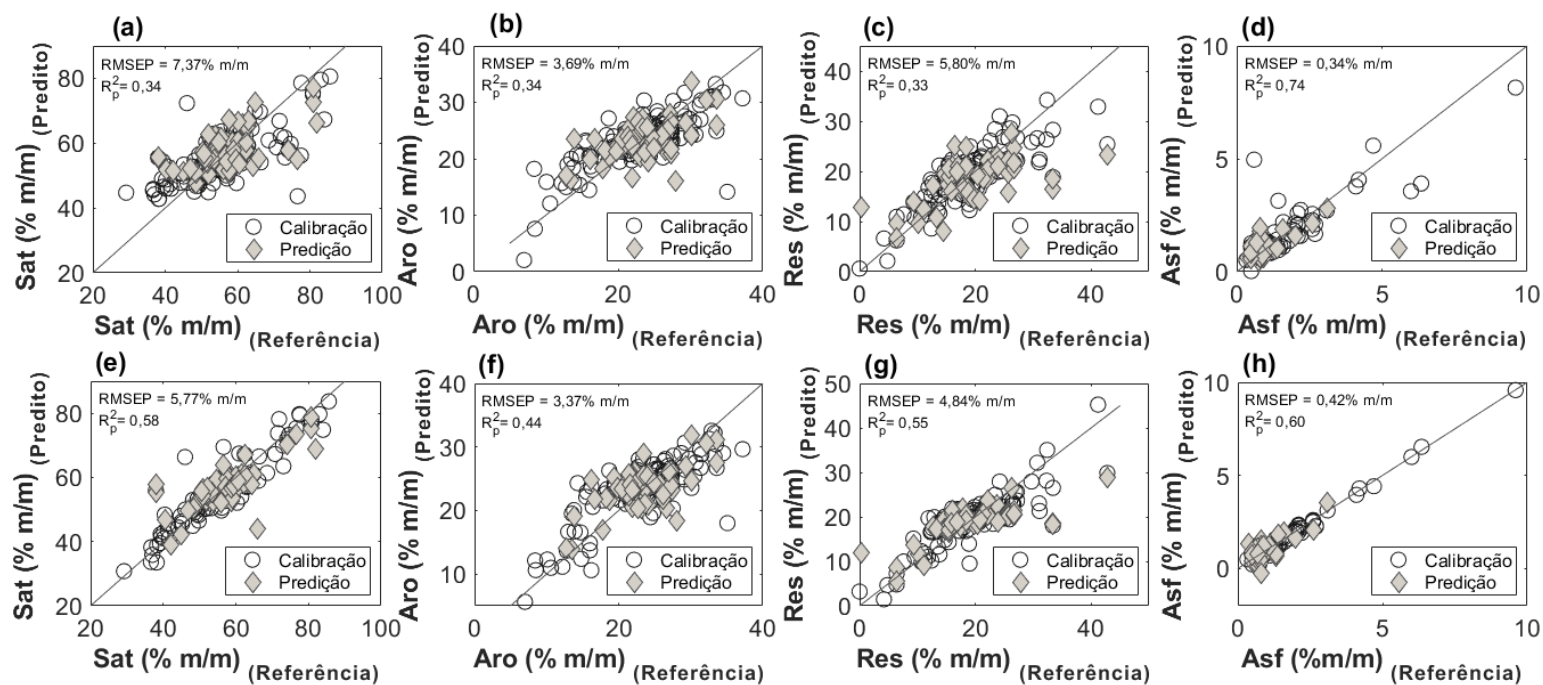


Figura 9.1.Curvas de calibração dos modelos PLS Sat, Aro, Res e Asf, respectivamente, para (a-d) Micronir I e (e-h) FT-NIR I.

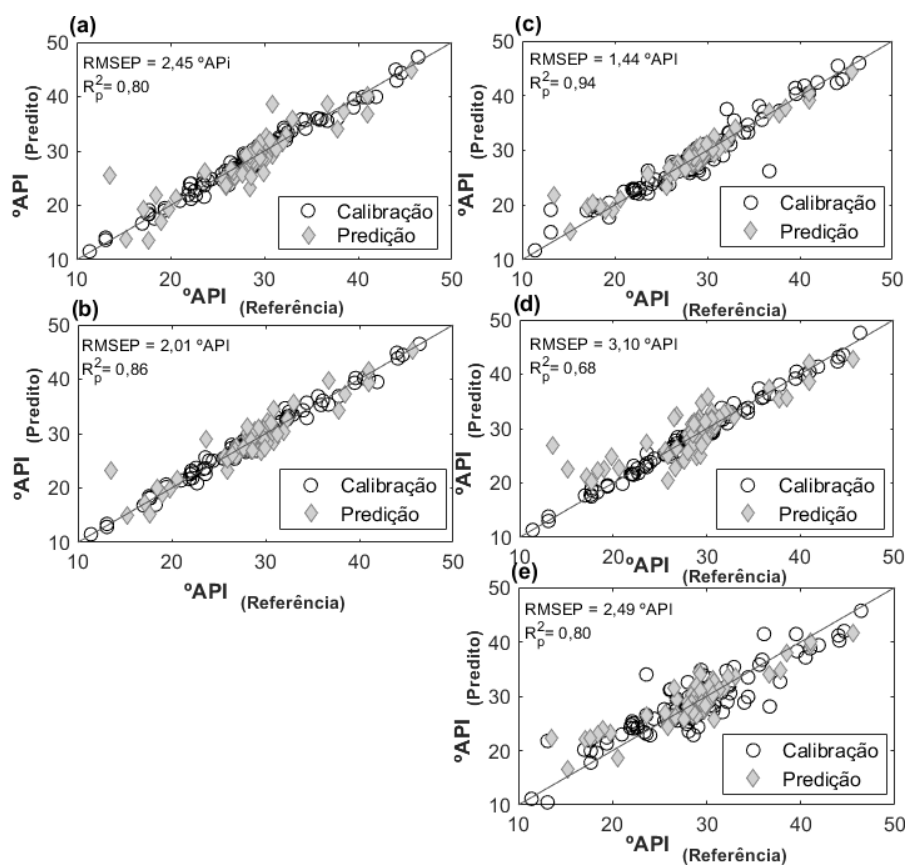


Figura 9.2. Curvas de calibração dos modelos PLS para estimar °API com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.

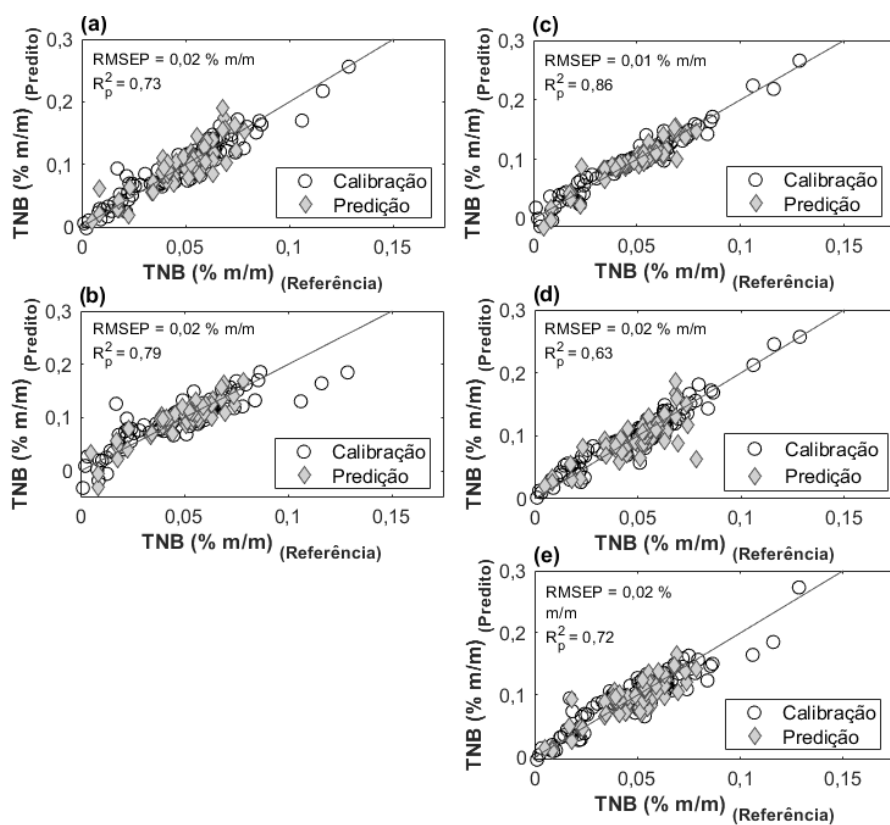


Figura 9.3. Curvas de calibração dos modelos PLS para estimar TNB (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.

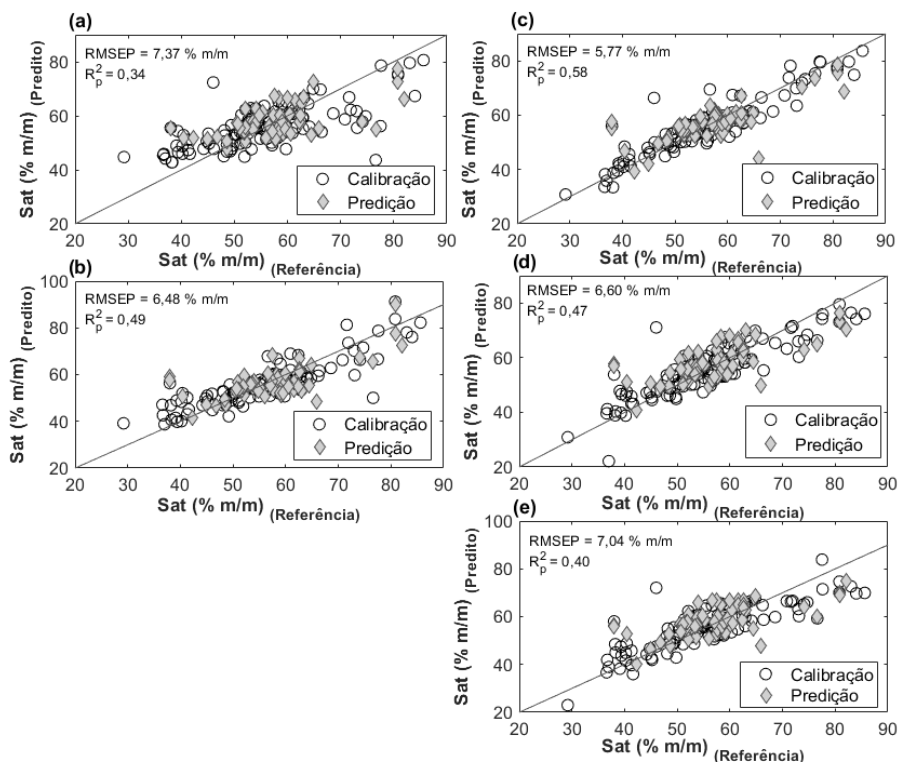


Figura 9.4. Curvas de calibração dos modelos PLS para estimar Sat (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.

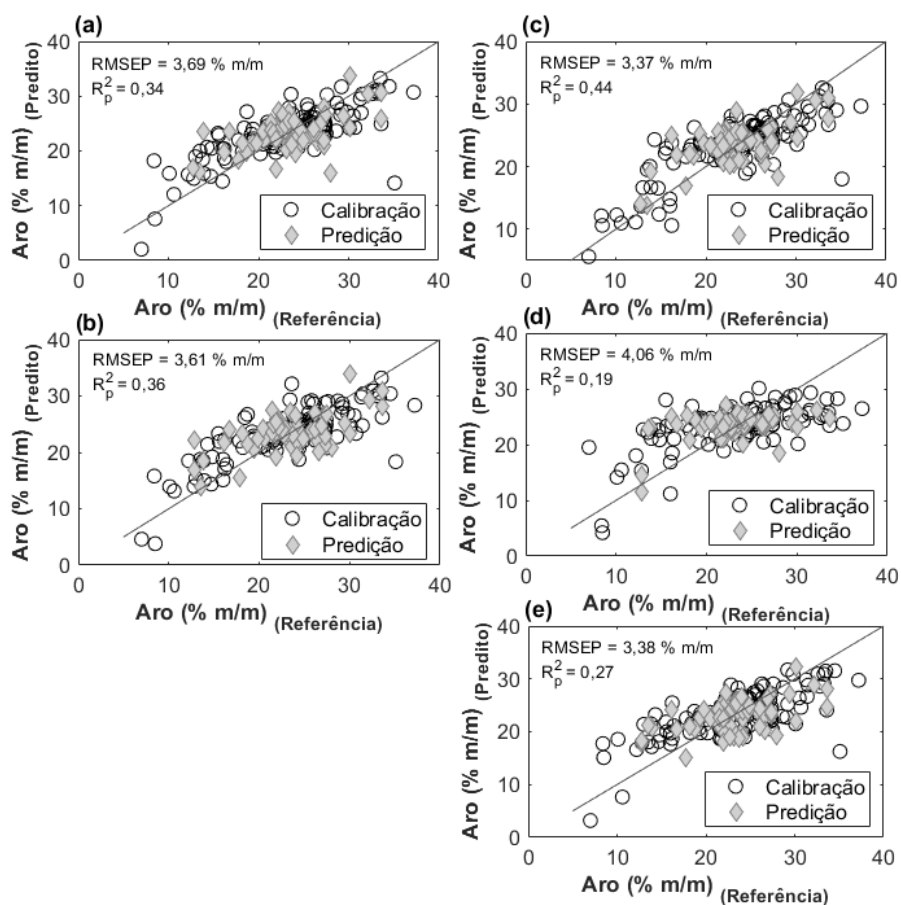


Figura 9.5. Curvas de calibração dos modelos PLS para estimar Aro (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.

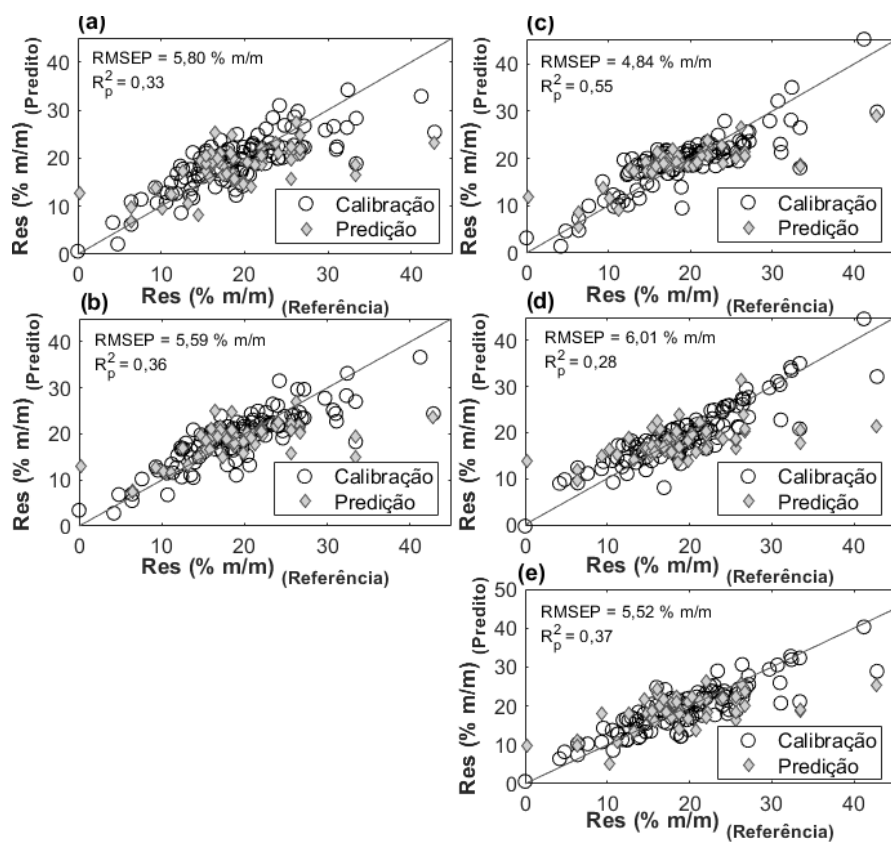


Figura 9.6.Curvas de calibração dos modelos PLS para estimar Res (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.

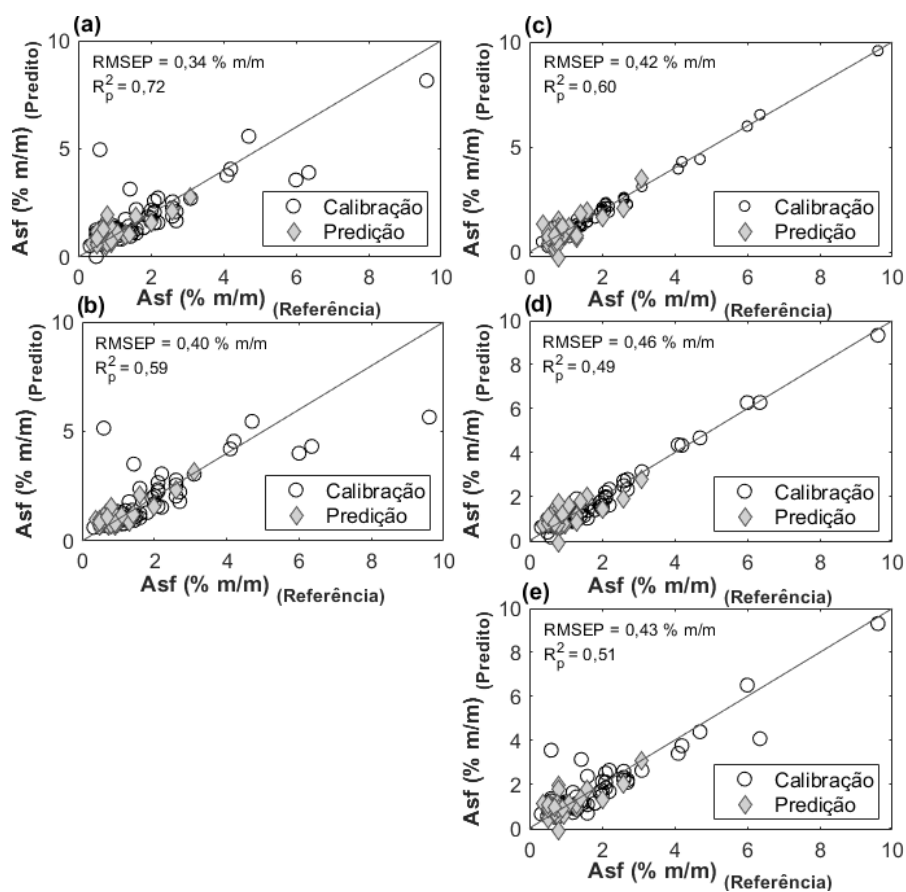


Figura 9.7. Curvas de calibração dos modelos PLS para estimar Asf (% m/m) com os dados (a) Micronir I, (b) Micronir II, (c) FT-NIR I, (d) FT-NIR II e (e) FT-NIR III.