

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

**Estudo de Alguns Parâmetros Físico-Químicos de
Petróleos Utilizando Modelagem Multivariada
Associada a Espectroscopia no Infravermelho**

Karla Pereira Rainha

Dissertação de Mestrado em Química

**Vitória
2014**

Karla Pereira Rainha

**Estudo de Alguns Parâmetros Físico-Químicos de
Petróleos Utilizando Modelagem Multivariada
Associada a Espectroscopia no Infravermelho**

Dissertação 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 Mestre em
Química, na área de Química
Analítica.

Orientador: Prof. Dr. Eustáquio
Vinícius Ribeiro de Castro

**VITÓRIA
2014**

Estudo de Alguns Parâmetros Físico-Químicos de Petróleos Utilizando Modelagem Multivariada Associada a Espectroscopia no Infravermelho

Karla Pereira Rainha

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

Aprovado(a) em 20/11/2014 por:

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

**Prof. Dr. Marcone Augusto Leal de Oliveira
Universidade Federal de Juiz de Fora**

**Prof. Dr. Reginaldo Bezerra dos Santos
Universidade Federal do Espírito Santo**

**Universidade Federal do Espírito Santo
Vitória, Novembro de 2014.**

A minha mãe, Rute Pereira Rainha,
por todo amor dedicado e esforço para me tornar a pessoa que sou.

AGRADECIMENTOS

A Deus, pela sabedoria e discernimento.

A minha amada mãe, Rute Pereira Rainha, por toda paciência que tive comigo, por todo amor e força, por toda motivação e amizade que me deu durante toda minha vida.

Ao Prof. Dr. Eustaquio Vinicius Ribeiro de Castro pela amizade, confiança e oportunidade de crescimento profissional e pessoal.

A banca, por aceitarem o convite.

A equipe do NCQP/LabPetro por toda amizade, em especial Suzy e Carlão.

Ao grupo da Quimiometria, em especial Paulo e Júlia.

Ao PPGQUI/UFES pela oportunidade, em especial Prof. Valdemar e Dirce.

A CAPES e FAPES pelo financiamento.

Ao CENPES/PETROBRAS.

A todos que contribuíram de alguma maneira em minha formação pessoal e profissional.

*“A matemática é o alfabeto no qual
Deus escreveu o universo.”*
Galileu Galilei (1565-1642)

LISTA DE FIGURAS

Figura 1: Matriz de dados originais.	23
Figura 2: Sistema de eixos da PCA.....	24
Figura 3: Espectroscópio infravermelho Perkin Elmer - Spectrum 400.	32
Figura 4: Célula de multi-ATR para leituras na região do MIR.	33
Figura 5: Célula de transreflectância com petróleo utilizada para leitura na região do NIR.	33
Figura 6: Espectros MIR do conjunto de calibração.	39
Figura 7: Espectros NIR do conjunto de calibração.	39
Figura 8: Seleção da região MIR para modelagem da propriedade NAT.	40
Figura 9: Seleção do número de variáveis latentes nas regiões MIR e NIR para NAT.....	41
Figura 10: Avaliação de leverage em MIR e NIR para NAT.	42
Figura 11: Valores obtidos para número de acidez total do método de referência versus a predição: (a) MIR, (b) NIR. Amostras de calibração (o), amostras de predição (*).	44
Figura 12: Resposta calculada para baixa acidez pelo modelo PLS-DA utilizando espectros MIR. Amostras (●) de treinamento e (o) amostras de teste com baixa acidez; (♦) treinamento e (◇) amostras teste com alta acidez.	46
Figura 13: Resposta calculada para baixa acidez pelo modelo PLS-DA utilizando espectros NIR. Amostras (●) de treinamento e (o) amostras de teste com baixa acidez; (♦) treinamento e (◇) amostras de teste com alta acidez.....	46
Figura 14: Seleção da região espectral em MIR para determinação de NAN. .	49
Figura 15: Valores obtidos para número de acidez naftênica do método de referência versus a predição: (a) MIR, (b) NIR. Amostras de calibração (o), amostras de predição (*).	50
Figura 16: RMSECV versus número de variáveis do Teor de Nitrogênio Total para MIR e NIR.....	52

Figura 17. Região espectral para o modelo MIR do Nitrogênio Total.	53
Figura 18: Análise de leverage do Nitrogênio Total.	54
Figura 19: Modelo de regressão MIR para Nitrogênio Total.	55
Figura 20: Regressão com barras de incerteza para Teor de Nitrogênio Total.	55
Figura 21: Região espectral para o modelo MIR do Nitrogênio Básico.	58
Figura 22: RMSECV versus número de variáveis do Nitrogênio Básico.	58
Figura 23: Análise de leverage do Nitrogênio Básico.	59
Figura 24: Modelo de regressão MIR para Nitrogênio Básico.	60
Figura 25: Intervalos de confiança do modelo para a predição do Teor de Nitrogênio Básico.	60

LISTA DE TABELAS

Tabela 1: Composição Elementar do Óleo Cru.	15
Tabela 2. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação do Número de Acidez Total.....	43
Tabela 3. Resultados dos testes de tendência nos resíduos e de permutação para Número de Acidez Total (NAT).	47
Tabela 4. Resultados do modelo de classificação PLS-DA....	43
Tabela 5. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação de Número de Acidez Naftênica (NAN).	48
Tabela 6. Resultados dos testes de tendência nos resíduos e de permutação para Número de Acidez Naftênica (NAN).....	49
Tabela 7. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação de Teor de Nitrogênio Total (TNT)..	56
Tabela 8. Resultados dos testes de tendência nos resíduos e de permutação de Teor de Nitrogênio Total (TNT)..	53
Tabela 9. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação de Teor de Nitrogênio Básico (TNB).....	56
Tabela 10. Resultados dos testes de tendência nos resíduos e de permutação de Teor de Nitrogênio Básico (TNB).....	57

LISTA DE ABREVIATURAS E SIGLAS

AN - Ácidos Naftênicos

API - American Petroleum Institute

ASTM - American Society Test and Method

DERIVMEAN - Derivação com suavização Savitz-Golay e dados centrados na média

D_{60/60°F} - Densidade Relativa a 60/60 °F

iPLS - Partial Least Squares for intervals (Mínimos Quadrados Parciais por intervalos)

IUPAC - International Union of Pure and Applied Chemistry

LabPetro - Laboratório de Pesquisa e Desenvolvimento de Metodologias para Análises de Petróleo

MIR - Middle Infrared Reflectance (Espectroscopia na região do infravermelho médio)

NAN - Número de Acidez Naftênica

NAT - Número de Acidez Total

NCQP - Núcleo de Competências em Química do Petróleo

NIR - Near Infrared Reflectance (Espectroscopia na região do infravermelho próximo)

PETROBRAS - Petróleo Brasileiro S/A.

PLS - Partial Least Squares (Mínimos Quadrados Parciais)

PLS-DA - Partial Least Squares Discriminant Analysis (Mínimos Quadrados Parciais com Discriminante Linear)

RMSEC - Root Mean Square Error of Calibration (Raiz quadrada do erro quadrático médio de calibração)

RMSECV - Root Mean Square Error of Cross Validation (Raiz quadrada do erro quadrático médio de validação cruzada)

RMSEP - Root Mean Square Error of Prediction (Raiz quadrada do erro quadrático médio de predição)

siPLS - Partial Least Squares for Intervals Synergism (Mínimos Quadrados Parciais por sinergismo de intervalos)

SNV - Standard Normal Variate (Variação Normal Padrão)

TNB - Teor de Nitrogênio Básico

TNT - Teor de Nitrogênio Total

RESUMO

Título: Estudo de Alguns Parâmetros Físico-Químicos de Petróleos Utilizando Modelagem Multivariada Associada a Espectroscopia no Infravermelho

Discente: Karla Pereira Rainha

Orientador: Prof. Dr. Eustáquio Vinícius Ribeiro de Castro

Métodos tradicionais de caracterização de petróleos são laboriosos, caros e demandam muito tempo para serem realizados, além de impactos ambientais que podem provocar. Estudos têm mostrado que a quimiometria associada a espectroscopia no infravermelho tem dado resultados bastante interessantes nos parâmetros de caracterização e validação de petróleos e seus derivados. Os métodos espectroscópicos podem ser uma alternativa eficaz na caracterização de petróleos como meio de monitoramento e determinações de parâmetros físico-químicos devido a sua potencialidade, praticidade e rapidez analítica. Na primeira parte da metodologia, foram realizadas medidas espectroscópicas de 99 amostras de petróleo utilizando NIR e MIR como também determinações de propriedades físico-químicas de petróleos, tais como número de acidez total (NAT), número de acidez naftênica (NAN), teor de nitrogênio total (TNT) e teor de nitrogênio básico (TNB). Na segunda etapa, foram extraídas informações relevantes dos métodos analíticos empregados por meio de técnicas quimiométricas, como PCA, PLS, iPLS e siPLS. O RMSEP do modelo obtido para cada propriedade foi 0,1964 mg KOH.g⁻¹ para NAT, 0,3547 mg KOH.g⁻¹ para NAN, 0,0381% m/m para TNT, 0,0073% m/m para TNB. Os modelos para os parâmetros de acidez podem ser melhorados, pois não passaram nos testes de tendência nos resíduos e de permutação. Os modelos para as propriedades relacionada aos compostos nitrogenados obtiveram resultados satisfatórios, pois obedeceram todos os critérios de validação do protocolo de modelagem seguido. A contribuição deste trabalho foi o desenvolvimento de modelos quimiométricos através da combinação entre as informações obtidas pela espectroscopia MIR e NIR com os resultados das análises de caracterização de petróleo, de forma que os modelos podem antecipar os resultados de inferência das características físico-químicas de um óleo, sem a necessidade da realização de uma caracterização completa.

Palavras-chave: Espectroscopia de Infravermelho. Petróleo.

ABSTRACT

Title: Study of Some Oil Physicochemical Parameters Using Multivariate Modeling Associated with Infrared Spectroscopy of Oil Properties Linked to Chemometrics Spectroscopic Techniques

Student: Karla Pereira Rainha

Teacher: PhD Eustáquio Vinícius Ribeiro de Castro

Traditional methods of characterizing oils are laborious, expensive, require a long time to be realized and may result in environmental impacts. Studies have shown that the chemometrics associated with infrared spectroscopy has given very promising results in the characterizing parameters and validation of petroleum and its derivatives. Spectroscopic methods may be an effective alternative in the characterization of oil as a means of monitoring and determining physical and chemical parameters due to its potential, practical and analytical speed. In the first part of the methodology, spectroscopic measurements were performed in 99 oil samples using NIR and MIR and physical-chemical properties of oil were determined, such as total acid number (NAT), number of naphthenic acid (NAN), total nitrogen content (TNT) and basic nitrogen (TNB). In the second stage, the relevant information were extracted from the analytical methods employed by chemometric techniques such as PCA, PLS, iPLS and siPLS. The RMSEP of the model obtained for each property was 0.1964 mg KOH.g⁻¹ for NAT, 0.3547 mg KOH.g⁻¹ for NAN, 0.0381% m/m for TNT, 0.0073% m/m for TNB. The models for the acidity parameters can be improved, because they fail the trend tests on waste and permutation. The models for the properties related to nitrogen compounds obtained satisfactory results, as they passed all the validation criteria of the followed modeling protocol. The contribution of this work was the development of chemometric models by combining the information obtained by MIR and NIR spectroscopy with the results of oil characterization analysis, so that the models can anticipate the results by inference of the physicochemical characteristics of an oil, without the need to perform a complete characterization.

Keywords: Infrared Spectroscopic. Petroleum.

SUMÁRIO

1. INTRODUÇÃO	14
1.1. PETRÓLEO	14
1.1.1. VISCOSIDADE	17
1.1.2. DENSIDADE.....	17
1.2. A IMPORTÂNCIA DA CARACTERIZAÇÃO DO PETRÓLEO.....	18
1.3. ACIDEZ NAFTÊNICA.....	19
1.4. TEOR DE NITROGÊNIO.....	19
1.5. TÉCNICAS ESPECTROSCÓPICAS	20
1.5.1. INFRAVERMELHO PRÓXIMO	21
1.5.2. INFRAVERMELHO MÉDIO	21
1.6. QUIMIOMETRIA.....	22
1.6.1. TÉCNICAS QUIMIOMÉTRICAS.....	22
1.6.2. MÍNIMOS QUADRADOS PARCIAIS – PLS.....	23
1.6.3. MÍNIMOS QUADRADOS PARCIAIS COM DISCRIMINANTE LINEAR – PLS-DA	25
2. OBJETIVOS.....	27
2.1. OBJETIVO GERAL	27
2.2. OBJETIVOS ESPECÍFICOS	27
3. METODOLOGIA	28
3.1. CARACTERIZAÇÃO FÍSICO-QUÍMICA.....	29
3.1.1. NÚMERO DE ACIDEZ TOTAL	29
3.1.2. NÚMERO DE ACIDEZ NAFTÊNICA.....	30
3.1.3. TEOR DE NITROGÊNIO TOTAL	30
3.1.4. TEOR DE NITROGÊNIO BÁSICO	31
3.2. INSTRUMENTAÇÃO.....	32
3.2.1. INFRAVERMELHO MÉDIO	32
3.2.2. INFRAVERMELHO PRÓXIMO	33
3.3. MODELOS QUIMIOMÉTRICOS	34
3.3.1. PROTOCOLO DE MODELAGEM	34
4. RESULTADOS E DISCUSSÃO	38
4.1. DETERMINAÇÃO DE NÚMERO DE ACIDEZ TOTAL (NAT)	38
4.1.1. MODELO DE CLASSIFICAÇÃO POR PLS-DA.....	45
4.2. DETERMINAÇÃO DE NÚMERO DE ACIDEZ NAFTÊNICA (NAN).....	47

4.3. DETERMINAÇÃO DO TEOR DE NITROGÊNIO TOTAL (TNT)	51
4.4. DETERMINAÇÃO DO TEOR DE NITROGÊNIO BÁSICO (TNB).....	55
5. CONCLUSÕES	61
6. REFERÊNCIAS BIBLIOGRÁFICAS	63

1. INTRODUÇÃO

Métodos tradicionais de caracterização e avaliação em petróleos e seus derivados, tais como nafta leve e pesada, querosene, gasolina, gasóleo leve e pesado, resíduo atmosférico e resíduo de vácuo, normalmente são laboriosos, caros e demandam muito tempo para serem realizados, além de impactos ambientais que podem provocar. Em uma avaliação completa de petróleo é possível ultrapassar a marca de 500 ensaios, podendo demorar até mais de um ano para finalização (DIAS, 2005).

Às vezes, tem-se urgência no resultado de determinadas análises para uma tomada de decisão e isto fica prejudicado pela forma como as análises são feitas. Diante disto, surge a necessidade de se desenvolver técnicas mais rápidas, econômicas, confiáveis e que provoquem menos impacto ambiental.

Estudos tem mostrado que a quimiometria associada a alguns métodos espectroscópicos, como o infravermelho, tem dado resultados bastante interessantes nos parâmetros de caracterização de petróleos e seus derivados.

Técnicas espectroscópicas são metodologias que podem fornecer de forma rápida um grande número de informações sobre o conjunto de amostragem. Apresentam uma série de atrativos como, por exemplo, a possibilidade de análise sem preparação da amostra; a rapidez para a obtenção dos espectros, cálculos e apresentação dos resultados; o fato de ser um método não destrutivo, a possibilidade de guardar a amostra como testemunha (PASQUINI, 2003).

Neste trabalho foram utilizados dados espectrais de infravermelho médio (MIR) e próximo (NIR) no desenvolvimento de modelos quimiométricos para previsão de algumas propriedades físico-químicas utilizadas na caracterização e avaliação de petróleos.

1.1. PETRÓLEO

A palavra petróleo vem do latim *petra* (pedra) e *oleum* (óleo), o seu uso é conhecido desde as civilizações antigas. No Brasil, a história do petróleo começa

em 1858 com a assinatura do Decreto nº 2266, que concedia o direito a extração de material betuminoso com a finalidade de produzir querosene. Em 1953 Getúlio Vargas promulgou a criação da Petrobras, iniciando efetivamente as pesquisas de petróleo no Brasil. Atualmente, a produção diária de petróleo pela Petrobras ultrapassa 2,5 milhões de barris de óleo equivalente por dia. (THOMAS et al., 2001; site: PETROBRAS).

O petróleo é uma substância oleosa, inflamável, alguns são pretos com variação até a coloração castanha, opacos, com baixa viscosidade e densidade, e quantidade considerável de gás em solução (SPEIGHT, 2002).

Sua diversidade é devido à origem, idade do reservatório e profundidade da extração do poço (PARISOTTO, 2007 apud ANDRADE, 2009). Como a matriz é muito complexa, não existem dois petróleos idênticos. Estas diferenças na composição determinam os produtos que serão obtidos através do refino e como este processo será executado. (FILGUEIRAS, 2011)

O petróleo é composto por uma mistura complexa de hidrocarbonetos e compostos orgânicos contendo outros átomos como oxigênio, enxofre, nitrogênio e metais como ferro, zinco, cobre, chumbo, cobalto e em maior incidência níquel e vanádio. O estado físico é predominantemente líquido, levando em consideração as condições normais de temperatura e pressão (FILGUEIRAS, 2011; ANDRADE, 2009). A Tabela 1 contém a típica distribuição elementar do óleo cru.

Tabela 1: Composição elementar do óleo cru típico.

Elemento	Porcentagem em Massa
Hidrogênio	11 - 14
Carbono	83 - 87
Enxofre	0,06 - 8
Nitrogênio	0,11 - 1,7
Oxigênio	0,1 – 2
Metais	até 0,3%

Fonte: Thomas, Fundamentos de Engenharia de Petróleo, 2001.

Os principais elementos presentes no petróleo são carbono e hidrogênio,

no entanto o enxofre, nitrogênio e oxigênio, embora ocorram com menor frequência que os dois primeiros, são muito relevantes e aparecem em forma de compostos orgânicos (THOMAS, 2001).

Os hidrocarbonetos presentes no petróleo são classificados em saturados, insaturados e aromáticos. Os alcanos ou parafínicos são os hidrocarbonetos saturados, que são aqueles cujos átomos de carbono são unidos por ligações simples ao maior número de átomos de hidrogênio, formando cadeias lineares (parafínicos normais), ramificadas (iso-parafínicos) ou cíclicas (naftênicos). As olefinas é a denominação do grupo dos hidrocarbonetos insaturados, que apresentam pelo menos um ligação carbono-carbono dupla ou tripla. Os chamados arenos são os hidrocarbonetos aromáticos, que apresentam pelo menos um anel de benzeno em sua estrutura (THOMAS, 2001).

O oxigênio é o principal heteroátomo responsável pela acidez, coloração, odor, formação de gomas e corrosividade do petróleo e frações. Ele aparece em formas complexas como ácidos carboxílicos, fenóis, cresóis, ésteres, amidas, cetonas e benzofuranos. Concentra-se principalmente nas frações pesadas (THOMAS, 2001).

Os compostos nitrogenados aparecem majoritariamente como piridinas, quinolinas, pirróis, indóis, porfirinas e compostos policíclicos. Assim como os compostos oxigenados, também se concentram nas frações pesadas. Alto teor de nitrogênio no petróleo promove aumento da retenção de água nas emulsões e favorecendo a contaminação dos catalisados. Seus produtos finais são instáveis ocasionando formação de gomas e alteração de cor (THOMAS, 2001).

O enxofre é o terceiro elemento mais abundante na composição elementar do petróleo, ocorrendo principalmente em forma de sulfetos, polissulfetos, benzotiofenos e derivados. Compostos sulfurados são extremamente indesejados no petróleo por diversos motivos. Um exemplo é a produção de produtos que em combustão liberam SO_2 e SO_3 , que são gases poluentes (THOMAS, 2001).

A classificação dos petróleos é de grande interesse desde os geoquímicos até os refinadores. Para realizar esta classificação, as propriedades físico-químicas são determinadas de acordo com os seus constituintes e a partir destes

parâmetros é possível identificar o potencial de produção de um poço, estimar qual o tipo de composto predominante no óleo, propor ajustes nas condições do processo de produção e refino, avaliar a qualidade do petróleo e de seus derivados. O preço do petróleo é dimensionado principalmente a partir das medidas de viscosidade e densidade (FILGUEIRAS, 2011; ANDRADE, 2009).

1.1.1. VISCOSIDADE

A viscosidade é dada pela resistência de um fluido ao escoamento a uma certa temperatura. Consistindo na medida das forças internas de atrito (cisalhamento) do fluido em movimento. A energia para movimentar um petróleo mais viscoso será maior em relação a um petróleo mais fluido. (MACHADO, 2002)

Dentre as viscosidades que podem ser determinadas, a viscosidade cinemática é a mais importante para a indústria petrolífera, onde esta é levada em consideração para o dimensionamento dos equipamentos e também está ligada à qualidade do óleo. (THOMAS, 2001)

A viscosidade cinemática do petróleo pode ser determinada pelo produto do tempo de escoamento medido pela constante do capilar do viscosímetro. E através desta viscosidade é possível verificar se o índice de viscosidade, que consiste em medir a variação da viscosidade cinemática do óleo em relação a temperatura. Quanto maior este índice, menor será a variação de viscosidade com a variação de temperatura. (MACHADO, 2002; THOMAS, 2001)

1.1.2. DENSIDADE

A densidade é a propriedade físico-química mais utilizada na caracterização de petróleos. É definida como a relação entre a massa específica da substância a uma determinada temperatura e a massa específica de um padrão em outra temperatura, geralmente utiliza-se a água como padrão. (DIAS, 2005)

Normalmente a densidade é expressa em duas condições padronizadas: a densidade relativa ($d_{20/4}$) e o °API a 60/60°F.

A densidade relativa ($d_{20/4}$) é a razão entre a massa específica do petróleo

a 20 °C e a massa específica da água a 4 °C. (SPEIGHT, 2002)

A densidade expressa através do °API é mais utilizada na indústria do petróleo e é calculado a partir da densidade relativa a 60/60 °F. Esta medida foi criada pelo *American Petroleum Institute* (API) e calculado a partir da seguinte equação 1.

$$^{\circ}\text{API}_{60/60\text{ }^{\circ}\text{F}} = \frac{141,5}{D_{60/60^{\circ}\text{F}}} - 131,5 \quad (1)$$

Este parâmetro é baseado na densidade e na temperatura, pois quanto maior a temperatura menor a densidade e por tanto, maior o °API.

É um dos critérios para estabelecer o preço do petróleo bruto e estimar o rendimento das frações que podem ser obtidas deste óleo em seu refino.

1.2. A IMPORTÂNCIA DA CARACTERIZAÇÃO DO PETRÓLEO

Tradicionalmente, atribui-se como caracterização o conjunto de determinações experimentais realizadas para identificar as características específicas do petróleo. De modo geral, a forma mais efetiva de se conhecer estas propriedades é a partir da execução de experimentos laboratoriais (DIAS, 2005). A caracterização dos parâmetros físico-químicas do petróleo é realizada por métodos analíticos convencionais, consumindo muito tempo e conhecimento em diferentes técnicas (ANDRADE, 2009).

A caracterização do petróleo pode ser motivada pelas mais variadas razões, assim são diversos os tipos ou modelos de experimentos que podem ser realizados para um óleo (DIAS, 2005). Para fins tributários, por exemplo, a legislação brasileira determina uma valoração cujo cálculo é baseado em apenas três propriedades: densidade, curva de destilação e teor de enxofre do petróleo (ANP, 20 00 apud DIAS, 2005). No entanto, um amplo conhecimento das principais propriedades do óleo viabiliza a tomada de decisão de forma rápida durante o processamento que visa o aumento da produção. Desta forma, as análises para uma caracterização completa do perfil de um petróleo pode ultrapassar a marca de 500 experimentos (DIAS, 2005).

1.3. ACIDEZ NAFTÊNICA

Essencialmente, os ácidos naftênicos são ácidos carboxílicos que contêm um ou mais anéis fundidos que apresentam saturação. No entanto, este termo também está sendo designado para descrever outros ácidos orgânicos, inclusive os que possuem anéis aromáticos, cadeias lineares e ramificadas. Geralmente estes ácidos são monocarboxilados, mas também é possível encontrar di e tricarboxilados (COLATI, 2012; GRUBER et al., 2012).

Geralmente, o teor de acidez é determinado pelo método titulométrico chamado Número de Acidez Total (NAT) e é expressa em miligramas de hidróxido de potássio suficientes para neutralizar a acidez de um grama de petróleo (PARISOTTO, 2007). Este método apenas indica o grau de corrosividade da amostra e não agrega informação sobre a qualidade e natureza dos ácidos, pois é uma metodologia muito genérica. (GRUBER et al., 2012).

O Número de Acidez Naftênica (NAN) é um método muito mais específico, pois primeiro ocorre a extração dos ácidos naftênicos e depois a análise. A etapa de extração é dificultada pelo fato de o petróleo ser uma matriz muito complexa. (GRUBER et al., 2012).

Estes ácidos são uma problemática para o refino e craqueamento, pois promovem a corrosão em linhas de transferência, colunas e condensadores da destilação (GRUBER et al., 2012). Estudos mostram que este fenômeno ocorre principalmente, por óleos com altos valores de acidez naftênica à temperatura elevada, ou seja, valores superiores a 0,5 mg KOH/g para NAT e temperaturas na faixa de 220 a 440 °C (GRUBER et al., 2012; PARISOTTO, 2007).

1.4. TEOR DE NITROGÊNIO

No Brasil é dado ênfase na aferição dos teores de nitrogênio, pois em altos teores podem provocar envenenamento dos catalisadores e os compostos nitrogenados podem precipitar podendo ocorrer a formação de depósitos em seus derivados (THOMAS, 2001).

A investigação e identificação dos compostos nitrogenados no petróleo é muito importante, para a redução dos efeitos da toxicidade que ele apresenta as

refinarias, ao meio ambiente e a saúde da população (Li, Liang, Hu et al., 2003).

O Teor de Nitrogênio Total (TNT), ASTM D 4629, é um indicativo das quantidades que podem ser encontradas do nitrogênio básico na amostra. O nitrogênio básico pode reagir com compostos ácidos e desta forma reduzir a atividade do catalisador (Li, Liang, Hu et al., 2003). O método UOP 269 para Teor de Nitrogênio Básico é realizado por uma titulação potenciométrica.

1.5. TÉCNICAS ESPECTROSCÓPICAS

Os métodos espectroscópicos podem ser uma alternativa eficaz na caracterização de petróleos como meio de monitoramento e determinações de parâmetros físico-químicos devido a sua potencialidade, praticidade e rapidez analítica. As técnicas espectroscópicas são capazes de fornecer informações sobre o óleo bruto a nível molecular, mas dificilmente o sinal analítico instrumental fornecerá a informação quantitativa de interesse diretamente. Para converter os sinais analíticos do instrumento à informação de interesse quantitativo, faz-se necessário a utilização da quimiometria (FILGUEIRAS, 2011; PEINDER, 2009).

Há relatos e patentes indicando bons resultados para a quantificação de propriedades em derivados de petróleo utilizando técnicas como UV-Vis, Infravermelho Próximo (NIR) e Médio (MIR), Raman, Ressonância Magnética Nuclear (RMN), Espectrometria de Massas (MS) e Técnicas Cromatográficas (PEINDER, 2009). No entanto, em petróleo bruto que é uma matriz mais complexa, de difícil análise e por apresentar problemáticas em sua modelagem os seus trabalhos são escassos. (PEINDER, 2009)

Na literatura são encontrados alguns trabalhos sobre a quantificação de propriedades em petróleo utilizando técnicas quimiométricas associadas com infravermelho próximo e médio. No entanto, alguns métodos sugeridos têm avaliado apenas a validação cruzada e os erros de predição, o que é um problema para a utilização prática dos modelos.

1.5.1. INFRAVERMELHO PRÓXIMO

Na espectroscopia de infravermelho próximo - NIR (do inglês, Near Infrared), o intervalo de comprimento de onda está compreendido entre 12800-4000 cm^{-1} . Nesta faixa espectral é possível observar transições harmônicas (sobretons) e combinações das transições fundamentais. As principais informações espectrais são correspondentes às vibrações das ligações C-H, N-H, S-H e O-H, ou seja, ligações covalentes que apresentam átomo de hidrogênio. (ANDRADE, 2009; BUENO, 2004)

As bandas de absorvidade normalmente são amplas e pouco seletivas, assim é difícil atribuir os sinais analíticos a compostos específicos. Esta dificuldade pode ser contornada quando se utiliza o NIR vinculado as técnicas quimiométricas. O vidro pode ser usado como material de célula, tornando a análise mais simples e fácil. (ANDRADE, 2009; PEINDER, 2009)

Comparativamente, a quantidade de trabalhos publicados utilizando a espectroscopia NIR aplicado a derivados de petróleo é muito mais extensa do que a aplicação direta no petróleo bruto. (BUENO, 2004)

1.5.2. INFRAVERMELHO MÉDIO

Espectroscopia de infravermelho na região do médio - MIR (do inglês, Mid Infrared) abrange a região de comprimentos de onda 4000-200 cm^{-1} . É uma técnica de caracterização molecular ampla e bem aceita. Teoricamente, estes espectros são ricos em informação. As vibrações moleculares são definidas pela massa dos átomos e pela força de ligação entre eles, que aumenta com a resistência de união e diminui com a massa atômica. E ainda, as vibrações são sensíveis às mudanças de interação intramoleculares e intermoleculares. (PAVIA, LAMPMAN, KRIZ et al., 2010)

A intensidade de uma banda de absorção no infravermelho depende da mudança no momento dipolar durante a vibração: uma mudança grande no momento dipolar dará um aumento forte na absorção e uma pequena alteração para uma banda fraca. No entanto, espectroscopia no infravermelho médio é útil na determinação de grupos funcionais polares, tais como ligações C=O, N-H e O-H. Sendo que grupos menos polares como olefinas alifáticas, ligações C-H em

aromáticos e vibrações C-C também exibem bandas características, tornando esta técnica muito valiosa para a análise do óleo bruto. (PAVIA, LAMPMAN, KRIZ et al., 2010; PEINDER, 2009)

1.6. QUIMIOMETRIA

Segundo Kolwalski (apud Ferreira, 1999, p. 724), a quimiometria é “[...] uma disciplina da química que emprega métodos matemáticos e estatísticos para planejar ou selecionar experimentos de forma otimizada e para fornecer o máximo de informação química com a análise dos dados obtidos”. (FERREIRA, 1999)

A quimiometria é a área da química destinada à análise de dados químicos de natureza multivariada. Os métodos multivariados são os mais adequados, quando se trata de análise de dados, pois permite um estudo com várias variáveis ao mesmo tempo, não importando a existência ou ausência de diferenças entre elas nem a existência de correlação nos dados. (FERREIRA, 1999)

1.6.1. TÉCNICAS QUIMIOMÉTRICAS

As técnicas de análise multivariadas podem ter as mais diversas aplicações. Os métodos quimiométricos podem ser divididos em três tipos principais: planejamento de experimentos, método de análise exploratória e de calibração multivariada. A escolha de um ou de outro depende essencialmente da informação que se deseja obter. (BRERETON, 2003)

A análise multivariada requer a organização do conjunto de dados em estudo numa matriz X ($n \times m$), ilustrada pela Figura 1, onde as linhas desta matriz representam o conjunto das amostras e as colunas, o das variáveis medidas. Uma matriz de dados X , contendo m variáveis obtidas para n amostras ou objetos, pode ser graficamente representada por n pontos num espaço m -dimensional, ou melhor, é possível representar esta matriz espacialmente, onde cada variável é uma dimensão no espaço e cada amostra é um ponto neste espaço. (PARREIRA, 2003)

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & & x_{2p} \\ \vdots & & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

Figura 1: Matriz de dados originais. Fonte: Parreira, 2003

Para facilitar o tratamento de dados de um conjunto com muitas variáveis, são utilizados métodos que possibilitem uma melhor visualização destas dimensões. (PARREIRA, 2003)

1.6.2. MÍNIMOS QUADRADOS PARCIAIS – PLS

A principal característica da calibração multivariada é a possibilidade em usar diversas variáveis simultaneamente ($x_1, x_2, \dots, x_n$) para quantificar outra variável de interesse (Y) (PASSAMAI, 2004).

O PLS (*Partial Least Squares*), Mínimos Quadrados Parciais, é a técnica mais utilizada para calibração multivariada. Pode ser definida como uma manipulação matemática com objetivo de reescrever o conjunto de dados sem perda substancial de informação. Esta pode ser dividida em duas etapas: modelagem e validação. Na primeira etapa as matrizes X e Y do conjunto de calibração são decompostas simultaneamente em uma soma das variáveis latentes. As variáveis latentes são equivalentes as componentes principais da PCA, pelo fato da construção delas serem feitas a partir de informações contidas no vetor das variáveis dependentes. A calibração, segunda etapa, é a otimização da relação entre as matrizes X e Y com a finalidade de melhor descrever as amostras de interesse. O número de variáveis latentes a ser usado é determinado pela validação cruzada (ALEME, 2008; PARREIRA, 2003; PASSAMAI, 2004).

É utilizado o termo parcial, pois a solução de mínimos quadrados não se aplica a qualquer conjunto de *loadings*. Cada variável latente da matriz X é modificada de modo que a covariância entre ela e o vetor da matriz Y satisfaça

o critério de maximização do produto (ALEME, 2008).

As equações 2 e 3 demonstram as decomposições das matrizes X e Y em uma soma de h variáveis latentes (Figura 2). De forma que, T e U são as matrizes dos escores de X e Y respectivamente, P e Q são as matrizes dos pesos e os resíduos E e F. As matrizes X (dados espectrais) e Y (parâmetro físico-químico de interesse) estabelecem uma correlação linear a partir da equação 4, onde B é a matriz de coeficientes da regressão (Brereton, 2003; Filgueiras, 2011).

$$X_{(m \times n)} = T_{(m \times h)} \cdot P^t_{(h \times n)} + E_{(m \times n)} \quad (2)$$

$$Y_{(m \times k)} = U_{(m \times h)} \cdot Q^t_{(h \times k)} + F_{(m \times k)} \quad (3)$$

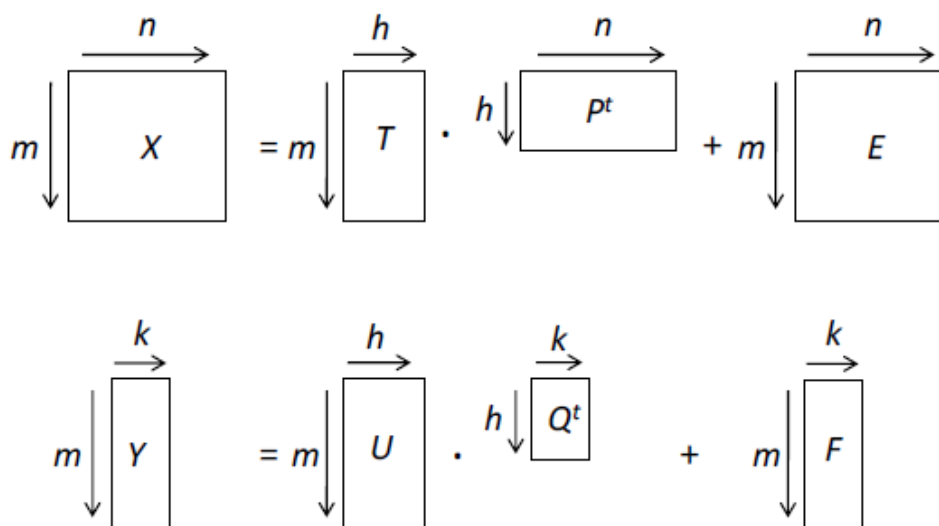


Figura 2: Decomposição de matrizes por PLS. Fonte: Brereton, 2003.

$$U_{(m \times h)} = T_{(m \times h)} \cdot B_{(h \times h)} \quad (4)$$

O PLS é sempre uma ferramenta importante, no entanto é necessário algum conhecimento prévio do sistema de dados a ser trabalhado. A regressão multivariada por PLS pode ser muito robusta, desde que as amostras externas tenham características semelhantes as amostras constituintes do modelo, pois as previsões a serem realizadas são totalmente baseadas em estatística. Quando os dados espectrais apresentam características muito diferentes não é possível assegurar a qualidade dos resultados obtidos (BRERETON, 2003).

Quando é aplicado PLS a técnicas espectroscópicas é comum usar

seleção de variáveis como PLS por intervalos (iPLS), que seleciona intervalos espectrais e constrói um modelo PLS cada um dos intervalos do espectro. E também o PLS por sinergismo de intervalos (siPLS), que é a associação de intervalos para melhor descrever a propriedade investigada. Estes métodos facilitam a visualização geral das regiões espectrais, podendo assim serem observados os intervalos que apresentam maiores e menores intensidades de absorbância. (ROCHA, 2009)

Ao construir um modelo, deve-se assegurar de que o conjunto de calibração seja homogêneo. As amostras que demonstrarem anomalias, também conhecidas como *outliers*, devem ser removidas para não prejudicar a previsão. (FERREIRA, 1999)

Para usar o modelo obtido pela PLS é necessário avaliar e validar, para assim poder descrever novas amostras. A validação cruzada (*cross validation*) é baseada na avaliação da magnitude da raiz quadrada dos erros médios de calibração cruzada (RMSECV). Nesse processo buscam-se os melhores valores de número de variáveis latentes no qual a raiz quadrada dos erros médios de previsão (RMSEP) seja menor e o coeficiente de correlação (R^2) seja maior, preferencialmente próximo a 1. (PASSAMAI, 2004; Filgueiras, 2011)

1.6.3. MÍNIMOS QUADRADOS PARCIAIS COM DISCRIMINANTE LINEAR – PLS-DA

O método dos Mínimos Quadrados Parciais com Discriminante Linear (PLS-DA - *Partial Least Squares Discriminant Analysis*) é um método de regressão combinado com uma técnica de classificação. Sua vantagem é que as variáveis latentes são modeladas e são encontrados diferentes padrões de dados, proporcionando a visualização gráfica de diferentes grupos (BALLABIO, CONSONNI, 2013).

Todos os cálculos realizados na execução do PLS são feitos novamente e pelas variáveis latentes, que são combinações lineares dos dados, as amostras são distribuídas em classes. É necessário estabelecer o limiar que será o ponto onde as classes se dividirão. Para o caso de dois grupos, só é preciso apenas um limiar ou ponto crítico.

Na PLS-DA o vetor de cada classe é composto por linhas que são as amostras e colunas que são as informações de cada classe. A atribuição de cada classe obedece a probabilidade de a amostra pertencer a uma ou outra classe. O cálculo por amostra é feito para todas as classes, aquele que apresentar maior probabilidade numérica será a classificação da amostra (BALLABIO, CONSONNI, 2013).

2. OBJETIVOS

2.1. OBJETIVO GERAL

Desenvolver modelos multivariados para determinação de alguns parâmetros físico-químicos para caracterização de petróleos, utilizando dados de medidas espectroscópicas.

2.2. OBJETIVOS ESPECÍFICOS

- Realizar medidas no NIR e MIR de petróleos com propriedades distintas;
- Desenvolver modelos de regressão multivariada para determinação de número de acidez total, número de acidez naftênica, teor de nitrogênio total e teor de nitrogênio básico;
- Comparar a eficiência de diversos modelos multivariados e as técnicas espectroscópicas associadas;
- Avaliar o potencial dos modelos de calibração para inferir as propriedades em amostras de petróleo;
- Validar os modelos.

3. METODOLOGIA

Este capítulo tem como finalidade descrever os métodos para obtenção dos dados das propriedades físico-químicas, espectroscópicas e a modelagem. Assim serão apresentados os procedimentos experimentais, equipamentos e as técnicas quimiométricas utilizadas na realização deste projeto. Algumas informações sobre procedimentos foram omitidas, pois os dados são de propriedade do CENPES, o que os torna sigilosos.

Inicialmente foram feitas medidas espectroscópicas utilizando NIR e MIR, como também determinações das propriedades físico-químicas de petróleos. Dentre os parâmetros de caracterização de petróleos utilizados em uma avaliação, foram selecionados quatro para estudo: Número de Acidez Total (NAT), Número de Acidez Naftênica (NAN), Teor de Nitrogênio Total (TNT) e Teor de Nitrogênio Básico (TNB). A escolha dos parâmetros foi feita com base em estudos e resultados preliminares.

Na etapa posterior foram extraídas informações relevantes dos métodos empregados por meio de quimiometria. O desenvolvimento de cada modelo para determinação dos parâmetros seguiu o protocolo interno elaborado no grupo de Quimiometria para avaliação e validação dos modelos multivariados.

Em seguida, foi realizado um estudo comparativo entre os modelos obtidos através dos dados de NIR e MIR, na tentativa de buscar o mais adequado para resolução da maioria dos problemas e identificando a forma que melhor descreve o conjunto de dados.

3.1. CARACTERIZAÇÃO FÍSICO-QUÍMICA

As amostras de petróleo bruto, um total de 99, utilizadas no desenvolvimento dos modelos quimiométricos são provenientes de diversos campos de produção do território brasileiro com o °API variando de 11,2 a 40,5. Os resultados dos parâmetros físico-químicos foram obtidos no Centro de Pesquisa e Desenvolvimento da Petrobras (CENPES), no Rio de Janeiro, Brasil. Para um melhor entendimento dos resultados obtidos, os métodos para obtenção dos parâmetros serão rapidamente descritos a seguir.

3.1.1. NÚMERO DE ACIDEZ TOTAL

Esta propriedade físico-química foi aferida experimentalmente seguindo a norma ASTM D 664 - *Standard Test Method for Acid Number of Petroleum Products by Potentiometric Titration*. A determinação aconteceu por meio de uma titulação potenciométrica com titulador automático Metrohm com uma unidade de dosagem, agitador magnético e o eletrodo Solvotrode easyClean. A solução alcoólica de hidróxido de potássio 0,1 mol.L⁻¹ foi o titulante. O eletrodo utilizado foi o de membrana com solução de cloreto de lítio saturada em etanol. As amostras foram solubilizadas em uma mistura de solventes contendo tolueno, propan-2-ol (isopropanol) e uma pequena porção de água. Foram utilizados aproximadamente 40 gramas de cada amostra por análise. Os resultados são dados em mg KOH/g de amostra.

A repetibilidade e reprodutibilidade do método padrão depende da média ($\bar{y}_i$) de resultados por amostra i , e são dadas pelas equações:

$$repetibilidade_i = 0.044(\bar{y}_i + 1) \quad (5)$$

$$reprodutibilidade_i = 0.141(\bar{y}_i + 1) \quad (6)$$

A variação das determinações do número de acidez total estava na faixa de 0,03 a 3,24 mg KOH/g.

3.1.2. NÚMERO DE ACIDEZ NAFTÊNICA

Este parâmetro é um método interno da Petrobras, assim não se tem muitas informações quanto ao procedimento do mesmo. Sabe-se que esta quantificação de ácidos naftênicos em petróleo ou frações envolve a extração em fase sólida (SPE) e depois os extratos são lidos por absorção no infravermelho médio.

Uma curva de calibração é construída em torno do máximo de absorvidade no comprimento de onda 1710 cm^{-1} e partir daí é realizado um cálculo para determinação da propriedade. Os resultados são expressos na mesma unidade do número de acidez total, mg KOH/g de amostra.

A reprodutibilidade é da mesma ordem de grandeza da ASTM D664 e a repetibilidade média é de 5%, sendo este valor consistente com a acidez total da amostra correspondente.

A faixa de variação do número de acidez naftênica foi de 0,02 a 2,51 mg KOH/g.

3.1.3. TEOR DE NITROGÊNIO TOTAL

Esta metodologia segue a norma ASTM D 4629 - *Standard Test Method for Trace Nitrogen in Liquid Petroleum Hydrocarbons by Syringe/Inlet Oxidative Combustion and Chemiluminescence Detection*.

O método é adequado para análise de óleos e destilados contendo 0,3 a 100 mg/Kg de nitrogênio total. Inicialmente constrói-se uma curva de calibração abrangendo a faixa de operação. O volume de injeção da amostra pode variar entre 3 e 100 mL, de acordo com o instrumento Antek série 9000. A amostra é introduzida em um equipamento com fluxo de gás inerte (hélio ou argônio), onde ela é vaporizada e enviada a um local com alta temperatura. Neste local, o oxigênio é introduzido e o nitrogênio que estava ligado organicamente é convertido em óxido nítrico (NO) e depois em dióxido de nitrogênio (NO₂). O decaimento da excitação do NO₂ gera uma luz que é detectada pelo dispositivo fotomultiplicador e o sinal resultante é o teor de nitrogênio total contido na amostra analisada. A média da triplicata é reportada como resultado final e o teor

de nitrogênio da amostra é expresso em mg/kg.

A repetibilidade e reprodutibilidade da norma ASTM correspondente dependem da média de dois resultados (X) e são dadas pelas seguintes equações:

$$repetibilidade = 0,1825 (X)^{0,5149} \quad (7)$$

$$reprodutibilidade = 0,804 (X)^{0,5149} \quad (8)$$

O intervalo de abrangência do teor de nitrogênio total foi de 0,002 a 0,470 % m/m.

3.1.4. TEOR DE NITROGÊNIO BÁSICO

Este método é utilizado para determinar o nitrogênio básico em petróleo e destilados usando titulação potenciométrica com titulador automático Metrohm, seguindo a norma UOP 269 - *Nitrogen Bases in Hydrocarbons by Titration*. A solução de ácido perclórico 0,02 N é utilizada como titulante e os solventes são tolueno e ácido acético glacial.

Na referida norma foi dada uma tabela indicando a quantidade de amostra a ser pesada de acordo com a quantidade de nitrogênio básico a ser esperada, a variação está na faixa de 1 a 100 g. A amostra foi aquecida em placa térmica e agitadora para garantir a homogeneidade do sistema. Foi adicionado de 25 ou 50 mL de tolueno, seguindo a indicação da norma de concentração de nitrogênio básico na amostra. Foram adicionados, lentamente, 75 mL de ácido acético glacial se não houver precipitação. Caso ocorra precipitação o total de ácido acético glacial deve ser a quantidade de amostra mais o volume de tolueno adicionado. Esta mistura é titulada com solução de ácido perclórico 0,02 N e anotado o teor de nitrogênio básico da amostra em questão.

O intervalo de trabalho para teor de nitrogênio básico foi de 0,016 a 0,168 % m/m.

3.2. INSTRUMENTAÇÃO

Os ensaios espectroscópicos foram realizados no NCQP/LabPetro - Núcleo de Competências em Química do Petróleo, localizado na UFES, campus de Goiabeiras, no Espírito Santo.

3.2.1. INFRAVERMELHO MÉDIO

A aquisição dos espectros de infravermelho na região do médio ocorreu através de um espectroscópio do fabricante Perkin Elmer, modelo Spectrum 400 (Figura 3) e o acessório utilizado foi uma célula horizontal de ATR constituída de ZnSe 45° fabricado pela Pike Technologies (Figura 4).

Foram coletados 32 *scans* para cada amostra, em triplicata autêntica, com resolução de 4 cm^{-1} e intervalo de leitura de 4000-650 cm^{-1} . O tempo para analisar as triplicatas foi, aproximadamente, 10 minutos e foi necessário 1 mL da amostra de interesse.



Figura 3: Espectroscópio infravermelho Perkin Elmer - Spectrum 400.

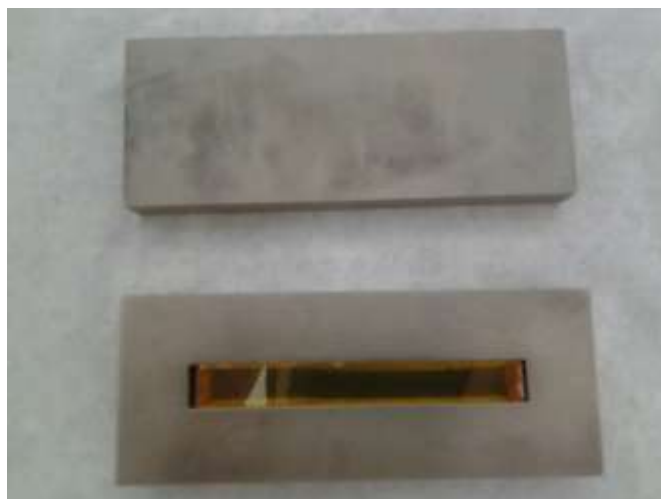


Figura 4: Célula de multi-ATR para leituras na região do MIR.

3.2.2. INFRAVERMELHO PRÓXIMO

Os espectros de infravermelho próximo foram obtidos em um equipamento Perkin Elmer, modelo Spectrum 400, célula de transreflectância Icrá Liquid Sampler do mesmo fabricante (Figura 5).

Cada amostra foi analisada em triplicata autêntica e foram coletados 128 *scans* com resolução de 8 cm^{-1} e intervalo de leitura de $10000\text{--}4000\text{ cm}^{-1}$. O volume de amostra total foi de aproximadamente 0,5 mL e o tempo de análise das triplicatas foi 1,5 minutos para NIR.



Figura 5: Célula de transfectância com petróleo utilizada para leitura na região do NIR.

3.3. MODELOS QUIMIOMÉTRICOS

Para estimar os parâmetros de caracterização do petróleo a partir de técnicas espectroscópicas selecionadas, foi necessário a criação de modelos de regressões multivariadas que relacionaram as respostas instrumentais com cada propriedade físico-química desejada. A construção dos modelos quimiométricos esteve baseada nos seguintes métodos: PLS, iPLS, siPLS e PLS-DA.

Os modelos foram desenvolvidos utilizando o software Matlab® versão 7.8 (The Mathworks, Natick, MA, USA), com o pacote itoolbox versão 8 (iToolbox for Matlab, USA) e algoritmos internos desenvolvidos pelo Grupo de Quimiometria do LabPetro (FILGUEIRAS, ALVES, SAD et al., 2014a).

3.3.1. PROTOCOLO DE MODELAGEM

Durante o desenvolvimento deste trabalho o grupo de Quimiometria elaborou um protocolo de modelagem e padronização para a predição de propriedades utilizando dados de MIR e NIR. Este contém a sequência avaliativa para a construção dos modelos, identificação de *outliers*, critérios de aceitação e alternativas para otimização dos modelos obtidos. Esta metodologia diminui a subjetividade dos modelos e aumenta a transparência em suas construções. O protocolo interno de modelagem MIR e NIR foi baseado em estatística básica, na norma ASTM E1655 e no artigo de OLIVIERI e colaboradores (2006) que descreve a estimação de incertezas e figuras de mérito em modelos multivariados.

3.3.1.1. AVALIAÇÃO DOS ESPECTROS

Este tipo de avaliação é visual, ou seja, os espectros foram analisados com a finalidade de serem observados o comportamento geral dos óleos. Os petróleos que apresentaram perfil diferente do conjunto amostral foram excluídos.

Após esta etapa, foi construído um modelo de classificação por Análise de Componentes Principais - PCA (do inglês, Principal Component Analysis) para

averiguação da dispersão das amostras. Desta forma, as réplicas deveriam estar próximas no gráfico de *scores*. Caso contrário, também seriam desconsideradas.

3.3.1.2. DIVISÃO DOS CONJUNTOS CALIBRAÇÃO E PREVISÃO

Cerca de 70% das amostras foram previamente selecionados para a calibração definida, enquanto os 30% restantes para o conjunto de previsão usando o procedimento de ordenar os valores de referência em ordem crescente e escolher uma amostra a cada três, método também conhecido como *persianas*. Foi observado para que as réplicas de cada amostra fossem mantidas nos mesmo conjunto.

3.3.1.3. PRÉ-PROCESSAMENTO DOS ESPECTROS

Os métodos de pré-processamento utilizados foram a variação normal padrão (SNV) e derivação com suavização Savitz-Golay e dados centrados na média (DERIVMEAN).

3.3.1.4. CONSTRUÇÃO DO MODELO

Os modelos PLS foram construídos com a faixa espectral completa MIR e NIR, posteriormente os espectros foram divididos através do iPLS em 10, 20, 30, 40, 50 e 60 intervalos e com o siPLS estes foram combinados entre si em 2 e 3 intervalos.

3.3.1.5. VALIDAÇÃO DO MODELO

Para a validação interna foi utilizada a validação cruzada *venetian blinds* nos modelos testados, estabelecendo um máximo de 10 variáveis latentes. As amostras são separadas de forma sistemática para o cálculo do RMSECV.

Através da observação do gráfico RMSECV pelo número de variáveis latentes (VLs) foi possível selecionar o menor número de variáveis que representavam o modelo sem perda de informação.

Nesta etapa outra condição deveria ser atendida, apenas os modelos que apresentassem porcentagem de variância explicada acima de 70% para as matrizes X e Y poderiam prosseguir na avaliação.

3.3.1.6. IDENTIFICAÇÃO DE AMOSTRAS ANÔMALAS

Os *outliers* foram identificados analisando os valores de *leverage* e resíduo de calibração, quando a amostra apresentava alto valor numérico para os dois simultaneamente, esta era retirada do modelo e obrigatoriamente o processo de validação citado acima era refeito.

3.3.1.7. AVALIAÇÃO DO MODELO

A avaliação dos modelos se deu pela determinação e análise dos seguintes requisitos: valor da correlação, curva obtida entre os valores previstos versus medidos, erros de validação cruzada e previsão.

O valor numérico de R^2 deveria ser superior a 0,8. Assume-se que valores inferiores a este indicam baixa qualidade preditiva do modelo.

Os pontos amostrais deveriam estar próximos a curva dos valores reais pelos valores de referência. Pontos distantes da curva demonstram inexatidão do resultado.

Por fim, os valores de RMSECV (erro médio quadrático de validação cruzada) e RMSEP (erro médio quadrático de previsão) não deveriam ser estatisticamente diferentes, ou seja, apresentar uma diferença grande em questão de unidades. Estes parâmetros indicam a grandeza dos erros associados aos resultados obtidos.

Os modelos escolhidos foram submetidos a testes estatísticos, e tiveram suas figuras de mérito calculadas, como o intervalo de confiança, sensibilidade, seletividade, limite de detecção e limite de quantificação.

O teste de bias foi aplicado para verificação da presença de erros sistemáticos e para a avaliação de tendência nos resíduos do modelo, foi aplicado o teste de permutação. Quando algum destes testes indicava erro significativo, era realizada uma nova avaliação com a modificação do número

VL.

Muitos modelos foram construídos para cada uma das propriedades físico-químicas estudadas e com base no melhor desempenho foi selecionado independentemente um para MIR e um para NIR por parâmetro. O trabalho seguiu com as comparações entre estes modelos, com a finalidade de ser escolhido apenas um para determinação de tal propriedade de interesse.

4. RESULTADOS E DISCUSSÃO

Nesta parte do trabalho serão apresentados os resultados obtidos durante o desenvolvimento deste estudo.

Quatro propriedades de caracterização de petróleo foram os alvos de interesse deste projeto, que são: Número de Acidez Total (NAT), Número de Acidez Naftênica (NAN), Teor de Nitrogênio Total (TNT), Teor de Nitrogênio Básico (TNB). Estes parâmetros foram escolhidos, pois demonstraram melhores resultados em estudos preliminares realizados com os dados disponíveis.

Foi selecionado um modelo independente para prever cada um dos parâmetros físico-químicos analisados, onde o escolhido passou por testes estatísticos para avaliar a eficiência da sua resposta. Foram calculadas algumas figuras de mérito como intervalo de confiança, sensibilidade, seletividade, limite de detecção (LD), limite de quantificação (LQ), além destes também foi aplicado teste para verificar a presença de erros sistemáticos e tendência nos resíduos. Por fim, os modelos definidos para MIR e NIR por parâmetro foram comparados entre si.

Modelos de calibração multivariada para determinação de propriedades relacionados a compostos ácidos e nitrogenados vem sendo estudados por outros pesquisadores pelo mundo (Jingyan, 2012; Li, 2003). No entanto, são escassos e pouco é dito sobre avaliação de modelos para aplicação industrial. Desta maneira, a proposta deste trabalho foi estudar o desenvolvimento de regressões para cada parâmetro de interesse e simultaneamente verificar o máximo potencial de eficiência em termos de predição.

4.1. DETERMINAÇÃO DE NÚMERO DE ACIDEZ TOTAL (NAT)

O total de 79 óleos brutos foram utilizados como grupo amostral para o desenvolvimento dos modelos para NAT. Os espectros MIR e NIR são mostrados nas Figura 6 e 7, respectivamente. Foi aplicado o método de derivação para corrigir a variação da linha de base para as duas regiões do infravermelho.

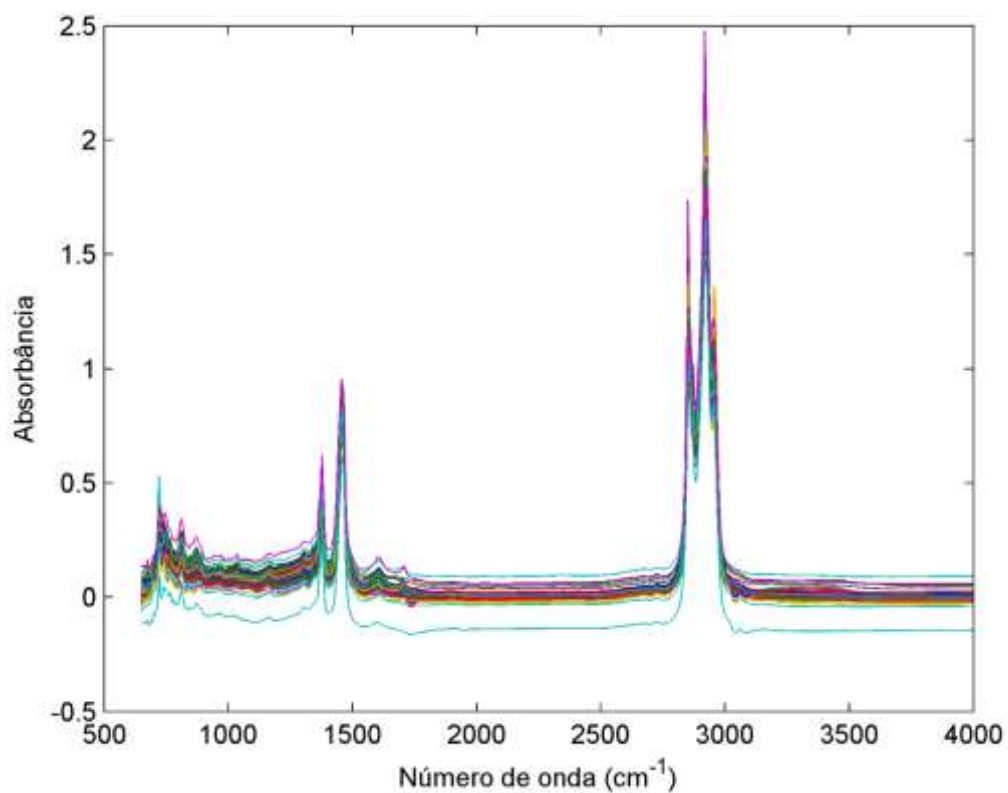


Figura 6: Espectros MIR do conjunto de calibração.

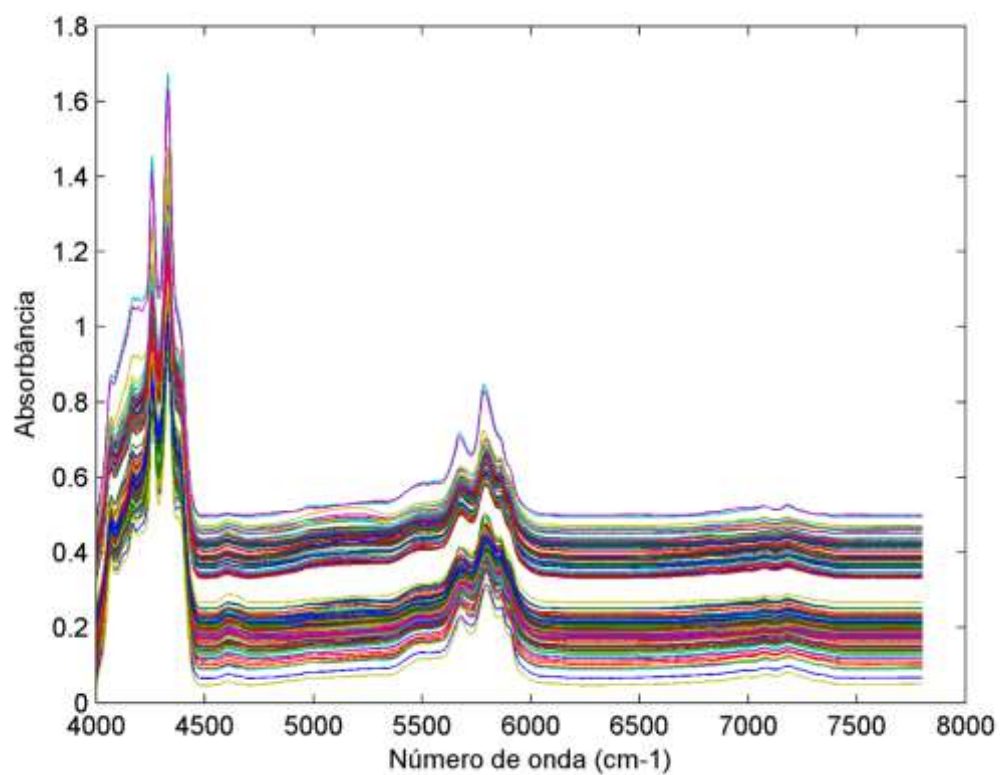


Figura 7: Espectros NIR do conjunto de calibração.

Os modelos PLS foram construídos com seleção de variáveis pelo método de intervalos e sinergismo de intervalos aplicado a partir de dados espectrais MIR e NIR. Dentre todas as regiões analisadas e combinações realizadas, a região espectral 1765-1208 cm^{-1} obteve os melhores resultados. Esta região mostrada na Figura 8 apresenta a característica de deformação axial C=O e deformação angular C-O (PAVIA, LAMPMAN, KRIZ et al., 2010). Este método de seleção de variáveis apresentou uma redução de 3351 para 558 variáveis. Para a região NIR, o melhor modelo foi obtido com espectros inteiros, totalizando 3001 variáveis na matriz X.

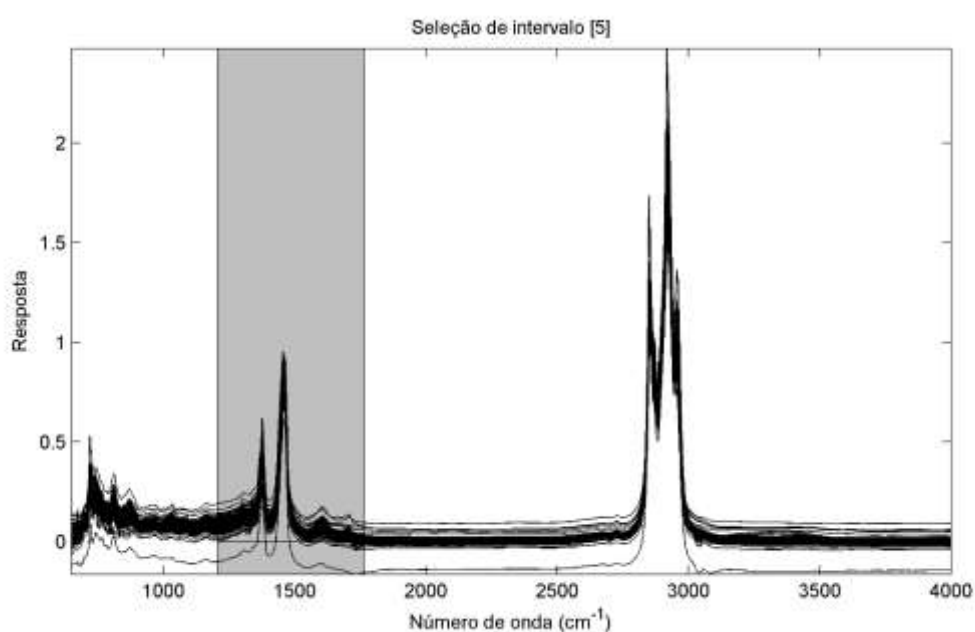


Figura 8: Seleção da região MIR para modelagem da propriedade NAT.

O número de variáveis latentes (VL's) foi escolhido tomando como base os gráficos apresentados na Figura 9. Esta imagem demonstra o decréscimo do valor de RMSECV em relação ao aumento no número de variáveis latentes por região espectral, quando o erro de calibração cruzada para de diminuir significativamente é determinada a quantidade de variáveis a serem utilizadas no modelo de calibração. Foi selecionado 5 variáveis para o modelo MIR e 6 para NIR, quantidades aceitas pelo protocolo utilizado.

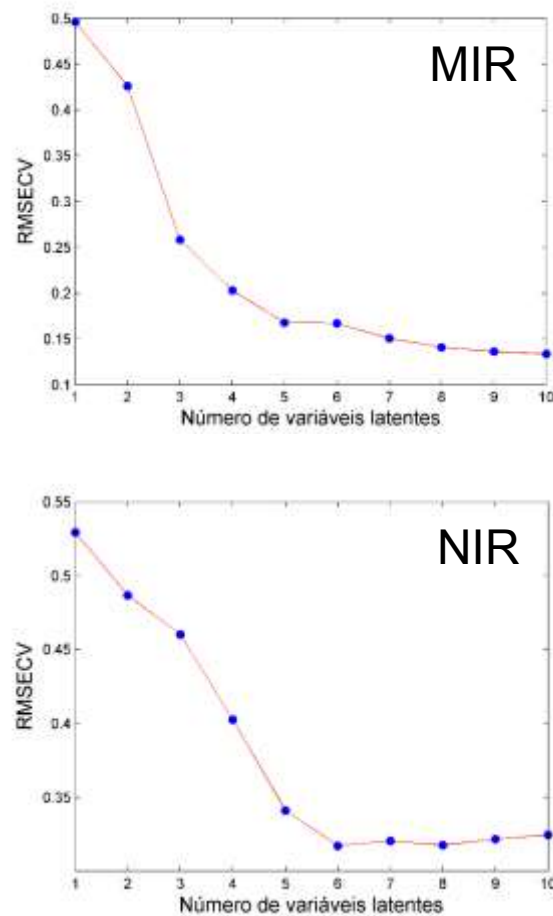


Figura 9: Seleção do número de variáveis latentes nas regiões MIR e NIR para NAT.

O *leverage* detecta a influência de cada amostra no modelo, assim quanto mais uma amostra se difere do perfil médio das amostras na regressão, maior será o seu valor de *leverage* (FERREIRA, 1999). Quando uma amostra apresentava altos valores de *leverage* e resíduo ela foi considerada um *outlier*, e foi retirada do modelo. Segue a Figura 10, representando os *leverages* por amostras do modelo final para MIR e NIR.

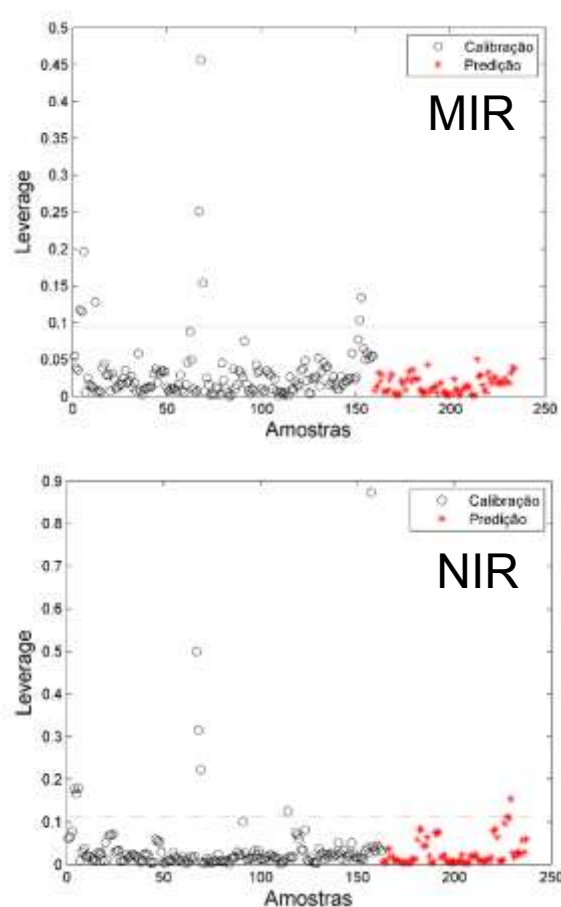


Figura 10: Avaliação de leverage em MIR e NIR para NAT.

A repetibilidade e reprodutibilidade do método padrão ASTM D 664 dependem do valor medido NAT para o óleo de acordo com as equações 5 e 6. Para 25 amostras do conjunto de teste, as reprodutibilidades dos resultados experimentais para o NAT foram na faixa de 0,040 a 2,78 mg de KOH.g⁻¹. As faixas teóricas para repetibilidade e reprodutibilidade variam de 0,046 a 0,166 mg KOH.g⁻¹ e 0,147 a 0,533 mg KOH.g⁻¹, respectivamente. Os RMSEPs obtidos para modelos com MIR e NIR foram 0,196 e 0,203 mg de KOH.g⁻¹ (Tabela 2). Estes valores estão próximos do valor mais baixo de reprodutibilidade, mostrando o potencial de técnicas espectroscópicas para a determinação da NAT.

Tabela 2. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação do Número de Acidez Total.

Parâmetro\Técnica espectroscópica	MIR	NIR
Região espectral (cm^{-1})	1765-1208	10000-4000
Número de variáveis latentes	5	6
Pré-processamento	derivmean	derivmean
RMSEC ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,1561	0,1935
RMSECV ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,1677	0,3172
RMSEP ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,1964	0,2022
R^2 calibração	0,9546	0,9320
R^2 calibração cruzada	0,9456	0,8138
R^2 predição	0,9019	0,8672
Bias calibração ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,0000	0,0000
Bias predição ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,0314	0,0041
Sensibilidade ($\text{mg KOH}\cdot\text{g}^{-1}$) ⁻¹	0,0059	0,0026
LD ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,1365	0,1228
LQ ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,4551	0,4092

Os coeficientes do modelo construído com dados MIR são ligeiramente melhores do que os obtidos com espectros NIR. Na Figura 11, demonstra a boa relação linear entre os valores de NAT e a referência.

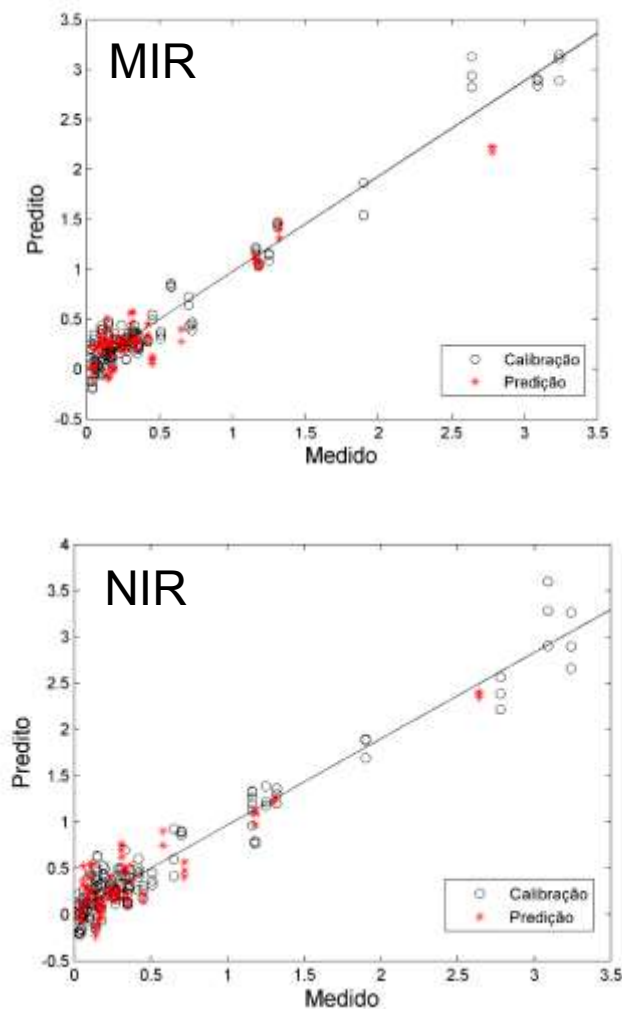


Figura 11: Regressão para NAT em espectros MIR e NIR.

Variabilidade dos dados matriz X e Y para o modelo MIR foi 92,54% e 95,46%, respectivamente. Na Tabela 3 seguem os resultados dos testes de verificação da presença de erros sistemáticos e de tendência nos resíduos. Para esta propriedade de petróleo, não foram encontrados modelos sem tendência nos resíduos. No entanto, os modelos até então construídos não foram descartados, diante da dificuldade em modelar este parâmetro.

Tabela 3. Resultados dos testes de tendência nos resíduos e de permutação para Número de Acidez Total (NAT).

Técnica espectroscópica	MIR	NIR
Tendência calibração	$t_{cal} = 0,00 < t_{tab} = 1,97$	$t_{cal} = 0,00 < t_{tab} = 1,97$
Tendência previsão	$T_{cal} = 1,39 < t_{tab} = 1,99$	$T_{cal} = 0,00 < t_{tab} = 1,99$
p valor calibração	0,0246	0,0034
p valor previsão	0,0000	0,0000

Os limites de detecção para a acidez são da mesma ordem de grandeza do valor mínimo de reprodutibilidade do método de referência para as amostras de teste. No entanto, o limite de modelos de quantificação foram de 0,455 e 0,409 mg KOH.g⁻¹ para o modelo construído com espectros MIR e NIR, respectivamente. Observa-se que a maioria dos resultados de NAT são inferiores a 0,5 mg KOH.g⁻¹. Isto implica que só se mede o número de acidez naftênica em óleos a partir de 0,4 mg KOH.g⁻¹, pois o interesse da indústria é o monitoramento de petróleos com alta acidez.

Jingyan *et al.* (2012) determinou o NAT usando espectroscopia de MIR e de regressão PLS utilizando 280 amostras. Obteve um RMSEP de 0,16 mg KOH.g⁻¹, mas com uma característica de importação, a maioria dos valores de NAT é abaixo de 0,4 mg KOH.g⁻¹. Para a indústria do petróleo é muito importante determinar o NAT para inferir a tendência do óleo a ser corrosivo. Óleos com NAT superior a 0,4 mg KOH.g⁻¹ tendem a ser corrosivo. Este é considerado o ponto crítico onde o petróleo com alto teor de acidez pode colocar em risco a integridade da unidade de destilação, linhas de transferência e equipamentos de refino. Assim, foi construído um modelo de classificação utilizando o método PLS-DA para a determinação de petróleo com baixo índice de acidez (abaixo de 0,4 mg KOH.g⁻¹).

4.1.1. MODELO DE CLASSIFICAÇÃO POR PLS-DA

O modelo classificatório PLS-DA com MIR e NIR foi construído com 3 e 5 variáveis latentes, respectivamente. As Figuras 12 e 13 mostram os resultados dos cálculos para determinação do grupo de petróleo com baixa acidez usando dados espectrais de NIR e MIR.

Comparativamente ao modelo de regressão para NAT por PLS com dados MIR, o modelo classificatório apresentou melhores resultados.

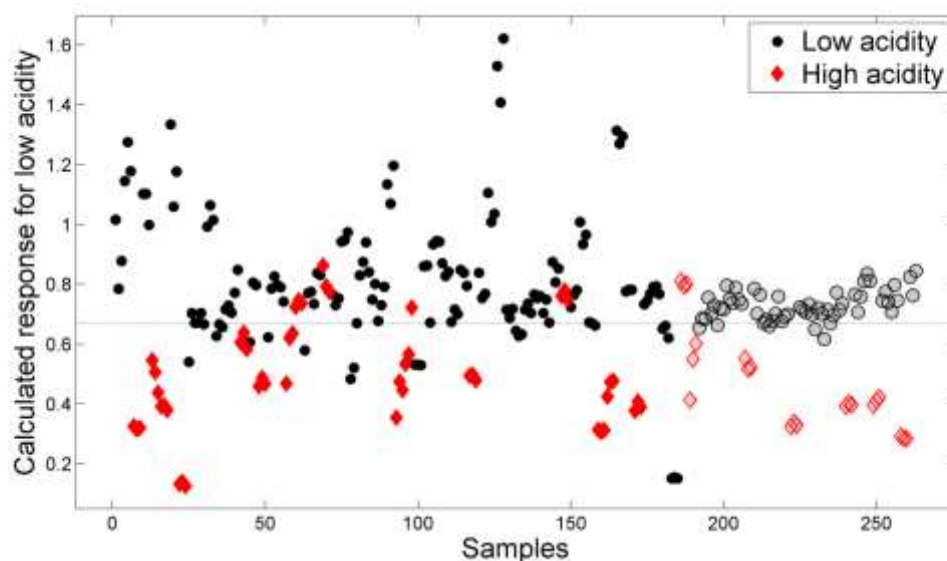


Figura 12: Resposta calculada para baixa acidez pelo modelo PLS-DA utilizando espectros MIR. Amostras (●) de treinamento e (○) amostras de teste com baixa acidez; (◆) treinamento e (◇) amostras teste com alta acidez.

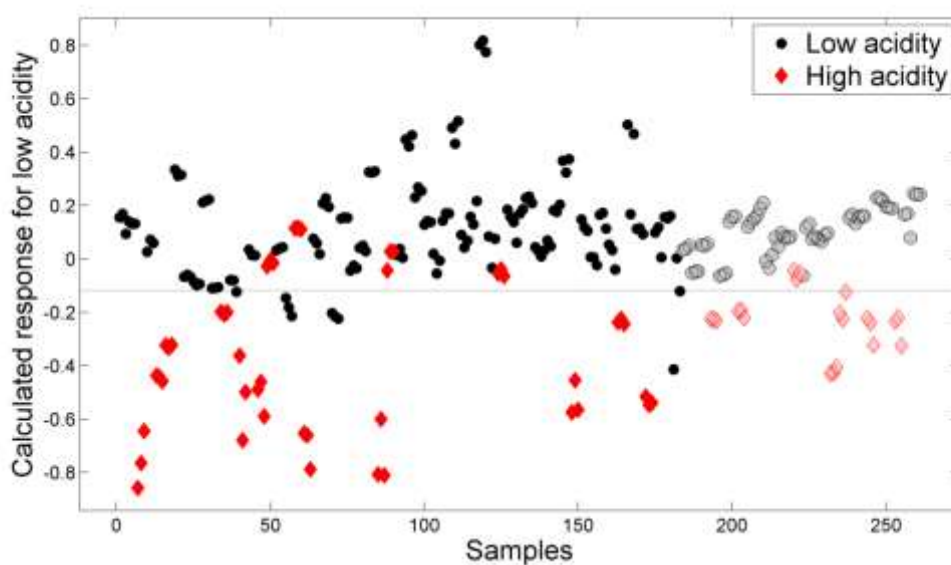


Figura 13: Resposta calculada para baixa acidez pelo modelo PLS-DA utilizando espectros NIR. Amostras (●) de treinamento e (○) amostras de teste com baixa acidez; (◆) treinamento e (◇) amostras de teste com alta acidez.

A sensibilidade do modelo usando espectros MIR foram ligeiramente superiores aos resultados obtidos utilizando os espectros de NIR (Tabela 4). No

entanto, a sensibilidade e precisão para o modelo com espectros NIR foi melhor, onde este apresentou especificidade e acurácia mais elevados.

Tabela 4. Resultados do modelo de classificação PLS-DA.

Parâmetros	MIR		NIR	
	Calibração	Predição	Calibração	Predição
Sensibilidade	0,81	0,86	0,73	0,86
Especificidade	0,92	0,98	0,93	1,00
Acurácia	0,89	0,92	0,89	0,96

Na Figura 12, observa-se que três amostras de previsão com elevada acidez são classificadas como baixa acidez, ou seja, estão classificadas erroneamente. Na Figura 13 também há falsos negativos para as amostras da previsão. Assim, o modelo classificatório para NAT é mais eficaz para os espectros de NIR para a identificação de amostras com elevada acidez.

Para a indústria de petróleo é muito importante determinar amostras com um NAT alto, porque estas amostras contém uma sub-classe de ácido, chamado de ácido naftênico (AN). O AN é a resposta pela corrosão naftênicos nas refinarias. É usual medir o Número de Acidez Naftênica (NAN) apenas em petróleo bruto com elevado NAT.

4.2. DETERMINAÇÃO DE NÚMERO DE ACIDEZ NAFTÊNICA (NAN)

Na Tabela 5 são mostrados os resultados obtidos para os melhores modelos na região espectral MIR e NIR para a determinação do Teor de Acidez Naftênica (NAN). Foram utilizadas 38 amostras, sendo 26 para o conjunto de calibração e 12 para o conjunto de predição.

Tabela 5. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação de Número de Acidez Naftênica (NAN).

Parâmetro\Técnica espectroscópica	MIR	NIR
Região espectral (cm^{-1})	3165-2980 2607-2236	10000-4000
Número de variáveis latentes	5	4
Pré-processamento	derivmean	derivmean
RMSEC ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,3579	0,1888
RMSECV ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,3878	0,2733
RMSEP ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,8453	0,3547
R^2 calibração	0,6972	0,9036
R^2 calibração cruzada	0,6182	0,7852
R^2 predição	0,1137	0,7713
Bias calibração ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,0000	0,0000
Bias predição ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,4211	0,1725
Sensibilidade ($\text{mg KOH}\cdot\text{g}^{-1}$) ⁻¹	0,0006	0,0064
LD ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,1311	0,0445
LQ ($\text{mg KOH}\cdot\text{g}^{-1}$)	0,4370	0,1482

A derivada com suavização Savitz-Golay e dados centrados na média apresentou melhores resultados para duas regiões independentes. O modelo com toda a região espectral NIR foi escolhido (3001 variáveis), enquanto para o MIR o siPLS selecionou três regiões 3165-2980 cm^{-1} (deformação angular O-H), 2607-2422 cm^{-1} e 2421-2236 cm^{-1} (estiramento C-H) (Figura 14), totalizando uma redução de 3351 para 558 variáveis.

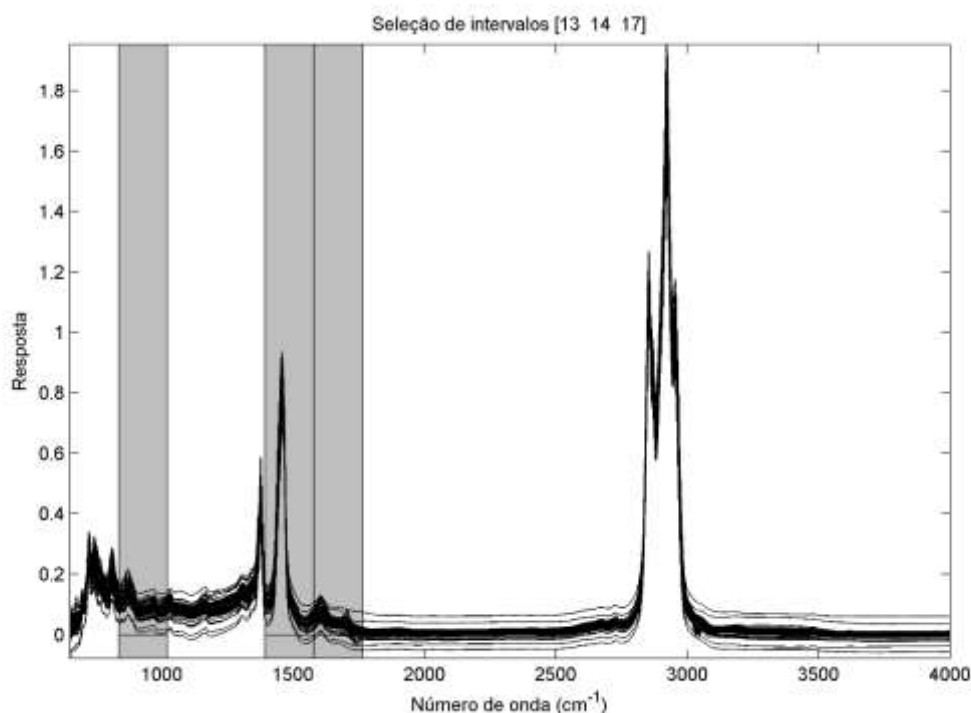


Figura 14: Seleção da região espectral em MIR para determinação de NAN.

Na Tabela 6, estão os resultados da análise de erros sistemáticos e de tendência nos resíduos de calibração e predição. Os dois modelos indicaram a presença de erros sistemáticos na previsão, além de também apresentarem tendência nos resíduos.

Tabela 6. Resultados dos testes de tendência nos resíduos e de permutação para Número de Acidez Naftênica (NAN).

Técnica espectroscópica	MIR	NIR
Tendência calibração	$t_{cal} = 0,00 < t_{tab} = 1,99$	$t_{cal} = 0,00 < t_{tab} = 1,99$
Tendência previsão	$T_{cal} = 3,40 > t_{tab} = 2,03$	$T_{cal} = 3,20 > t_{tab} = 2,04$
p valor calibração	0,0000	0,0000
p valor previsão	0,4733	0,0000

O modelo construído com espectros NIR mostra superioridade comparado ao modelo com os resultados de espectros de MIR. O RMSEC e RMSEP para o modelo com espectros NIR foram $0,189 \text{ mg KOH.g}^{-1}$ e $0,355 \text{ mg de KOH.g}^{-1}$ em contraste com $0,358 \text{ mg de KOH.g}^{-1}$ e $0,845 \text{ mg de KOH.g}^{-1}$ para o modelo com espectros MIR.

Na Figura 15 são mostradas as relações lineares dos valores medidos e previstos dos modelos MIR e NIR. As respostas para o modelo MIR parecem estar aleatoriamente distribuídas, com um coeficiente de correlação 0,697 e 0,114 para os conjuntos de calibração e de predição, respectivamente. No entanto, o modelo com espectros NIR demonstram uma boa relação linear, com um R^2 de 0,904 e 0,771 para a calibração e predição, respectivamente. O bloco X do modelo NIR explica 92,58% da variabilidade dos dados e o bloco Y 90,36%. Embora o modelo NIR seja melhor que o MIR, este ainda apresenta tendência nos resíduos e erros sistemáticos devendo ser utilizado com cautela e ainda estudado métodos para melhorá-lo.

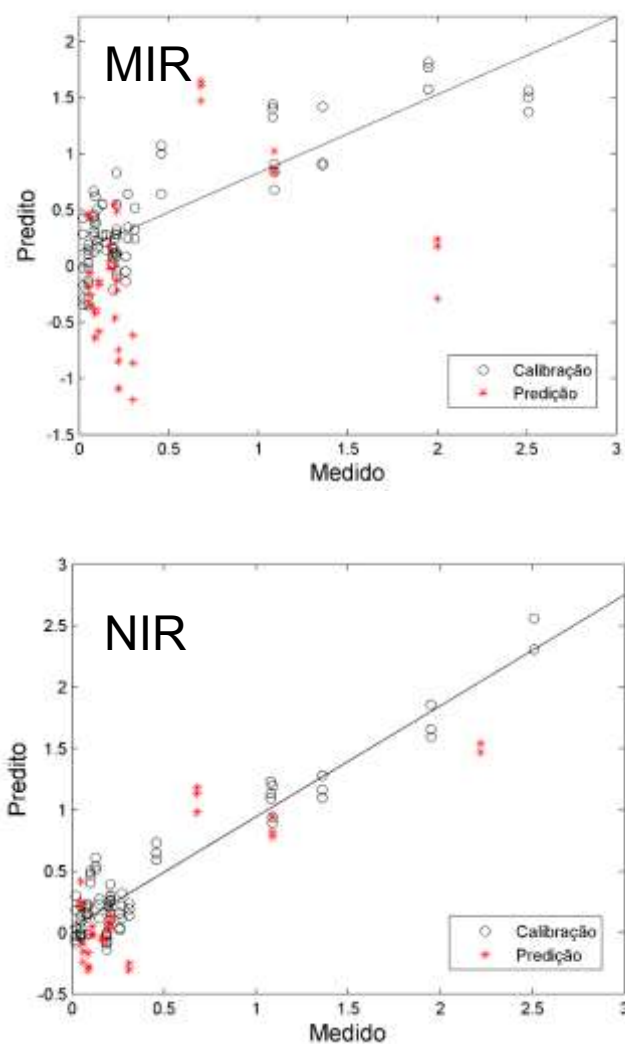


Figura 15: Regressão para NAN em espectros MIR e NIR.

Os limites de detecção (LD) e de quantificação (LQ) para o modelo com espectros NIR foram baixos, 0,045 mg de KOH.g⁻¹ e 0,148 mg de KOH.g⁻¹ respectivamente, mostrando o potencial da técnica NIR associada com calibração multivariada para determinar o Número de Acidez Naftênica no petróleo bruto.

4.3. DETERMINAÇÃO DO TEOR DE NITROGÊNIO TOTAL (TNT)

Na predição do Teor de Nitrogênio Total (TNT) os modelos que apresentaram melhores resultados estão descritos na Tabela 7, onde aquele construído com espectros MIR obteve um desempenho comparativamente superior ao NIR e, portanto, essa foi a técnica espectroscópica escolhida para a previsão desta propriedade.

Tabela 7. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação de Teor de Nitrogênio Total (TNT).

Parâmetro\Técnica espectroscópica	MIR	NIR
Região espectral (cm ⁻¹)	1648-1565 850-784	6000-5500
Número de variáveis latentes	5	5
Pré-processamento	derivmean	Snv
RMSEC (% m/m)	0,0320	0,0406
RMSECV (% m/m)	0,0331	0,0415
RMSEP (% m/m)	0,0381	0,0443
R ² calibração	0,9068	0,8525
R ² calibração cruzada	0,8964	0,8397
R ² predição	0,8383	0,7622
Bias calibração (% m/m)	0,0000	0,0000
Bias predição (% m/m)	0,0039	0,0071
Sensibilidade (% m/m)	0,0129	0,9472
LD (% m/m)	0,0286	0,0196
LQ (% m/m)	0,0952	0,0654

Foram utilizadas 76 amostras, em triplicatas, no desenvolvimento das

regressões, sendo 52 para o conjunto de calibração e 24 para a predição. Nas duas regiões espectrais foram utilizados o mesmo número de variáveis latentes, 5, como segue na Figura 16.

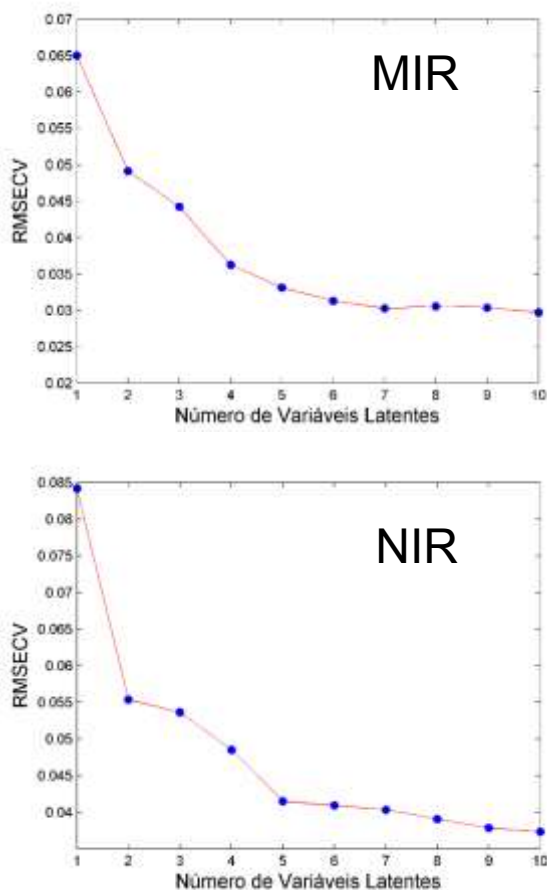


Figura 16: RMSECV versus número de variáveis do Teor de Nitrogênio Total para MIR e NIR.

Analizando ambos modelos através dos testes de verificação de tendência nos resíduos e pelo teste permutação ficou comprovado que o modelo MIR atende a todos os requisitos para aplicação a dados externos (Tabela 8). Nenhum dos modelos indicou a presença de erros sistemáticos, pois os valores de t calculado foram menores que os valores de t tabelado. No entanto, o modelo NIR apresentou tendência no resíduo de calibração (p valor 0,0022), ou seja, valor inferior a 0,05.

Tabela 8. Resultados dos testes de tendência nos resíduos e de permutação de Teor de Nitrogênio Total (TNT).

Técnica espectroscópica	MIR	NIR
Tendência calibração	$t_{cal} = 0,00 < t_{tab} = 1,97$	$t_{cal} = 0,00 < t_{tab} = 1,97$
Tendência previsão	$T_{cal} = 0,87 < t_{tab} = 1,99$	$T_{cal} = 1,39 < t_{tab} = 1,99$
p valor calibração	0,2362	0,0022
p valor previsão	0,0679	0,3149

Abaixo segue a Figura 17 mostrando as regiões espectrais selecionadas na faixa do infravermelho médio ($1648\text{-}1565\text{ cm}^{-1}$ e $850\text{-}784\text{ cm}^{-1}$). Estes intervalos correspondem a ligação N-H (deformação angular) que resultam em bandas nas faixas de $1640\text{-}1560\text{ cm}^{-1}$ e 800 cm^{-1} (PAVIA, LAMPMAN, KRIZ et al., 2010). O percentual de variância acumulada para o modelo com este conjunto de dados foi de 95,00% na matriz X e 90,68% na matriz Y.

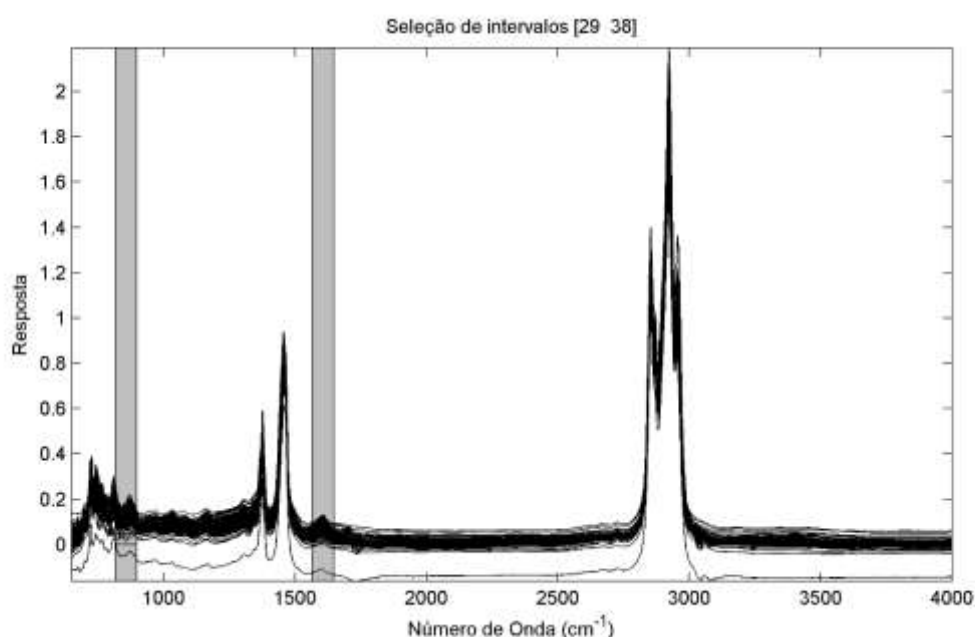


Figura 17. Região espectral para o modelo MIR do Teor de Nitrogênio Total.

Foi analisado o *leverage* e o resíduo para cada amostra, quando os dois resultados apresentavam altos valores, simultaneamente, a triplicata dessa amostra era retirada do modelo. A Figura 18 mostra o gráfico de *leverage* versus amostras para o modelo MIR que passou em todos os critérios de seleção do

protocolo.

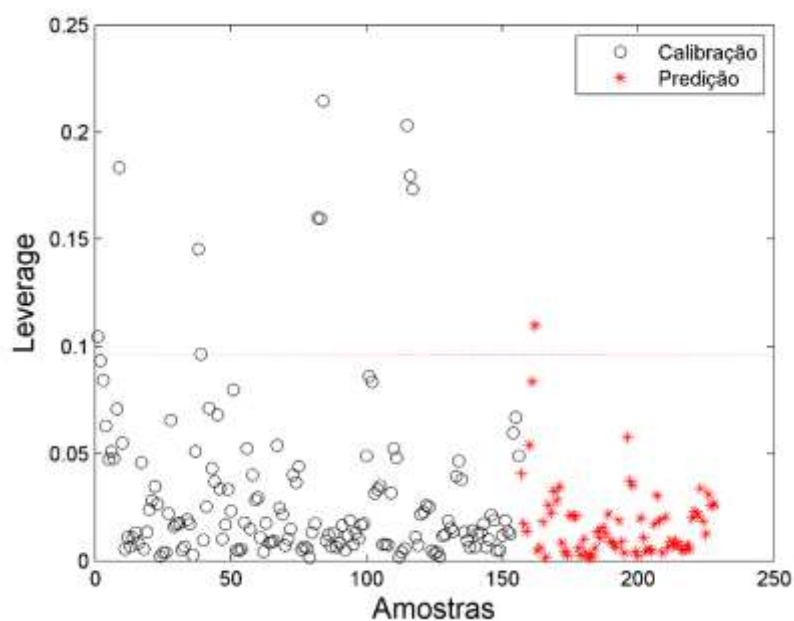


Figura 18: Análise de *leverage* do Teor de Nitrogênio Total.

Nas Figuras 19 e 20 são encontradas a regressão e as barras de incerteza, respectivamente, para a determinação do Teor de Nitrogênio Total na região do MIR. Os resultados são satisfatórios (Tabelas 7 e 8), com um erro de predição de 0,0381% m/m e as regiões espectrais escolhidas caracterizam a presença de nitrogênio, desta forma mostrando compatibilidade ao ser relacionado quimiometricamente a propriedade em estudo.

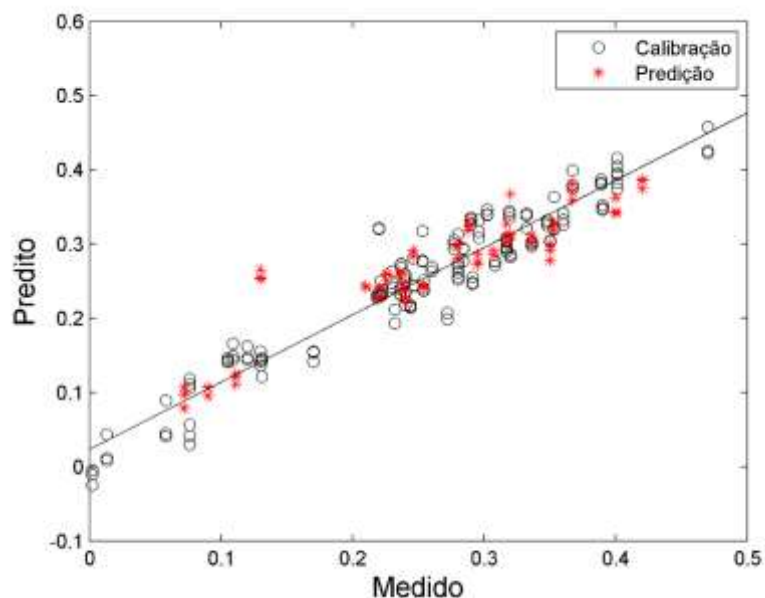


Figura 19: de regressão MIR para Teor de Modelo Nitrogênio Total.

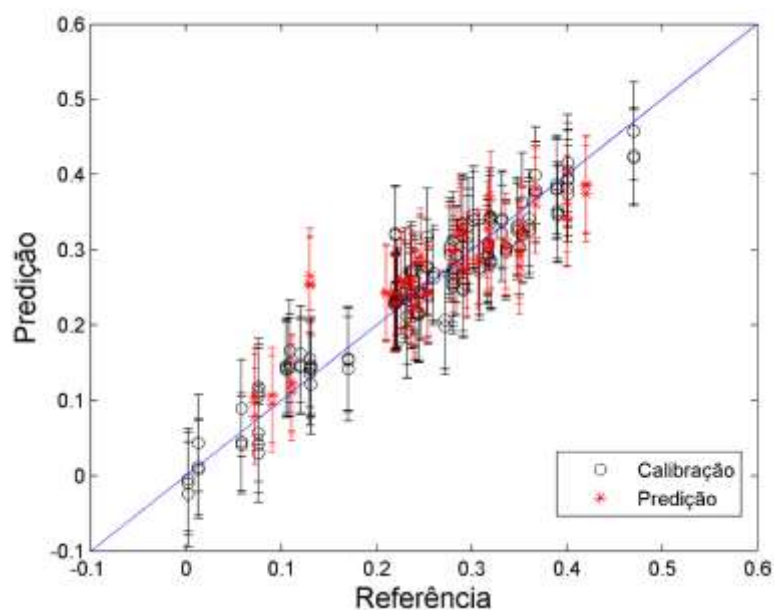


Figura 20: Regressão com barras de incerteza para Teor de Nitrogênio Total.

4.4. DETERMINAÇÃO DO TEOR DE NITROGÊNIO BÁSICO (TNB)

Para determinação do Teor de Nitrogênio Básico por PLS foi disponibilizado um conjunto contendo 68 amostras em triplicata, onde estas foram divididas em dois grupos: 46 amostras de calibração e 22 amostras de predição. A partir destes dados foram desenvolvidos diversos modelos de

regressão multivariada para as duas regiões do infravermelho em estudo. Dentre todos os modelos construídos, aqueles com seleção de variáveis usando siPLS com combinação de 2 intervalos apresentaram melhores resultados para MIR e NIR. A seleção de variáveis por siPLS em MIR promoveu a redução de 3351 para 333 e em NIR de 3001 para 596.

A regressão gerada através de dados MIR obteve um desempenho mais satisfatório do que o NIR e, portanto, essa foi a técnica escolhida para a predição desta propriedade físico-química de petróleo. Os parâmetros escolhidos para a construção desse modelo e os resultados apresentados estão listados abaixo, onde é possível observar que há concordância com todos os critérios previamente estabelecidos pelo protocolo. Na Tabela 9 seguem os parâmetros estatísticos adquiridos para MIR e NIR.

Tabela 9. Resultados dos parâmetros estatísticos dos modelos PLS para quantificação de Teor de Nitrogênio Básico (TNB).

Parâmetro\Técnica espectroscópica	MIR	NIR
Região espectral (cm ⁻¹)	3496-3329 1651-1485	6098-5800 5798-5500
Número de variáveis latentes	7	6
Pré-processamento	SNV	SNV
RMSEC (% m/m)	0,0094	0,0100
RMSECV (% m/m)	0,0099	0,0103
RMSEP (% m/m)	0,0073	0,0111
R ² calibração	0,9330	0,9222
R ² calibração cruzada	0,9211	0,9139
R ² predição	0,9466	0,8601
Bias calibração (% m/m)	0,0000	0,0000
Bias predição (% m/m)	0,0002	0,0007
Sensibilidade (% m/m)	7,5968	2,3385
LD (% m/m)	0,0030	0,0035
LQ (% m/m)	0,0099	0,0118

O pré-processamento aplicado aos dois modelos foi a variação normal padrão (SNV), que é um método empregado para a correção do espalhamento

multiplicativo.

Embora os modelos MIR e NIR tenham apresentado resultados muito semelhantes, após a averiguação da presença de erros sistemáticos e tendência nos resíduos gerados ficou comprovado que o modelo MIR não apresenta nenhum tipo de tendência e o NIR têm problema de tendência na calibração. O p valor do modelo de calibração em NIR foi 0,032, menor do que o valor de referência (0,05). A Tabela 10 apresenta os valores para os cálculos acima citados.

Tabela 10. Resultados dos testes de tendência nos resíduos e de permutação de Teor de Nitrogênio Básico (TNB).

Técnica espectroscópica	MIR	NIR
Tendência calibração	$t_{cal} = 0,00 < t_{tab} = 1,98$	$t_{cal} = 0,00 < t_{tab} = 1,98$
Tendência previsão	$T_{cal} = 1,95 < t_{tab} = 2,00$	$T_{cal} = 0,50 < t_{tab} = 1,98$
p valor calibração	0,0544	0,0320
p valor previsão	0,1545	0,2494

As regiões espectrais para a construção do modelo na região do infravermelho médio estão destacadas na Figura 21, onde também é possível observar os espectros das amostras, em triplicata, que foram utilizadas no conjunto de calibração. A faixa 3496-3329 cm^{-1} indica deformação axial N-H em compostos alifáticos e a região 1651-1485 cm^{-1} corresponde ao dobramento da ligação N-H, deformação angular (PAVIA, LAMPMAN, KRIZ et al., 2010).

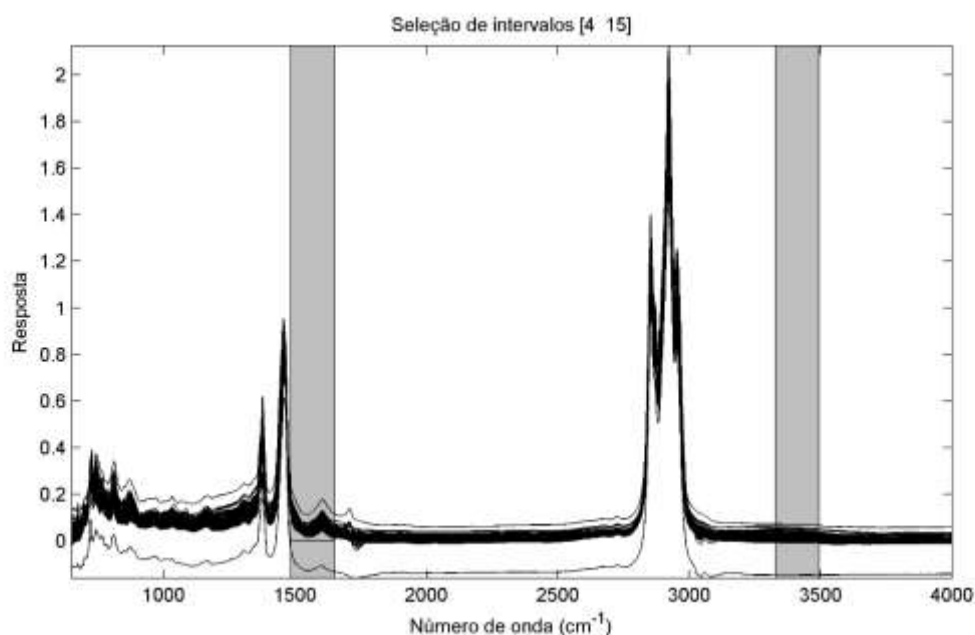


Figura 21: Região espectral para o modelo MIR do Teor de Nitrogênio Básico.

O número de variáveis latentes foi escolhido analisando o gráfico RMSECV versus número de variáveis latentes (Figura 22), onde 7 variáveis representaram muito bem a variabilidade dos dados. O percentual de variância acumulada na matriz X para o modelo foi 95,84% e na matriz Y 93,30%.

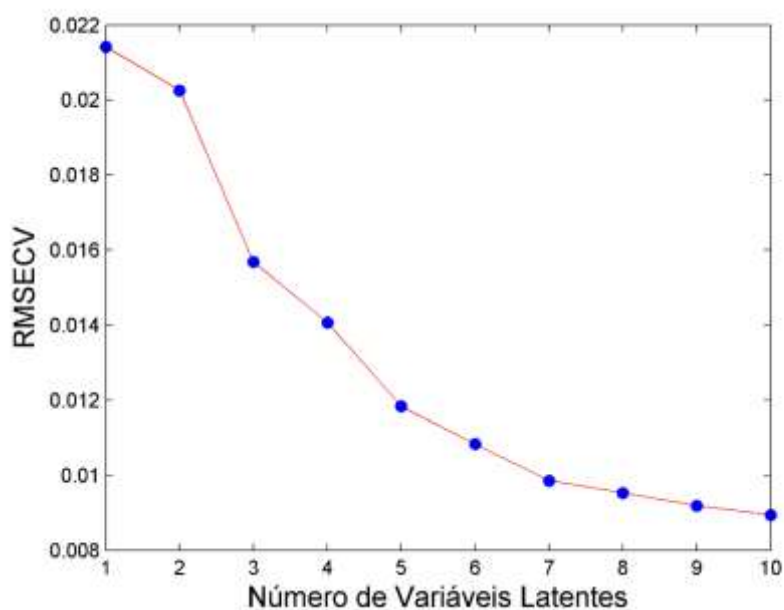


Figura 22: RMSECV versus número de variáveis do Teor de Nitrogênio Básico em MIR.

A Figura 23 mostra o gráfico de *leverage* versus amostras para o modelo MIR que passou em todos os critérios de seleção do protocolo.

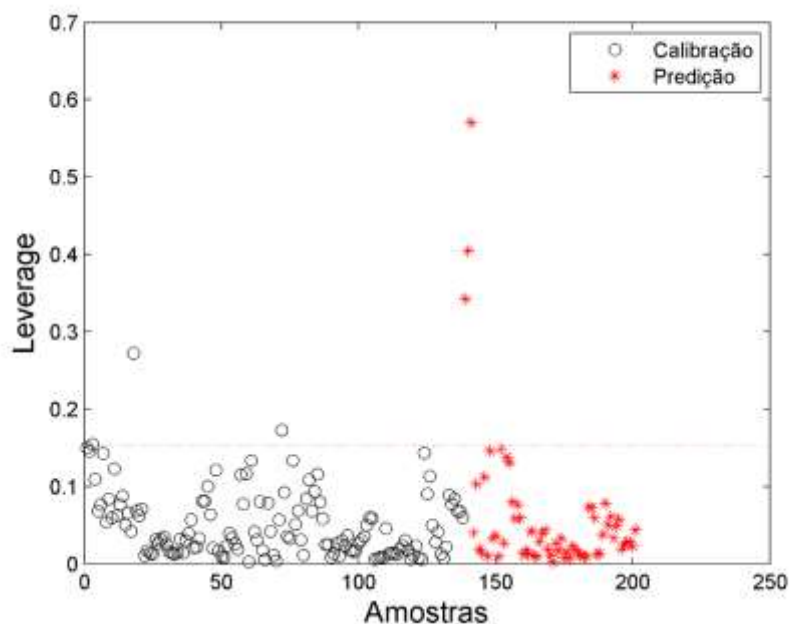


Figura 23: Análise de leverage do Teor de Nitrogênio Básico para MIR.

A relação estabelecida entre os valores medidos versus preditos, e os respectivos intervalos de confiança do modelo MIR selecionado são mostrados nas Figuras 24 e 25, respectivamente.

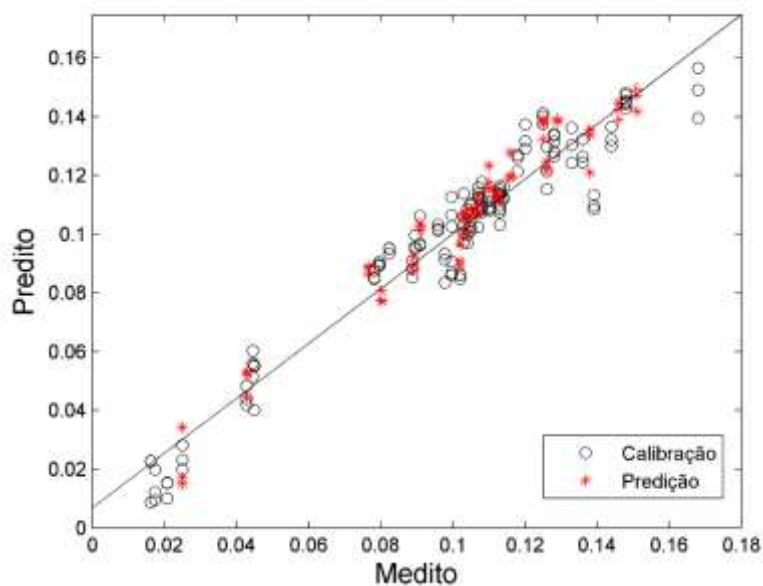


Figura 24: Modelo de regressão MIR para Teor de Nitrogênio Básico.

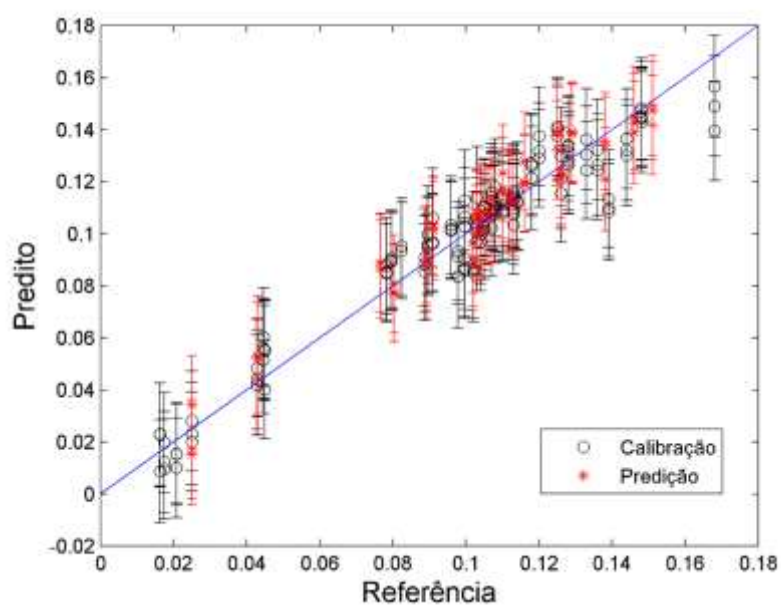


Figura 25: Intervalos de confiança do modelo para a predição do Teor de Nitrogênio Básico.

A partir dos resultados apresentados acima foi possível verificar que o modelo está estatisticamente aprovado, pois atendeu a todos os parâmetros previamente estabelecidos para validação e aceitação dos modelos.

5. CONCLUSÕES

Neste trabalho, quatro parâmetros físico-químicos de petróleos foram estimados por infravermelho médio e próximo usando mínimos quadrados parciais de regressão e análise discriminante. A regressão multivariada por PLS foi utilizado para estimar o Número de Acidez Total (NAT) apresentando limites de quantificação superior a $0,40 \text{ mg KOH.g}^{-1}$, que é o valor que a indústria do petróleo geralmente aceita como o limite para o óleo a ser considerada não-corrosivo. Assim, um modelo de classificação para o NAT é mais viável porque permite a distinção de óleos com baixa e alta acidez. O método PLS-DA apresenta-se como uma boa estratégia para este fim, que permite a obtenção de modelos com alta precisão utilizando espectros infravermelhos MIR ou NIR.

O modelo PLS de previsão para o Número de Acidez Naftênica (NAN) em óleos brutos apresenta bons resultados quando se utiliza os espectros de infravermelho próximo, com um erro de predição de $0,355 \text{ mg de KOH.g}^{-1}$, o inverso da sensibilidade analítica de $0,015 \text{ mg de KOH.g}^{-1}$ e baixo limite de quantificação, $0,148 \text{ mg de KOH.g}^{-1}$. O método ASTM D664 não inclui a determinação de ácido naftênico, mas o erro de predição que pode ser obtido a partir do modelo PLS é da mesma ordem de grandeza que a reprodutibilidade do método ASTM D664 para o número total de acidez.

A predição do Teor de Nitrogênio Total em petróleos apresentou ótimos resultados quando foram utilizados os espectros de infravermelho médio, com um erro de predição de $0,0381\% \text{ m/m}$ e baixo limite de quantificação, $0,0952\% \text{ m/m}$. Este modelo não apresenta erros sistemáticos e erros de tendência, passando por todos os critérios de avaliação e apresentando valores satisfatórios para as figuras de mérito calculadas.

O Teor de Nitrogênio Básico também apresentou modelo que foi avaliado positivamente em todos os parâmetros de avaliação. Valor de RMSEP $0,0073\% \text{ m/m}$ e limite de quantificação $0,0030\% \text{ m/m}$.

A obtenção de espectros infravermelhos de petróleo bruto é relativamente simples e rápida. Agilidade nos resultados de inferência com certa confiabilidade é atraente para a indústria de petróleo, a metodologia pode ser usada na rotina

industrial com o objetivo de antecipar e estimar os resultados de acidez em pequenas quantidades de óleo.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALEME, H. G. **Utilização de curvas de destilação e métodos quimiométricos na previsão da origem e na determinação do teor alcoólico e massa específica da gasolina automotiva.** Dissertação de Mestrado. Departamento de Química da Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

ANDRADE, G. H. **Estudo da espectroscopia na região do infravermelho médio e próximo para previsão das propriedades do petróleo e emulsão de petróleo do tipo água em óleo.** Departamento de Química da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009. 1-8 p.

ANDRADE J. M.; FRESCO P.; MUNIATEGUI S.; PRADA D. Comparison of oil spillages using mid-IR indexes and 3-way procrustes rotation, matrix-augmented principal components analysis and parallel factor analysis. *Talanta*, **77**, 863-869, 2008.

ASTM Method D 4629-12. Standard Test Method for Trace Nitrogen in Liquid Petroleum Hydrocarbons by Syringe/Inlet Oxidative Combustion and Chemiluminescence Detection. West Conshohocken, PA, USA: American Society for Testing and Materials, 2012.

ASTM Method D 664-11a. Standard Test Method for Acid Number of Petroleum Products by Potentiometric Titration. West Conshohocken, PA, USA: American Society for Testing and Materials, 2011.

ASTM Method E 1655. Standard Practices for Infrared Multivariate Quantitative Analysis. West Conshohocken, PA, USA: American Society for Testing and Materials, 2005.

BALLABIO D.; CONSONNI V. Classifications tools in chemistry. Part 1: linear models. PLS-DA. *Analytical Methods*, **5**, 3790-3798, 2013.

BRERETON R. G. **Chemometrics:** Data Analysis for the laboratory and chemical plant. 1ª Ed. West Sussex: John Wiley & Sons, 2003. 297-317 p.

BUENO A. F. **Caracterização de petróleo por espectroscopia no infravermelho próximo.** Dissertação de Mestrado. Departamento de Química

da Universidade Estadual de Campinas, Campinas, 2004.

COLATI, K. A. P. **Extração e Caracterização de Ácidos Naftênicos por Espectrometria de Massas de Altíssima Resolução e Exatidão (ESI-FT-ICR MS)**. Dissertação de Mestrado. Departamento de Química da Universidade Federal do Espírito Santo, Vitória, 2012. 19 p.

DIAS, J. C. M. **Estudo exploratório das propriedades de caracterização de petróleos brasileiros**. Dissertação de Mestrado. Departamento de Química da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005. 10 p.

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*, **51**, 127-137, 2006.

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Quimiometria I: calibração multivariada, um tutorial. *Química Nova*, **22**, 724-731, 1999.

FILGUEIRAS, P. R. **Determinação da composição de blends de petróleo utilizando FTIR-ATR e calibração multivariada**. Departamento de Química da Universidade Federal do Espírito Santo, Vitória, 2011. 20-22 p.

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*, **116**, 123-130, 2014.

FILGUEIRAS P. R.; ALVES J. C. L.; SAD C. M. S.; CASTRO E. V. R.; DIAS J. C. M.; POPPI R. J. Evaluation of trends in residuals of multivariate calibration models by permutation test. *Fuel*, **133**, 33-41, 2014a.

GARY, J. H.; HANDWERK, G. E. **Petroleum Refining: Technology and Economics**. 4^a Ed. New York: Marcel Dekker, Inc, 2001. 21-24, 28-30 p.

GRUBER L. D. A.; DAMASCENO F. C.; CARAMÃO E. B.; JACQUES R. A.; GELLERA M.; CAMPOS M. C. V. Ácidos Naftênicos no Petróleo. *Química Nova*, **35**, 1423-1433, 2012.

JINGYAN L.; XIAOLI C.; SONGBAI T. Research on determination of total acid number of petroleum using mid-infrared attenuated total reflection spectroscopy. *Energy Fuels*, **26**, 5633-5637, 2012.

LI B.; LIANG Y.; HU Y.; DU Y.; SONG Y.; CUI H. Evaluation of gas chromatography-mass spectrometry in conjunction with chemometric resolution for identification of nitrogen compounds in crude oil. *Talanta*, **61**, 803-809, 2003.

MACHO S.; LARRECHI M. S. Near-infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry. *Trends in Analytical Chemistry*, **21**, 779-806, 2002.

MELÉNDEZ L. V.; LACHE A.; ORREGO-RUIZ J. A.; PACHÓN Z., MEJÍA-OSPINO E. Prediction of the SARA analysis of Colombian crude oils using ATR-FTIR spectroscopy and chemometric methods. *Journal of Petroleum Science and Engineering*, **90**, 56-60, 2012.

MÜLLER A. L. H.; PICOLOTO R. S.; MELLO P. A.; FERRÃO M. F.; SANTOS M. F. P.; GUIMARÃES R. C. L.; MÜLLER E. I.; FLORES E. M. M. Total sulfur determination in residues of crude oil distillation using FT-IR/ATR and variable selection methods. *Spectrochimica Acta Part A*, **89**, 82-87, 2012.

OLIVIERI AC, FABER NKM, FERRÉ J, BOQUÉ R, KALIVAS JH, MARK H. Uncertainty estimation and figures of merit for multivariate calibration. *IUPAC*, **78**, 633-661, 2006.

PARISOTTO, G. **Determinação do número de acidez total em resíduos de destilação atmosférica e de vácuo do petróleo empregando a espectroscopia no infravermelho (ATR-FTIR) e calibração multivariada.** Dissertação de Mestrado. Departamento de Química da Universidade Federal de Santa Maria, Santa Maria, 2007. 4-5 p.

PARREIRA T. F. **Utilização de métodos quimiométricos em dados de natureza multivariada.** Dissertação de Mestrado. Departamento de Química da Universidade Estadual de Campinas, Campinas, 2003.

PASQUINI C.; BUENO A. F. Characterization of petroleum using near-infrared

spectroscopy: Quantitative modeling for the true boiling point curve and specific gravity. *Fuel*, **86**, 1927-1934, 2007.

PASQUINI, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, **14**, 198-219, 2003.

PAVIA D. L.; LAMPMAN G. M.; KRIZ G. S.; VYVYAN J. R. Introdução à Espectroscopia. 4ª Ed. São Paulo: Cengage Learning, 2010. 15-20, 61-62, 72-73 p.

PEINDER, P. **Characterization and classification of crude oils using a combination of spectroscopy and chemometrics**. PhD thesis. Utrecht University, Netherlands, 2009. 10-12 p.

PEINDER P.; VISSER T.; WAGEMANS R.; BLOMBERG J.; CHAABANI H.; SOULIMANI F.; WECKHUNYSEN B. M. Sulfur speciation of crude oils by partial least squares regression modeling of their infrared spectra. *Energy Fuels*, **24**, 557-562, 2010.

Perfil. Disponível em: www.petrobras.com.br. Acessado em 30 de outubro de 2014.

ROCHA J. T. C. **Emprego de Espectrometria no Infravermelho e Métodos Quimiométricos para Identificação e Quantificação de Petróleos a Partir de Misturas de Frações de Diesel**. Dissertação de Mestrado. Departamento de Química da Universidade Federal do Espírito Santo, Vitória, 2009.

SATYA S.; ROEHNER R. M.; DEO M. D.; HANSON F. V. Estimation of properties of crude oil residual fractions using chemometrics. *Energy and Fuels*, **21**, 998-1005, 2007.

SILVA A. L.; SILVA A. M. S.; RIBEIRO J. C.; MARTINS F. G.; SILVA F. A.; SILVA C. M. Chromatographic and spectroscopic analysis of heavy crude oil mixtures with emphasis in nuclear magnetic resonance spectroscopy: A review. *Analytica Chimica Acta*, **707**, 18-37, 2011.

SPEIGHT, J. G. **Handbook of Petroleum Product Analysis**. 1º Ed. New Jersey:

Wiley-Interscience, 2002. 1-10 p.

THOMAS, J. E. **Fundamentos de Engenharia de Petróleo**. 2ª Ed. Rio de Janeiro: Editora Interciência, 2001. 1-13 p.

UOP Method 269-10. Nitrogen Bases in Hydrocarbons by Titration. West Conshohocken, PA, USA: American Society for Testing and Materials, 2010.