

LÍDIA RENATA ZANÃO

Desenvolvimento de método para previsão de parâmetros físico-químicos do óleo diesel comercial por cromatografia gasosa ultrarrápida – CG-UR e quimiometria

Tese apresentada ao Instituto de Química, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Doutor em Química

Orientador: Prof. Dr. Rodrigo Sequinel

Co-orientador: Prof. Dr. Danilo Luiz Flumignan

Araraquara

2019

#### FICHA CATALOGRÁFICA

Zanão, Lídia Renata  
Z27d      Desenvolvimento de método para previsão de parâmetros físico-químicos do óleo diesel comercial por cromatografia gasosa ultrarrápida - CG-UR e quimiometria / Lídia Renata Zanão. – Araraquara : [s.n.], 2019  
218 f. : il.

Tese (doutorado) – Universidade Estadual Paulista, Instituto de Química  
Orientador: Rodrigo Sequinel  
Coorientador: Danilo Luiz Flumignan

1. Combustíveis diesel. 2. Quimiometria. 3. Cromatografia a gás. 4. Análise multivariada. 5. Controle de qualidade. I. Título.

**CERTIFICADO DE APROVAÇÃO**

**TÍTULO DA TESE:** "Desenvolvimento de método para previsão de parâmetros físico-químicos do óleo diesel comercial por cromatografia gasosa ultrarrápida – CG-UR e quimiometria"

**AUTORA:** LÍDIA RENATA ZANÃO

**ORIENTADOR:** RODRIGO SEQUINEL

Aprovada como parte das exigências para obtenção do Título de Doutora em QUÍMICA, pela Comissão Examinadora:

Prof. Dr. RODRIGO SEQUINEL 

Departamento de Engenharias e Exatas / Universidade Federal do Paraná - UFPR - Palotina

  
Prof. Dr. KELLY JOHANA DUSSAN MEDINA

Departamento de Bioquímica e Tecnologia Química / Instituto de Química - UNESP - Araraquara

  
Prof. Dr. EDILENE CRISTINA FERREIRA

Departamento de Química Analítica / Instituto de Química - UNESP - Araraquara

  
Prof. Dr. ARISTEU GOMES TININI

Campus de Matão / Instituto Federal de São Paulo - IFSP - Matão

  
Profa. Dra. MARIA LUCIA RIBEIRO

Departamento Meio Ambiente / Universidade de Araraquara - UNIARA - Araraquara

Araraquara, 14 de março de 2019

## **DADOS CURRICULARES**

### **1. Identificação**

Nome: Lídia Renata Zanão

Nome em citações bibliográficas: ZANÃO, L.R.

### **2. Formação acadêmica**

Doutorado em Química (2014 – 2019). Instituto de Química, Instituto de Química, Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), Araraquara.

Orientador: Prof. Dr. Rodrigo Sequinel

Co-orientador: Prof. Dr. Danilo Luiz Flumignan

Bolsista: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Formação Complementar Pedagógica para não licenciados em Química (2017 - 2018). Universidade Cruzeiro do Sul.

Mestrado em Ciências (2011 – 2013). Departamento de Química, Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo (USP), Ribeirão Preto.

Orientador: Prof. Dr. Anderson Rodrigo Moraes de Oliveira.

Bolsista: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Graduação em Bacharelado em Química com Habilitação em Química Forense (2006 – 2011). Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo (USP), Ribeirão Preto.

Orientador: Prof. Dr. Celso Teixeira Mender Júnior.

Bolsista: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

### **3. Formação complementar**

#### **2018 – 2018**

Treinamento de Auditores Internos na norma ABNT ISO/IEC 17025:2017 segundo a norma ABNT NBR ISO 19011:2012. (Carga horária: 4h).

Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP.

#### **2018 - 2018**

Curso de Gestão da Qualidade Total (GQT) – Curso online (Carga horária: 15h)

Escola Virtual Portogente

#### **2018 - 2018**

Curso de Oratória – Curso online (Carga horária: 14h).

Escola Virtual Portogente

**2017 - 2017**

Curso de Matlab - Aplicações em Bioprocessos. (Carga horária: 20h).  
Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP. Araraquara, SP, Brasil,

**2015 - 2015**

Seminário Técnico de Cromatografia e Preparo de Amostras. (Carga horária: 8h).  
Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP. Araraquara, SP, Brasil.

**2012 - 2012**

Curso de operação cromatógrafo LC-20AT/ Shimadzu SINC. (Carga horária: 18h).  
Universidade de São Paulo. Ribeirão Preto, SP, Brasil.

**2010 – 2010**

Mini-curso de Controle de Dopagem no Esporte. (Carga horária: 4h)  
Universidade de São Paulo. Ribeirão Preto, SP, Brasil.

**2008 – 2008**

Mini-curso de Drogas Sintéticas – Principais Rotas e Formas de Identificação Forense (Carga horária: 4h).  
Universidade de São Paulo. Ribeirão Preto, SP, Brasil.

**2008 – 2008**

Mini-curso de Produção de Cachaça (Carga horária: 4h).  
Universidade de São Paulo. Ribeirão Preto, SP, Brasil.

**2007 – 2007**

Mini- curso de Cosméticos (Carga horária 4h)  
Universidade de São Paulo. Ribeirão Preto, SP, Brasil.

**2006 – 2006**

Extensão universitária em Identificação Forense de DNA Humano. (Carga horária: 8h).

**4. Atuação profissional**

**2018 – 2018:** Estagiária na escola municipal NEJA Irmã Edith em Araraquara, SP, atuando no ensino de Ciências para o Ensino Fundamental. Vigência: 05/02/18 a 05/06/2018 (Carga horária: 15h semanais)

**2017 – 2017:** Estagiária na escola estadual Doutor João Pires de Camargo em Araraquara, SP, atuando no ensino de Química para o Ensino Médio. Vigência: 20/08/2017 a 21/11/2017. (Carga horária: 15h semanais).

**2013- 2014:** Bolsista TT3 Fapesp na Faculdade de Odontologia de Ribeirão Preto/USP Processo 2013/15284-4. Vigência: 01/09/2013 a 30/04/2014. (Carga horária: 40 h semanais).

## 5. Projetos de pesquisa

**2014 – atual:** Desenvolvimento de método para previsão de parâmetros físico-químicos do óleo diesel comercial por cromatografia gasosa ultrarrápida - CG-UR e quimiometria.

**Descrição:** Este trabalho tem como objetivo desenvolver modelos quimiométricos com capacidade de estimar os seguintes parâmetros do óleo diesel: teor de biodiesel; massa específica; pontos de ebulição em 10%, 50%, 85% e 90% do volume recuperado; enxofre total e ponto de fulgor através de análises por cromatografia gasosa ultra-rápida visando a obtenção de uma metodologia alternativa confiável, viável, rápida, econômica e ambientalmente mais limpa para o monitoramento destes parâmetros.

**2013 – 2014:** Detecção da enzima 5-lipoxigenase e de osteoclastos na lesão periapical por meio de imunistoquímica e histoenzimologia.

**Descrição:** O objetivo desta pesquisa será detectar a presença da enzima 5-lipoxigenase e de osteoclastos durante o desenvolvimento e progressão da lesão periapical experimental, em camundongos C57BL/6. Serão utilizados 58 camundongos C57BL/6, nos quais serão induzidas lesões periapicais. Os animais serão divididos nos períodos experimentais de 7, 14, 21 e 28 dias após contaminação dos canais radiculares (n= 6 por período) e subdivididos em 3 grupos experimentais de acordo como o tratamento realizado. No Grupo 1 os animais serão tratados com um medicamento que previne a ativação da enzima 5-lipoxigenase (MK 886, Merck, Darmstadt, Germany), no Grupo 2 serão induzidas lesões periapicais e os animais não receberão o medicamento e no Grupo 3 serão utilizados dentes hígidos como controle (sem lesão periapical). Os tecidos periapicais obtidos serão avaliados em microscopia de luz, imunomarcados para a proteínas mediadoras da conversão do ácido araquidônico e a presença e atividade de osteoclastos serão determinadas pela contagem do número de osteoclastos por milímetro da extensão de tecido ósseo em reabsorção. A porcentagem de células positivamente marcadas será determinada e a análise estatística será realizada de acordo com a distribuição dos dados por meio do teste ANOVA ou Kruskal-Wallis ( $\pm = 0,05$ ).

Bolsista Fundação de Amparo – Treinamento Técnico.

Supervisor: Prof. Dr. Francisco Wanderley Garcia de Paula e Silva

**2011 – 2013:** Avaliação de fungos e complexos de salen na obtenção do metabólito quiral e ativo terbutalina.

**Descrição:** O projeto tem como objetivo avaliar a capacidade de fungos e catalisadores biomiméticos em bio(transformar) estereosseletivamente o fármaco bambuterol em seu metabólito quiral e ativo, a terbutalina. Como objetivos secundários para este projeto, têm-se o desenvolvimento de um método analítico estereosseletivo para a análise do fármaco bambuterol e de seu metabólito ativo terbutalina.

**2008 – 2009:** Caracterização forense de SNPs do gene SLC45A2 e sua relação com a pigmentação da pele e anexos em amostra da população brasileira.

**Descrição:** Este trabalho tem como objetivo o levantamento de frequências alélicas e genotípicas dos SNPs L374F, E272K e rs3756462 (T>C) do gene SLC45A2 com o intuito de verificar associações destes SNPs com a coloração da pele e anexos na população brasileira para previsões forenses de cena de crime.

## **6. Idiomas**

### **Inglês**

Compreende Bem, Fala Razoavelmente, Lê Bem, Escreve Razoavelmente.

### **Espanhol**

Compreende Razoavelmente, Fala Pouco, Lê Bem, Escreve Pouco.

## **7. Produção Bibliográfica**

### **Artigos completos publicados**

1. ZANÃO, L.R.; DOS SANTOS, B.C.D.B.; SEQUINEL, R.; FLUMIGNAN, D.L.; DE OLIVEIRA, J.E. Prediction of relative density, distillation temperatures, flash point, and cetane number of S500 diesel oil using multivariate calibration of gas chromatographic profiles. *Energy & Fuels*. 32(8): 8108-8114, 2018.

2. FRACASSO, N.C.A.; ANDRADE, E.S.; WIEZEL, C.E.V.; ANDRADE, C.C.F.; ZANÃO, L.R.; SILVA, M.S.; MARANO, L.A.; DONADI, E.A.; CASTELLI, E.C.; SIMÕES, A.L.; MENDES-JUNIOR, C.T. Haplotypes from the SLC45A2 gene are associated with the presence of freckles and eye, hair and skin pigmentation in Brazil. *Legal Medicine*. 25:43-51, 2017.

3. FRACASSO, N.C.A. ; ANDRADE, E.S. ; ANDRADE, C.C.F. ; ZANÃO, L.R. ; SILVA, M.S. ; MARANO, L.A. ; WIEZEL, C.E.V. ; DONADI, E.A. ; SIMÕES, A.L.; MENDES-JUNIOR, C.T. Association of SNPs from the SLC45A2 gene with human pigmentation traits in Brazil. *Forensic Science International. Genetics Supplement Series (Print)*, 4:342-343, 2013.

### **Resumos publicados em anais de congressos**

FRACASSO, N. C. A. ; ANDRADE, E. S. ; ANDRADE, C. C. F. ; ZANÃO, L. R. ; SILVA, M. S. ; MARANO, L.A. ; WIEZEL, C. E. V. ; DONADI, E.A. ; SIMOES, A. L. ; MENDES-JUNIOR, C.T. . Forensic applicability and association of haplotypes from the SLC45A2 gene with human pigmentation traits in Brazil. In: IV Congresso Brasileiro de Genética Forense, 2013, São Paulo. Saúde, Ética & Justiça. São Paulo, 2013.

ZANÃO, L.R.; FORTES, S.S.; DE OLIVEIRA, A.R.M. Otimização da microextração líquido líquido dispersiva na extração de analitos altamente polares em meio de cultura. Congresso Latino Americano de Cromatografia (COLACRO) XIV, 2012, Florianópolis. Livro de Resumos – COLACRO XIV. v. único. p. 31-31.

MENDES-JUNIOR, C.T. ; FRACASSO, N. C. A. ; ANDRADE, E. S. ; ANDRADE, C. C. F. ; ZANÃO, L. R. ; SILVA, M. S. ; MARANO, L. A. ; WIEZEL, C. E. V. ; DONADI, E. A. ; SIMOES, A. L. . Haplotypes from the SLC45A2 gene are associated with the presence of freckles and eye, hair and skin pigmentation in Brazil. In: 12th International Congress of Human Genetics / 61st Annual Meeting of The American Society of Human Genetics, 2011, Montreal. Program of the 12th International Congress of Human Genetics. Bethesda: The American Society of Human Genetics, 2011. p. 214-214.

VELHO, J. A. ; BRUNI, A. T. ; SANTOS, A. R. M. ; FERNANDES, A. V. M. ; POSSATO, B. ; FURLAN, D. M. ; DOMINGUES, F. P. ; DANIEL, G. R. ; ALCANTARA, G. K. S. ; SANTOS, I. F. S. ; MESSIAS, J. M. ; MARQUES, L. P. ; ZANAO, L. R. ; AGOSTINHO, T. C. ; ANDRADE, Y. A. P. . Análise de distância de disparos por residuografia em anteparos: o que o CSI não conta. In: II ENQFor- Encontro Nacional de Química Forense, 2010, Ribeirão Preto. Livro de Resumos- II ENQFor. Ribeirão Preto: DQ-FFCLRP/USP, 2010. v. único. p. 8-8.

VELHO, J. A. ; BRUNI, A. T. ; MARQUES, L. P. ; POSSATO, B. ; SANTOS, A. R. M. ; ZANAO, L. R. ; SANTOS, I. F. S. ; DANIEL, G. R. ; FURLAN, D. M. ; ALCANTARA, G. K. S. ; ANDRADE, Y. A. P. ; FERNANDES, A. V. M. ; DOMINGUES, F. P. . O ensino grafodocumentoscópico para alunos do curso Química Forense. In: II ENQFor- Encontro Nacional de Química Forense, 2010, Ribeirão Preto. Livro de Resumos- II ENQFor. Ribeirão Preto: DQ-FFCLRP/USP, 2010. v. ÚNICO. p. 9-9.

BRUNI, A. T.; POSSATO, B.; ZANAO, L. R.; OLIVEIRA, A. M. Análise de autos findos de homicídio: prática didática de Assistência Técnica dentro do Processo Penal. In: II ENQFor- Encontro Nacional de Química Forense, 2010, Ribeirão Preto. Livro de Resumos- II ENQFor. Reibeirão Preto: DQ-FFCLRP/USP, 2010. v. ÚNICO. p. 7-7.

MENDES-JUNIOR, C.T. ; FRACASSO, N. C. A. ; ZANÃO, L. R. ; ANDRADE, C. C. F.; SILVA, M. S. ; MARANO, L. A. ; WIEZEL, C. E. V. ; DONADI, E. A. ; SIMOES, A. L. . Alleles from SNPs at the SLC24A5 and SLC45A2 genes are associated with the presence of freckles and eye, hair and skin pigmentation in Brazil. In: 60th Annual Meeting of The American Society of Human Genetics, 2010, Washington. Program of the 60th Annual Meeting of The American Society of Human Genetics, 2010. p. 169-169.

ANDRADE, E. S.; FRACASSO, N. C. A.; ZANÃO, L. R.; MARANO, L. A.; SIMOES, A. L.; MENDES-JUNIOR, C.T. Caracterização forense de SNPs do gene SLC45A2 e sua relação com a pigmentação humana (APRESENTAÇÃO ORAL). In: III Seminário Nacional de DNA e Laboratórios Forenses, 2010, Brasília. Programação do Evento, 2010. p. 3-3.

### **Apresentações de trabalho**

1. ZANÃO, L.R. FORTES, S.S. OLIVEIRA, A.R.M. Otimização da microextração líquido líquido dispersiva na extração de analitos altamente polares em meio de cultura. 2012. (Apresentação de Trabalho/Congresso).



2. HABENSCHUS, M. D.; ZANÃO, L. R.; BOLZON, L. B.; DORES, M. A.; de Oliveira, A. R. M. Avaliação de Complexos de Salen na catálise enantiosseletiva do bambuterol e análise por HPLC. 2012. (Apresentação de Trabalho/Simpósio).
3. MENDES JUNIOR, C. T. ; FRACASSO, N. C. A. ; ANDRADE, E. S. ; ANDRADE, C. C. F. ; ZANÃO, L. R. ; SILVA, M. S. ; MARANO, L. A. ; WIEZEL, C. E. V. ; DONADI, E. A. ; SIMOES, A. L. . Haplotypes from the SLC45A2 gene are associated with the presence of freckles and eye, hair and skin pigmentation in Brazil. 2011. (Apresentação de Trabalho/Congresso).
4. ANDRADE, E. S.; ZANÃO, L. R.; WIEZEL, C. E. V.; SIMOES, A. L.; MENDES JUNIOR, C. T. Avaliação do valor preditivo de ancestralidade e de características relacionadas à pigmentação humana do SNP rs26722 (E272K) do gene SLC45A2 (APRESENTAÇÃO ORAL). 2011. (Apresentação de Trabalho/Congresso).
5. ZANÃO, L. R.; WIEZEL, C. E. V.; SIMOES, A. L.; MENDES JUNIOR, C. T. Caracterização forense de SNPs do gene SLC45A2 e sua relação com pigmentação da pele e anexos em amostras da população brasileira. 2010. (Apresentação de Trabalho/Simpósio).
6. ANDRADE, E. S. ; FRACASSO, N. C. A. ; ZANÃO, L. R. ; MARANO, L. A. ; SIMOES, A. L. ; MENDES JUNIOR, C. T. . Caracterização forense de SNPs do gene SLC45A2 e sua relação com a pigmentação humana (APRESENTAÇÃO ORAL). 2010. (Apresentação de Trabalho/Congresso).
7. MENDES JUNIOR, C. T. ; FRACASSO, N. C. A. ; ZANÃO, L. R. ; ANDRADE, E. S. ; SILVA, M. S. ; MARANO, L. A. ; WIEZEL, C. E. V. ; DONADI, E. A. ; SIMOES, A. L. . Alleles from SNPs at the SLC24A5 and SLC45A2 genes are associated with the presence of freckles and eye, hair and skin pigmentation in Brazil. In: 60th Annual Meeting of The American Society of Human Genetics. 2010. (Apresentação de Trabalho/Congresso).
8. VELHO, J. A. ; BRUNI, A. T. ; SANTOS, A. R. M. ; FERNANDES, A. V. M. ; POSSATO, B. ; FURLAN, D. M. ; DOMINGUES, F. P. ; DANIEL, G. R. ; ALCANTARA, G. K. S. ; SANTOS, I. F. S. ; MESSIAS, J. M. ; MARQUES, L. P. ; ZANÃO, L. R. ; AGOSTINHO, T. C. ; ANDRADE, Y. A. P. . Análise de distância de disparos por residuografia em anteparos: o que o CSI não conta. 2010. (Apresentação de Trabalho/Congresso).
9. VELHO, J. A. ; BRUNI, A. T. ; MARQUES, L. P. ; POSSATO, B. ; SANTOS, A. R. M. ; ZANÃO, L. R. ; SANTOS, I. F. S. ; DANIEL, G. R. ; FURLAN, D. M. ; ALCANTARA, G. K. S. ; ANDRADE, Y. A. P. ; FERNANDES, A. V. M. ; DOMINGUES, F. P. . O ensino grafodocumentoscópico para alunos do curso Química Forense. 2010. (Apresentação de Trabalho/Congresso).
10. BRUNI, A. T.; POSSATO, B.; ZANÃO, L. R.; OLIVEIRA, A. M. . Análise de autos findos de homicídio: prática didática de Assistência Técnica dentro do Processo Penal. 2010. (Apresentação de Trabalho/Congresso).

## **8. Participação em eventos/congressos/simpósios/oficinas/workshops**

1. Empreendedor Business Day e BPM Day Ribeirão Preto. 2016.
2. Seminário Técnico de Cromatografia HPLC/GC/preparação de amostras. 2015.

3. I Workshop em Avanços Recentes no Preparo de Amostras/FCFRP-USP. 2012.
4. XIV Congresso Latino-Americano de Cromatografia e Técnicas Relacionadas 2012.
5. 20º Simpósio Internacional de Iniciação Científica da USP (SIICUSP). 2012
6. 18º Simpósio Internacional de Iniciação Científica da USP (SIICUSP). 2010
7. II Encontro Nacional de Química Forense. 2010
8. I Simpósio Forense/Centro Universitário Barão de Mauá. 2009
9. I Encontro Nacional de Química Forense. 2008
10. Dia do Químico/FFCLRP-USP. 2008
11. XXXII Semana da Química/FFCLRP-USP. 2008
12. Dia do Químico/FFCLRP-USP. 2007

## **9. Supervisão Científica**

Maísa Daniela Habenschus - Avaliação de complexos de salen na catálise enantiosseletiva do bambuterol e análise por HPLC. 2011

Dedico este trabalho à minha família, amigos, ao meu querido Eurípedes e a quem possa ser útil na contribuição com a Ciência.

## AGRADECIMENTOS

Primeiramente, a Deus pela minha vida e as oportunidades em meu caminho e ao meu mentor e aos meus guias espirituais por todo o amparo, intuição e proteção nessa jornada.

À toda minha família, especialmente aos meus pais, por todo o amor, apoio, educação, suporte, por acreditarem nos meus sonhos e me incentivarem a realizá-los.

Ao meu querido namorado Eurípedes por sempre estar ao meu lado, me apoiar, ser paciente, compreensivo, por me fazer a cada dia ser uma pessoa melhor, e, claro, por me alimentar com suas deliciosas refeições.

Aos orientadores e co-orientador que convivi e aprendi nesses anos: Prof. Dr. José Eduardo de Oliveira, Prof. Dr. Rodrigo Sequinel e Prof. Dr. Danilo Luiz Flumignan, pela receptividade, pela oportunidade, pelos ensinamentos e por todo suporte acadêmico e psicológico.

Às colaborações preciosíssimas do Dr. Juliano Passaretti Filho, do auxílio indispensável do Msc Luís Augusto da Silva e do Prof. Dr. Leonardo Pezza por ceder seu laboratório e o *software Pirouette*.

Aos colegas do grupo de pesquisa: Antonio, Camila, Gabrielle, Larissa, Letícia, Maurílio e Weslei pela agradável convivência e por tornarem as dificuldades menos dolorosas e os momentos compartilhados mais felizes.

Aos funcionários e estagiários que passaram pelo CEMPEQC durante a realização deste trabalho: por todo o auxílio, convivência, café e conversas agradáveis.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES – pela bolsa concedida.

“O período de maior ganho em conhecimento e experiência é o período mais difícil da vida de alguém.” (Dalai Lama)

## RESUMO

O óleo diesel é o derivado de petróleo mais consumido no Brasil, principalmente no setor de transporte rodoviário, o qual predomina no país, e, por isso, garantir sua qualidade é imprescindível para o funcionamento adequado do motor e redução de emissões veiculares para a atmosfera. Os métodos oficiais de análise demandam bastante tempo e volume de amostra e, em um contexto onde a preocupação ambiental e econômica prevalecem, é interessante o desenvolvimento de metodologias que otimizem os processos, reduzindo tempo de análise e quantidade de amostra e reagentes. Motivado por essa problematização, este trabalho teve como objetivo o desenvolvimento de um método de cromatografia gasosa ultrarrápida com a máxima quantidade de picos cromatográficos possíveis e resolução global satisfatória para correlacionar as áreas destes picos com um banco de dados formado por resultados de ensaios oficiais dos parâmetros físico-químicos, a fim de criar modelos de calibração multivariada por PLS baseados nestas informações. As amostras utilizadas foram de diesel mistura B8 comercial e os parâmetros avaliados foram a massa específica, teor de enxofre, teor de biodiesel, ponto de fulgor e temperaturas de ebulição para amostras 10% destiladas (T10), 50% destiladas (T50), 85% destiladas (T85) e 90% destiladas (T90). Em seis meses, 637 amostras foram selecionadas de forma representativa por HCA e analisadas através do método cromatográfico otimizado, onde a separação cromatográfica ultrarrápida foi realizada em apenas 6,12 minutos com resolução global de 1,09 e possibilidade de identificação de toda a faixa de hidrocarbonetos presentes no diesel e dos ésteres de ácidos graxos do biodiesel. Este método foi validado e apresentou seletividade; linearidade adequada ( $r > 0,99$ ) sem falta de ajuste e heterocedasticidade; valores adequados segundo o MAPA de precisão intermediária e repetibilidade (menores que 5,0 e 4,8, respectivamente), de erro avaliados intradia e interdias (menores que 4,8 e 4,9, respectivamente) de limites de quantificação (52,62  $\mu\text{g/g}$ ) e detecção (15,95  $\mu\text{g/g}$ ); e robustez. 100 dessas amostras de óleo diesel (18 não conformes e 82 conformes) foram novamente selecionadas por HCA de forma a manter a representatividade do conjunto amostral disponível e facilitar a construção dos modelos PLS. Estes foram construídos utilizando as áreas dos 30 picos principais de cada amostra utilizada como matriz X e os dados dos ensaios normativos como matriz Y. Os melhores resultados foram com a utilização do pré-processamento autoescalado e o tratamento de normalização para as características temperaturas de destilação (T10, T50, T85 e T90); teor de biodiesel e teor de enxofre S10 e S500. Para a massa específica, foi utilizado o pré-processamento centrado na média e o tratamento de normalização. O modelo para o ponto de fulgor foi construído com apenas os 3 primeiros picos cromatográficos e o melhor resultado foi obtido centrando os dados na média e com transformação logarítmica. Em geral, os modelos apresentaram previsão dos parâmetros com boa correlação ( $r_{\text{val}} > 0,6033$ ) e erros de previsão e padronizados um pouco acima, porém similares aos valores de repetibilidade e reprodutibilidade obtidos pelos guias normativos ASTM, EN e NBR e fornecem, com o uso de apenas 0,1  $\mu\text{L}$  de diesel, a obtenção simultânea de 8 parâmetros físico-químicos de 5 amostras por hora, em contrapartida aos métodos oficiais, onde apenas a análise de 8 amostras são possíveis por dia. Isso torna o método atrativo e possível de ser utilizado como método de *screening* em rotinas laboratoriais, reduzindo o número de amostras que necessitam de análise por métodos oficiais e, com isso, diminuindo o tempo e volumes gastos de amostras.

**Palavras-chave:** diesel, cromatografia gasosa ultrarrápida, quimiometria, calibração multivariada.

## ABSTRACT

Diesel fuel is the most consumed petroleum derivative in Brazil, mainly in the road transport sector, which predominates in this country. In this way, guaranteeing its quality is essential for the proper functioning of the engine and the reduction of vehicular emissions to the atmosphere. The official methods of analysis of each characteristic of diesel oil quality require considerable sample time and volume, and in a context where environmental and economic concern prevail, it is interesting to develop methodologies that optimize processes, reducing analysis time, quantity sample and reagents. This work was motivated by this problematization and had the objective of developing an ultrafast gas chromatographic method with a maximum amount of chromatographic peaks and a satisfactory global resolution to correlate the areas of these peaks with a database formed by results of official tests of the physicochemical parameters in order to create multivariate calibration models by PLS based on this information. Commercial diesel B8 blend were used as samples and the parameters evaluated were the specific mass, sulfur content, biodiesel content, flash point and boiling temperatures for samples 10% distilled (T10), 50% distilled (T50), 85 % distilled (T85) and 90% distilled (T90). In six months, 637 samples were representatively selected by HCA and analyzed by the optimized chromatographic method, where the ultrafast chromatographic separation was performed in only 6,12 minutes with an overall resolution of 1.09 and possibility of identification of the entire range of hydrocarbons present in diesel and the fatty acid esters of biodiesel. This method was validated and presented selectivity; linearity ( $r > 0.99$ ) without lack of fit and heteroscedasticity; values in according to MAPA guide of intermediate accuracy and repeatability (less than 5.0 and 4.8, respectively), of intraday and inter-valued error (less than 4.8 and 4.9, respectively) of quantification limits (52, 62  $\mu\text{g/g}$ ) and detection (15.95  $\mu\text{g/g}$ ); and robustness. One hundred of these diesel samples (18 non-conforming and 82 conforming) were again selected by HCA in order to maintain the representativeness of the sample set available and to facilitate the construction of PLS models. These models were constructed using the areas of the 30 main peaks of each sample as matrix X and the data of the normative tests as matrix Y. The best results were with the use of self-scaling pre-processing and standardization treatment for the characteristics of distillation temperatures (T10, T50, T85 and T90); biodiesel content and sulfur content S10 and S500. For the relative density, the centered pre-processing and standardization treatment were used. The model for the flash point was constructed with only the first 3 chromatographic peaks and the best result was obtained by centering the data in the mean and applying logarithmic transformation. In general, the models presented a prediction of the parameters with a good correlation ( $r_{\text{val}} > 0.6033$ ) and predicted and standard errors slightly above, but similar to the values of repeatability and reproducibility obtained by the ASTM, EN and NBR normative guides. So, using only 0.1  $\mu\text{L}$  of diesel, it is possible provide simultaneously 8 physicochemical parameters of 5 samples per hour, in contrast to the official methods, where only the analysis of 8 samples are possible per day. This makes the method attractive and possible for been used as a screening method in laboratory routines, which will reduce the number of samples that need to be analyzed by official methods and will reduce the time of total analysis of all samples.

**Keywords:** diesel, ultrafast gas chromatography, chemometrics, multivariate calibration

## LISTA DE ILUSTRAÇÕES

Figura 1	Principais processos de refino do petróleo para a produção de derivados: Separação, conversão e tratamento	37
Figura 2	Torre de destilação atmosférica com as frações obtidas do petróleo após a separação física (gás combustível, GLP, nafta, querosene, gasóleo atmosférico e resíduo atmosférico)	38
Figura 3	Vendas, pelas distribuidoras, dos derivados combustíveis de petróleo (m <sup>3</sup> ) em 2017	43
Figura 4	Participação de fontes energéticas no consumo total no Brasil em 2017	43
Figura 5	Participação dos combustíveis no setor rodoviário em 2017	44
Figura 6	Não conformidades observadas no óleo diesel em 2017	50
Figura 7	Índice de não conformidade do óleo diesel encontrado no Brasil ao longo dos anos	50
Figura 8	Reação geral de transesterificação	51
Figura 9	Evolução do uso de biodiesel no Brasil.	55
Figura 10	Esquema dos principais componentes de um sistema cromatográfico a gás	62
Figura 11	Representação esquemática de uma coluna empacotada	65
Figura 12	Representação esquemática dos principais tipos de colunas capilares de tubo aberto	66
Figura 13	Dispositivo UFM. (A) Fotografia de um UFM da empresa Thermo Fischer Scientific Inc. ®. (B) Esquema dos componentes que integram o UFM: 1- elemento aquecedor, 2-conluna, 3-fibra cerâmica e 4-sensor de temperatura	68
Figura 14	Fotografia de um disposto UFM instalado em forno de uma cromatografia a gás convencional	69
Figura 15	Representação esquemática de um detector FID	73
Figura 16	Relação entre as disciplinas que definem a quimiometria	85
Figura 17	Representação vetor-linha da amostra x nos eixos das variáveis: 1, 2, e 3	88
Figura 18	Primeira (a) e segunda (b) derivada de um Gaussiano puro	91
Figura 19	Um sinal representado por dois picos sobrepostos (a) aplicado a primeira (b) e segunda (c) derivada.	91
Figura 20	Pontos experimentais para o planejamento composto central utilizando três fatores	98
Figura 21	Superfície de resposta linear ou planar	100
Figura 22	Superfície de resposta quadrática	101
Figura 23	Perfis de superfícies de resposta comumente encontrados: (a) Superfície com ponto de máximo; (b) Superfície com ponto de mínimo; (c) Superfície com ponto de sela	102
Figura 24	Dendograma obtido via HCA composto por 4 grupos a uma similaridade de 0.5 aplicada a análise de 75 amostras	110
Figura 25	Esquema de um processo de calibração convencional que descreve a relação entre as respostas instrumentais (X) e a informação concreta (y). As respostas são transformadas em informação química por meio do modelo matemático definido $y=f(x)$ .	112



Figura 26	Interpretação geométrica de um típico modelo PLS a partir da decomposição simultânea das matrizes de X e Y pelas variáveis latentes (VLs). a) um conjunto de dados composto por dez objetos (amostra) e três variáveis x e y. b) interpretação geométrica da primeira componente VLs no espaço X. c) interpretação geométrica da primeira componente VLs no espaço Y. d) relação interna entre as componentes VLs	117
Figura 27	Representação esquemática do processo de seleção de amostras pelo algoritmo de Kennard-Stone.	122
Figura 28	Resumo das etapas do procedimento experimental realizado neste trabalho	125
Figura 29	Cromatógrafo a gás com módulo <i>ultrafast</i> acoplado	128
Figura 30	Polaridade e composição das fases estacionárias das colunas a) UFM VF23MS; b) UFM PH SE54.	128
Figura 31	Cromatograma obtido sob as condições iniciais de análise, onde a temperatura inicial foi igual a 40 °C, taxa de aquecimento de 100 °C/min até a temperatura final de 250 °C; volume injetado de 0,1 µL e split de 1:100. Vazão de 0,5 mL.min <sup>-1</sup> . A coluna VF23MS foi utilizada como fase estacionária a) mostrando a integração dos picos; b) mostrando as escalas nos eixos.	134
Figura 32	Gráfico de Pareto das variáveis avaliadas no planejamento fatorial completo 2 <sup>5</sup> . Os efeitos lineares ou quadráticos em azul são considerados significativos ( $\alpha = 0,05$ ) para a contribuição com o aumento de picos cromatográficos. Os efeitos em cinza podem ser desprezados.	136
Figura 33	Gráfico normal dos efeitos calculados para o planejamento fatorial completo 2 <sup>5</sup> . Apenas os efeitos representados em vermelho são significativos para o modelo de separação cromatográfica. Os efeitos negativos mostram que estes fatores ou combinações entre fatores influenciam positivamente o modelo quando estão no menor nível e os efeitos positivos indicam que estes fatores e combinações apresentam maior influência quando estão no maior nível.	137
Figura 34	Cromatograma obtido do experimento número 20, onde temperatura final foi mantida no nível máximo e temperatura inicial, taxa de aquecimento e vazão foram mantidos no nível mínimo. A coluna empregada foi a de maior caráter apolar PH SE54.	138
Figura 35	Cromatograma obtido do experimento número 26, onde temperatura final e a vazão foram mantidos no nível máximo e temperatura inicial e taxa de aquecimento foram mantidos no nível mínimo. A coluna empregada foi a de maior caráter apolar PH SE54.	138
Figura 36	Cromatograma obtido do experimento número 12, onde temperatura inicial, temperatura final e taxa de aquecimento foram mantidos no nível máximo e a vazão foi mantido no nível mínimo. A coluna empregada foi a de caráter polar VF23MS	139
Figura 37	Cromatograma obtido do experimento número 15, onde a temperatura final, taxa de aquecimento e vazão foram mantidos	139

	no nível máximo e a temperatura inicial foi mantida no nível mínimo. A coluna empregada foi a de caráter polar VF23MS	
Figura 38	Curva de van Deemter para gases hidrogênio, hélio e nitrogênio, comparando as velocidades lineares ideais.	140
Figura 39	Cromatograma obtido do experimento número 15, onde a temperatura inicial e a temperatura final foram mantidas no ponto central e a taxa de aquecimento foi mantida no nível mínimo axial (-1,68). Coloque aqui a coluna e a vazão do gás de arraste.	142
Figura 40	Gráfico com comparação das respostas experimental e prevista (a) e gráfico de análise de resíduos (b).	147
Figura 41	Superfície de resposta (a) e Gráfico de contorno (b) de Resposta vs TF; Taxa de aquecimento.	148
Figura 42	Superfície de resposta (a) e Gráfico de contorno (b) de Resposta vs TF; TI	149
Figura 43	Superfície de resposta (a) e Gráfico de contorno (b) de Resposta vs TI; Taxa de aquecimento.	150
Figura 44	Curvas de van Deemter (altura do prato vs velocidade linear) ilustrando como o aumento da temperatura auxilia na velocidade linear ótima, permitindo separações mais rápidas	151
Figura 45	Desejabilidade da otimização do método para gerar o maior número de picos. Em vermelho, estão as condições consideradas desejáveis	152
Figura 46	Cromatograma obtido do experimento reproduzindo as condições de desejabilidade, onde a temperatura inicial e a taxa de aquecimento foram mantidas no menor nível (-1,68) e a temperatura final foi mantida no maior nível (1,68). A resolução global é de 1,08	153
Figura 47	Cromatograma obtido do experimento "Asc/Desc 9", onde a temperatura inicial foi igual a 30 °C, a taxa de aquecimento igual a 50 °C/min e temperatura final igual a 240 °C, com tempo total de análise igual a 5,40 minutos	155
Figura 48	Cromatograma obtido do experimento "Refino 7", onde a razão de split é igual a 1:200 e o volume de injeção é igual a 0,1 µL	156
Figura 49	Cromatograma obtido do experimento "Refino 5", onde a razão de split é igual a 1:100 e o volume de injeção é igual a 0,3 µL	157
Figura 50	Cromatograma obtido do experimento Hold Time 5, onde a temperatura inicial, igual a 35 °C, foi mantida por 0,7 minutos. O tempo total desta análise foi igual a 5,80 minutos	158
Figura 51	Cromatograma obtido do experimento "1", onde a temperatura inicial foi igual a 35 °C, mantida por 0,1 minuto e numa taxa de 45 °C/min até 280 °C e então numa taxa de aquecimento igual a 800°C/min até temperatura final igual a 340°C que se mantém por 0,5 minuto, com tempo total de análise igual a 6,12 minutos	160
Figura 52	Replicatas (n = 10) sobrepostas de cromatogramas obtidos do experimento "1", onde a temperatura inicial foi igual a 35 °C, mantida por 0,1 minuto e numa taxa de 45 °C/min até 280 °C e então numa taxa de aquecimento igual a 800 °C/min até temperatura final igual a 340 °C que se mantém por 0,5 minuto, com tempo total de análise igual a 6,12 minutos	160

Figura 53	Comparação dos cromatogramas obtidos nas condições de análise inicial (em vermelho) e final (em preto)	161
Figura 54	Cromatograma obtido de amostra de biodiesel de soja puro utilizando as condições do método cromatográfico otimizado	162
Figura 55	Comparação dos cromatogramas obtidos entre amostra de biodiesel de soja e a amostra de diesel comercial B8 utilizada no desenvolvimento do método de separação	162
Figura 56	Cromatograma obtido ao analisar a série homóloga parafínica (C8-C40) pelo método cromatográfico otimizado	163
Figura 57	Cromatograma obtido em cromatógrafo a gás convencional, onde a análise foi feita utilizando a coluna DB-Petro da J&W Scientific com temperatura inicial igual a 50 °C com taxa de 2 °C/min até 160 °C e taxa de 4 °C/min até 220 °C com taxa de aquecimento de 12 °C/min até 320 °C, mantendo por 30 minutos, totalizando um tempo de análise igual 140 minutos. O detector FID foi mantido a 300 °C	164
Figura 58	Curvas analíticas de cada hidrocarboneto para elucidação do parâmetro linearidade	167
Figura 59	Cromatograma referente à amostra de hidrocarbonetos na concentração de 52,62 µg.g <sup>-1</sup> e hexano como solvente, sendo possível visualizar que o hexano não interfere na seletividade dos hidrocarbonetos estudados	168
Figura 60	Gráfico de Pareto com os fatores estudados, onde TI é temperatura inicial, R1 é taxa de aquecimento 1 e R2 é taxa de aquecimento 2. Nota-se que nenhum efeito ou combinação entre efeitos mostrou ter influência sobre a resposta estudada, ou seja, a área total dos hidrocarbonetos no nível de significância ( $\alpha=0,05$ ).	175
Figura 61	Gráfico normal dos efeitos para os fatores estudados, onde TI é temperatura inicial, R1 é taxa de aquecimento 1 e R2 é taxa de aquecimento 2. Nota-se que nenhum efeito ou combinação entre efeitos mostrou ter influência sobre a resposta estudada, ou seja, a área total dos hidrocarbonetos no nível de significância ( $\alpha=0,05$ ).	175
Figura 62	Dendogramas com os resultados de HCA para seleção das amostras de diesel a cada mês: a) setembro/2017; b) outubro/2017; c) novembro/2017; d) dezembro 2017; e) janeiro/2018; f) fevereiro/2018.	177
Figura 63	Tratamento da linha de base de amostra de diesel a) primeira parte do cromatograma (até 2,0 min) antes da correção; b) primeira parte após correção; c) segunda parte do cromatograma (2,0 a 6,12 min) antes da correção; d) segunda parte após correção; e) cromatograma inteiro antes da correção; f) cromatograma inteiro após correção	179
Figura 64	Dendograma resultante de HCA para seleção de novo conjunto representativo de amostras de óleo diesel	180
Figura 65	Variação dos resultados de todos os parâmetros físico-químicos das amostras selecionadas	181

Figura 66 Regressões lineares obtidas com os dados de valores medidos vs previstos na calibração e validação cruzada nos modelos PLS de cada parâmetro 188

## LISTA DE TABELAS

Tabela 1	Consumo nacional de derivados de petróleo segundo atividade	26
Tabela 2	Composição de hidrocarbonetos presentes no óleo diesel	41
Tabela 3	Especificações do óleo diesel rodoviário brasileiro de acordo com a Resolução ANP nº 50/2013, já adicionadas as modificações ocasionadas pela Resolução ANP nº 69/2014	45
Tabela 4	Principais parâmetros que classificam os principais tipos de cromatografia gasosa de acordo com Magni e colaboradores	61
Tabela 5	Descrição da tabela ANOVA para regressão múltipla	104
Tabela 6	Regiões de amostragem e coleta sob a responsabilidade do CEMPEQC	126
Tabela 7	Normas e equipamentos utilizados para cada parâmetro físico-químico do diesel.	127
Tabela 8	Condições da análise cromatográfica inicial	129
Tabela 9	Níveis do planejamento fatorial completo 2 <sup>5</sup>	130
Tabela 10	Pontos utilizados na construção da curva analítica da etapa de validação, valores das concentrações nominais e reais dos hidrocarbonetos e massas de solução estoque (mix) e solvente utilizadas	131
Tabela 11	Resultados obtidos no planejamento fatorial completo 2 <sup>5</sup>	135
Tabela 12	Níveis do CDD com 3 fatores	141
Tabela 13	Resultados do planejamento CCD. A resposta está representada por picos cromatográficos.	141
Tabela 14	Análise de variância para o método cromatográfico	142
Tabela 15	Coeficientes codificados e efeitos	143
Tabela 16	Análise de variância para o novo modelo calculado para o método cromatográfico	144
Tabela 17	Coeficientes codificados e efeitos para o novo modelo calculado do método cromatográfico.	145
Tabela 18	Valores da resposta experimental e da resposta prevista pelo modelo de regressão	146
Tabela 19	Experimentos de exploração da superfície de resposta (ascendência e descendência)	153
Tabela 20	Variação de volume de injeção e razão de split	155
Tabela 21	Variação do hold time inicial	157
Tabela 22	Experimentos adicionando mais taxas de aquecimento	159
Tabela 23	Comparação dos métodos cromatográficos antes e depois da otimização	161
Tabela 24	Tempo retenção (min) de cada hidrocarboneto da série homóloga parafínica correspondente à faixa do óleo diesel	163
Tabela 25	Equação de regressão da curva analítica e seus parâmetros para cada hidrocarboneto	166
Tabela 26	Teste F para avaliação da seletividade do método	169
Tabela 27	Resultados dos ensaios de repetibilidade, precisão intermediária e exatidão intradia e interdia.	170
Tabela 28	Resultados do ensaio para verificar LQ	171

Tabela 29	Resultados dos experimentos de estabilidade em autoamostrador por 6h e por 24 h para os hidrocarbonetos avaliados	172
Tabela 30	Descrição dos experimentos do planejamento fatorial completo $2^4$ para o ensaio de robustez	173
Tabela 31	Resultados do planejamento fatorial completo $2^4$ para o ensaio de robustez do método desenvolvido	174
Tabela 32	Quantidade de amostras de óleo diesel selecionadas a cada mês	176
Tabela 33	Intervalo de valores de cada parâmetro físico-químico presente nas amostras do conjunto representativo de óleo diesel	182
Tabela 34	Valores dos parâmetros físico-químicos fora da especificação nas amostras selecionadas de óleo diesel	183
Tabela 35	Pré-processamento e transformações dos dados para os modelos de calibração	184
Tabela 36	Características gerais dos modelos de PLS obtidos nas melhores condições de pré-processamento, transformações e seleção de variáveis para cada parâmetro físico-químico do óleo diesel e comparação dos modelos com valores obtidos através dos guias de referência.	187
Tabela 37	Comparação dos modelos PLS obtidos por CG-UR e por CG	191

## LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
ANOVA	Análise de Variância
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustível
ANVISA	Agência Nacional de Vigilância Sanitária
API	American Petroleum Institute
ASTM	American Society for Testing and Materials
BX	Mistura diesel e biodiesel
B8	8% de biodiesel misturado ao diesel, em volume
CEMPEQC	Centro de Monitoramento e Pesquisa da Qualidade de Combustíveis, Biocombustíveis, Petróleo e Derivados
CCD	Planejamento Composto Central ( <i>Central Composite Design</i> )
CG	Cromatografia Gasosa
CG-MS	Cromatografia gasosa acoplada a detector por espectrometria de massas
CG-UR	Cromatografia Gasosa Ultrarrápida
CNPE	Conselho Nacional de Pesquisa Energética
CV	Coeficiente de variação
EN	Norma Europeia ( <i>European Standard</i> )
EPE	Empresa de Pesquisa Energética
FAME	Éster Metílico de Ácido Graxo ( <i>Fatty Acid Methyl Ester</i> )
FID	Detector por Ionização de Chama ( <i>Flame Ionization Detector</i> )
GLP	Gás Liquefeito do Petróleo
INMETRO	Instituto Nacional de Metrologia, Qualidade e Tecnologia
ISO	International Organization for Standardization
HCA	Análise Hierárquica de Agrupamento ( <i>Hierarchical Cluster Analysis</i> )
HDS	Hidrodesulfurização
HDT	Hidrotratamento
LD	Limite de Detecção
LQ	Limite de Quantificação
m/m	Massa/Massa
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
ME	Massa Específica
MME	Ministério de Minas e Energia
MRC	Material de Referência Certificado
MSR	Metodologia de Superfície de Resposta
NBR	Norma Brasileira
OGR	Óleos e Gorduras Residuais
PCA	Análise de Componentes Principais ( <i>Principal Components Analysis</i> )
PET	Politereftalato de etileno ( <i>Polyethylene terephthalate</i> )
PLOT	Porous-Layer Open Tubular
PLS	Mínimos Quadrados Parciais ( <i>Partial Least Squares</i> )
PMQC	Programa de Monitoramento de Combustíveis Líquidos
PNPB	Programa Nacional de Produção e Uso do Biodiesel
r	coeficiente de correlação
R <sup>2</sup>	coeficiente de determinação

Rs	Resolução
RSD	Erro padrão relativo ( <i>Relative Deviation Standard</i> )
RMSEC	Erro médio de calibração ( <i>Root Mean Square Error of Calibration</i> )
RMSEP	Erro médio de previsão ( <i>Root Mean Square Error of Prediction</i> )
RMSEV	Erro padrão de validação cruzada ( <i>Root Mean Square Error of Cross Validation</i> )
SCOT	<i>Support Coated Open Tubular</i>
T10	Temperatura de ebulição quando 10% da amostra foi destilada
T50	Temperatura de ebulição quando 50% da amostra foi destilada
T85	Temperatura de ebulição quando 85% da amostra foi destilada
T90	Temperatura de ebulição quando 90% da amostra foi destilada
TI	temperatura inicial
TF	temperatura final
t <sub>r</sub>	tempo de retenção
QAV	Querosene de Aviação
UFM	Módulo Ultrarrápido ( <i>Ultrafast Module</i> )
USP	<i>United States Pharmacopeia</i>
VL	Variáveis Latentes
v/v	Volume/Volume
WCOT	Wall Bonded Open Tube



## SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO</b>	26
<b>2.</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	32
<b>2.1</b>	<b>Petróleo</b>	32
2.1.2	Química do petróleo	34
2.1.3	Refino do petróleo	36
2.1.4	Óleo diesel	40
2.1.4.1	A importância do óleo diesel no Brasil	42
2.1.4.2	Especificações do óleo diesel rodoviário brasileiro	44
2.1.4.3	Qualidade do óleo diesel	46
2.1.4.4	Não conformidades no óleo diesel	49
2.1.5	Biodiesel	51
2.1.5.1	O biodiesel na matriz energética brasileira	52
<b>2.2</b>	<b>Cromatografia</b>	55
2.2.1	Um breve histórico	55
2.2.2	Cromatografia gasosa: aspectos gerais, parâmetros cromatográficos e instrumentação	57
2.2.2.1	Modalidades e definições de cromatografia gasosa rápida	59
2.2.2.2	Instrumentação	61
2.2.3	Validação	75
<b>2.3</b>	<b>Quimiometria</b>	83
2.3.1	Processamento prévio dos dados: organização dos dados	87
2.3.1.1	Processamento prévio dos dados: transformação	89
2.3.1.2	Processamento prévio dos dados: pré-processamento	92
2.3.2	Planejamento de experimentos	94
2.3.2.1	Planejamento fatorial	95
2.3.2.2	Planejamento composto central	97
2.3.2.3	Metodologia da superfície de resposta (MSR)	98
2.3.2.4	Análise de Variância ( <i>Analysis of Variance</i> – ANOVA)	103
2.3.3	Análise multivariada	104
2.3.3.1	Métodos de reconhecimento de padrões	107
2.3.3.1.1	Métodos não supervisionados	107
2.3.3.1.1.1	Análise hierárquica de agrupamentos (HCA)	108
2.3.3.2	Métodos de calibração multivariada	111
2.3.3.2.1	Regressão por mínimos quadrados parciais (PLS)	114
2.3.3.2.2	Algoritmo Kennard-Stones (K-S)	120
<b>3</b>	<b>OBJETIVOS</b>	124
<b>3.1</b>	<b>Objetivo geral</b>	124
<b>3.2</b>	<b>Objetivos específicos</b>	124
<b>4.</b>	<b>MATERIAIS E MÉTODOS</b>	125
<b>4.1</b>	<b>Coleta e armazenamento de amostras de óleo diesel</b>	126
<b>4.2</b>	<b>Análise dos parâmetros físico-químicos</b>	127
<b>4.3</b>	<b>Desenvolvimento de método de separação cromatográfica para o óleo diesel</b>	127
<b>4.4</b>	<b>Validação analítica do método cromatográfico desenvolvido e otimizado</b>	130
<b>4.5</b>	<b>Seleção de amostras de diesel por HCA</b>	132
<b>4.6</b>	<b>Análise cromatográfica das amostras selecionadas</b>	132

4.7	<b>Desenvolvimento de modelo de calibração multivariada pelo método PLS</b>	133
5	<b>RESULTADOS E DISCUSSÃO</b>	134
5.1	<b>Desenvolvimento e otimização do método cromatográfico</b>	134
5.1.2	Análise da amostra de diesel comercial por cromatografia gasosa convencional	164
5.2	<b>Validação analítica do método cromatográfico</b>	165
5.2.1	Linearidade	165
5.2.2	Seletividade	168
5.2.3	Precisão e exatidão	169
5.2.4	Limite de quantificação	171
5.2.5	Limite de detecção	172
5.2.6	Estabilidade no autoamostrador	172
5.2.7	Robustez	173
5.3	<b>Seleção de amostras de óleo diesel por HCA</b>	175
5.4	<b>Desenvolvimento de modelos de calibração multivariada</b>	177
5.4.1	Ajustes nos picos cromatográficos	177
5.4.1.1	Nova seleção de amostras de óleo diesel por HCA	179
5.4.2	Desenvolvimento dos modelos PLS	183
6	<b>CONCLUSÕES</b>	193
	<b>REFERÊNCIAS</b>	195
	<b>APÊNDICE A</b>	210
	<b>APÊNDICE B</b>	214
	<b>ANEXO A</b>	218

## 1. INTRODUÇÃO

Nos dias atuais, ainda há forte dependência de petróleo e seus derivados para atendimento à demanda energética mundial (EPE, 2018), onde sua participação é superior a 30% do consumo energético total, além da tendência a permanecer como a principal fonte energética mundial nos próximos anos (EPE, 2018). Logo, o argumento de alguns autores (MELO; PAULILLO; VIAN, 2007) de que haverá redução da participação do petróleo na matriz energética mundial devido a investimentos em outras fontes de energia soa um pouco tendencioso, de acordo com as projeções.

O cenário nacional brasileiro não destoa tanto assim do contexto mundial, embora exista uma maior tendência da participação das fontes renováveis na matriz energética brasileira, sobretudo, impulsionadas pelo Programa Nacional do Álcool (PROÁLCOOL) de 1970 (COSTA; PEREIRA; ARANDA, 2010). Neste sentido, os derivados de petróleo (óleo diesel, gasolina, querosene, GLP, óleo combustível e gás natural) também representam maior parcela na matriz nacional, correspondendo a 43% do consumo energético em 2017, sendo a participação desses derivados no consumo energético brasileiro acima da média mundial (EPE, 2018). O setor de maior demanda por derivados de petróleo, de longe, é o de transportes, com 65% de consumo, dos quais 92% desses, representam o transporte rodoviário, como pode ser verificado na Tabela 1.

**Tabela 1** - Consumo nacional de derivados de petróleo segundo atividade.

<b>Setores</b>	<b>Consumo (%)</b>	
Transportes	65	92 (rodoviário)
		5 (aéreo)
		2 (aquaviário)
		1 (hidroviário)
Uso não energético	13	
Industrial	10	
Residencial	6	
Agropecuário	5	
Outros	1	

**Fonte:** EPE (2018).

De acordo com o relatório anual da Empresa de Pesquisa Energética (EPE) referente ao ano de 2017, o principal derivado do petróleo que subsidia em grande parte a demanda energética brasileira é o óleo diesel, sobretudo, utilizado como

combustível no setor de transportes. Desta maneira, o crescimento da economia brasileira reflete diretamente na expansão da frota de veículos utilizados no transporte de *commodities*, bem como o aumento de veículos para o transporte de passageiros, o que resulta no aumento das emissões atmosféricas de compostos potencialmente tóxicos ao ambiente e ao ser humano.

Além das emissões de poluentes na atmosfera, um outro problema recorrente, não apenas no que se diz respeito ao óleo diesel, mas aos combustíveis de modo geral, são as adulterações. A prática da adulteração visa o lucro a partir da fraude na composição dos combustíveis. Desta forma, o diesel pode ser adulterado por meio da adição de álcool, solventes, óleos vegetais e água (CNT, 2012). A adulteração de combustíveis é um problema tanto para os governos que perdem muito dinheiro por meio da fiscalização quanto para os consumidores diante dos possíveis impactos negativos que um combustível adulterado ocasiona no desempenho do motor (VEMPATAPU; KANAUIA, 2017).

Uma das formas mais comuns de adulteração do diesel consiste na adição de um combustível doméstico de baixo custo como o querosene. O querosene tem um forte potencial para ser usado como adulterante no diesel devido à sobreposição de hidrocarbonetos ( $C_9$  e  $C_{19}$ ) no diesel pelos hidrocarbonetos ( $C_6$  e  $C_{16}$ ) do querosene (CUNHA et al, 2016). Além disso, biodiesel e óleos residuais (óleo lubrificante automotivo, óleo de fritura etc.) também foram relatados como adulterantes de diesel (CORGOZINHO; PASA; BARBEIRA, 2008; BRANDAO; BRAGAB; SUAREZ, 2012)

Na tentativa de solucionar os problemas de adulteração, de um modo geral, as distribuidoras de petróleo contam com programas de acompanhamento e controle dos produtos oferecidos nas redes credenciadas (CNT, 2012). Contudo, há necessidade de um aporte maior de investimentos em pesquisa com intuito de desenvolver métodos de análise mais rápida e fácil de combustíveis. O desenvolvimento de um método rápido e sensível para detecção e determinação rápida da concentração de biodiesel, por exemplo, é valioso para uma variedade de campos onde o biodiesel é misturado ao diesel tradicional (HUPP et al., 2018).

No entanto, a complexidade na composição de combustíveis derivados de petróleo, bem como de seus potenciais adulterantes, leva a um cenário desafiador no qual a imposição e a conformidade das normas padrão muitas vezes não são fáceis. Métodos padrão amplamente aceitos e adotados incluem as normas ISO (*International Organization for Standardization*), EN (*European Normalization*), ASTM (*American*

*Society for Testing and Materials*) e NBR (Normas Brasileiras), que abrangem uma variedade de métodos e procedimentos para verificar se os requisitos especificados relativos à qualidade de produtos de petróleo, incluindo os combustíveis, são atendidos e, assim, se estão em conformidade segundo a norma em questão, por meio da investigação de suas propriedades físico-químicas (PAULS, 2001a; HUPP *et al.*, 2018).

Para o óleo diesel, os critérios principais relacionados ao controle da qualidade são estabelecidos pela ANP. Atualmente, a Resolução ANP n° 50, de 2013 especifica as características desejáveis para este combustível e os métodos de análise, através de metodologias normativas como ASTM, EN e ABNT NBR constituídas principalmente por métodos físico-químicos, cromatográficos e espectroscópicos.

Os métodos físico-químicos utilizam uma maior quantidade de amostras (combustíveis) em comparação com os métodos cromatográficos e espectroscópicos, onde, exceto a diluição, nenhuma outra forma de pré-tratamento da amostra é necessária na maioria dos casos (ARMENTA; GARRIGUES; GUARDIA, 2008).

Tradicionalmente, a cromatografia gasosa (CG) tem sido o método preferido para elaboração de métodos analíticos aplicados em diversas análises de blendas de diesel e biodiesel, pois envolve uma instrumentação analítica relativamente simples, uma vez que uma coluna de caráter apolar e um programa de temperatura bastante simples permitem a identificação de componentes do diesel e, conseqüentemente, avalia a qualidade do processo de refino (YU, WANG; QIAN, 2005; BARMAN *et al.*, 2001; WEBER; WALSH; HARYNUKK, 2016; HUPP *et al.*, 2018; ZANÃO *et al.*, 2018).

O óleo diesel é considerado uma amostra de composição complexa, devido à presença de cadeias carbônicas maiores, menores e intermediárias, correspondentes à fração do petróleo com hidrocarbonetos entre 8 a 28 carbonos (SPEIGHT, 2006). Devido à essa complexidade, a análise de óleo diesel por CG geralmente é demorada e mesmo utilizando horas de análise, a separação entre os picos cromatográficos não se dá por completo. Dos Santos (2012) desenvolveu uma metodologia de detecção por FID com resolução global de 1,79, possibilitando a identificação de 481 picos e com tempo total de análise de 108 minutos para correlacionar com parâmetros da qualidade. Menezes e Cataluña (2008) desenvolveram uma metodologia também de detecção por FID com análise cromatográfica de 175 minutos para identificação de um padrão de hidrocarbonetos parafínicos na faixa do diesel, de maneira qualitativa para ser empregado em análise de emissão de poluentes do motor de ciclo diesel,

sem calcular número de picos ou resolução global. Chereptisa et al. (2003) desenvolveram um método com detecção por FID para determinar parâmetros da qualidade do óleo diesel por CG e obtiveram um tempo de análise de 140 min e, de acordo com reprodução deste método por Dos Santos (2009), resolução global de 2,65 e identificação de 373 picos. Faria et al. (2007) trabalharam em um método de CG-MS visando apenas quantificar o biodiesel em misturas com diesel e este método teve tempo total de análise de 28,7 minutos.

Em geral, análises de blendas de diesel e biodiesel realizadas em sistema CG convencional são geralmente demoradas, com tempos de execução de pelo menos trinta minutos para uma corrida cromatográfica (SHIBAHARA; YAMAMOTO; KINOSHITA, 2008; MARSHAL *et al.*, 2009; HUPP *et al.*, 2018) Por outro lado, métodos baseados na cromatografia gasosa ultrarrápida (CG-UR) utilizam também colunas apolares, mas com menor comprimento (~2 a 10 m) e diâmetro interno em torno de 0,05 a 0,32 mm que permitem análises de diesel e biodiesel em apenas pouco minutos (DAVID *et al.*, 1999; STRIEBICH, 1999; TRANCHIDA *et al.*, 2008; SEQUINEL *et al.*, 2010).

Outro ponto a favor da preferência do uso da CG, bem como a CG-UR em análises de diesel e biodiesel consiste na maior capacidade da técnica em identificar individualmente com maior clareza os vários componentes presentes na composição do combustível dos tais FAMES (*Fatty Acid Methyl Esters*), que consistem em uma mistura de ésteres metílicos de ácidos graxos variando de C<sub>6</sub> a C<sub>22</sub> o número de carbonos dependendo da matéria-prima empregada para produção do biodiesel, uma vez que em CG-UR o tempo de análise é expressivamente reduzido, e, geralmente, sem trazer prejuízos significativos à resolução (MONDELLO *et al.*, 2006; MASOOD; SALEM, 2008; PAULS, 2011b ; GODING *et al.*, 2013; MAZIVILA, 2018; TERCINI *et al.*, 2018).

Nos métodos relativos à separação CG-UR, um imenso volume de dados é gerado, logo, deve ser adequadamente processado por um *software* operacional, e devidamente tratado, pois essas etapas são essenciais para uma análise confiável a qual não se resume em apenas uma imensidão de dados sem correlação alguma. Portanto, os métodos baseados em CG-UR fornecem um nível superior de informação em comparação com outros modos de cromatografia, o que torna o processo de validação embora trabalhoso extremamente importante (MARSHAL *et al.*, 2009; VEMPATAPU; KANAUIA, 2017).

Correlações mais definitivas entre os dados podem ser obtidas usando procedimentos matemáticos que permitem a comparação estatística de cromatogramas, resultando em um aumento da confiança na associação ou distinção dos dados. Assim, métodos estatísticos e quimiométricos podem ser usados para agrupar amostras com base em semelhanças químicas, quantificando estatisticamente a correspondência de padrões (HUPP *et al.*, 2009).

Um dos pilares da quimiometria são os métodos multivariados de análises os quais são geralmente construídos para interpretar a variação química observada de amostra para amostra, na magnitude do sinal do detector nas variáveis correspondentes, por exemplo tempo de retenção, para um conjunto grande de dados (BEEBE; PELL; SEASHOLTZ, 1998; MALINOWSKI, 2002; GEMPERLINE, 2006).

O universo dos métodos multivariados de análises contempla diversos métodos quimiométricos, cada qual com sua limitação e potencial. Dentro deste universo, os métodos de reconhecimento de padrão e calibração multivariada são os mais utilizados em cromatografia. (SOUZA; POPPI, 2012; FERREIRA, 2015).

Marshall e colaboradores utilizaram a técnica de cromatografia gasosa acoplada ao espectrômetro de massas (CG-EM) e métodos quimiométricos baseados na PCA e na correlação por produto momentâneo de Pearson (PPMC) para criar um modelo de classificação de amostras de diesel. Os autores utilizaram um algoritmo de alinhamento do tempo de retenção para alinhar os cromatogramas antes da normalização da área do pico. O PPMC e a PCA foram então empregados para investigar associação e discriminação entre as amostras de diesel. Os autores observaram que os coeficientes PPMC para comparações pareadas do diesel puro foram geralmente superiores a 0,80, indicando uma forte correlação. No entanto, os gráficos de *loadings* do PCA apontaram para a presença de picos derivados indicando um possível desalinhamento do tempo de retenção, cuja magnitude pode indicar a extensão do desalinhamento (MARSHALL *et al.*, 2009).

Hupp e colaboradores realizaram um estudo com amostras de diesel utilizando CG-MS e os métodos quimiométricos PCA e PPMC. O tratamento quimiométrico apontou que os componentes aromáticos têm maior contribuição na variância dos modelos criados, isto é, são mais expressivos para a discriminar as amostras de diesel (HUPP *et al.*, 2009). Cramer *et al.* (2014) usaram dados obtidos por CG-EM e posteriormente tratados por PLS, para preverem muitas propriedades de combustíveis, incluindo viscosidade, densidade e índice de cetano e desenvolveram

um modelo único para um conjunto diversificado de combustíveis que vão do querosene ao diesel marítimo.

Freye e colaboradores investigaram a relação entre a composição química e várias propriedades físico-químicas de combustíveis querosene usando cromatografia gasosa bidimensional (2D) abrangente (GC x GC) acoplada ao detector de ionização por chama (FID) e os métodos PLS e de calibração cruzada para construir dois modelos preditivos de regressão. Os autores observaram que os erros do modelo previsto construído a partir da validação cruzada são ligeiramente menores do que o modelo criado a partir do PLS (FREYE *et al.*, 2016).

Um estudo conduzido por Kehimkar *et al.* (2015) foi publicado para estudar dados cromatográficos multidimensionais obtidos por CG 2D combinada com espectrômetro de massas com analisador de tempo de voo (CG-ATV) usando análise quimiométrica por PLS para modelar e prever dados da curva de destilação avançada de combustíveis utilizados em motores de propulsão. O erro quadrático médio de calibração variou de 0,1 a 0,5 ° C, e foi tipicamente abaixo de  $\pm 0,2$  ° C, para a calibração por PLS da modelagem da curva a partir dos dados cromatográficos, indicando um bom ajuste do modelo.

Hupp *et al.* (2018) utilizaram o método oficial D7758, originalmente desenvolvido para análise de diesel a partir da CG convencional, para desenvolver um método alternativo para analisar blends de diesel e biodiesel a partir de métodos quimiométricos e CG-UR. Os autores agruparam as amostras com base na concentração e no tipo de diesel utilizando a PCA. A concentração de biodiesel foi avaliada com precisão usando o método PLS em um conjunto de treinamento para uma faixa de biodiesel B0-B20, enquanto as previsões foram feitas com algum sucesso em um conjunto de incógnitas comerciais e laboratoriais. O algoritmo kNN foi usado para descrever e prever a concentração e para identificar misturas de biodiesel de diferentes matérias-primas. Por fim, os autores concluíram que a combinação de métodos quimiométricos aliados a CG-UR formam uma técnica eficaz e rápida (1,4 a 2,2 minutos) para a análise da fonte e composição do biodiesel em combustíveis mistos biodiesel-diesel.

Rocha e Sheen (2019) compararam a precisão e a robustez de dois modelos preditivos a partir de dois métodos quimiométricos, PLS e SVM (*Support Vector Machine*), para determinar como as propriedades físico-químicas de vários combustíveis derivados de petróleo dependem da composição. Para tanto, os autores



utilizaram CG-EM para investigar a composição e as normas padrão ASTM para analisar as propriedades físico-químicas. Verificou-se que o SVM é geralmente melhor para prever as propriedades físico-químicas, embora para um conjunto de dados mais abrangente, o desempenho dos modelos PLS possa ser melhorado. No entanto, ambos os modelos fornecem previsões robustas que podem ser usadas para fins regulatórios, econômicos e de segurança.

Por fim, são escassos os trabalhos publicados que reportam, especificamente, sobre a análise multivariada de dados a partir de métodos quimiométricos dentro de um contexto no qual a matriz seja o óleo diesel e a técnica utilizada para extrair informações da composição química das amostras seja a cromatografia gasosa ultrarrápida. A maioria dos trabalhos, dentro de uma perspectiva semelhante, utiliza a gasolina como matriz e as técnicas de espectroscopia no infravermelho com transformada de Fourier, espectroscopia no infravermelho próximo, ressonância magnética nuclear, além de técnicas mais convencionais de cromatografia gasosa, o que dá a este trabalho relevância para ser realizado.

Neste contexto, foi proposto nesse trabalho uma metodologia alternativa que seja capaz de prever alguns parâmetros físico-químicos (massa específica, teor de biodiesel, teor de enxofre, ponto de fulgor e temperaturas de destilação em 10%, 50%, 85% e 90% do volume recuperado) de amostras diesel comerciais a partir de modelos de calibração multivariada construídos pelo método PLS utilizando perfis obtidos por CG-UR.

## **2. REVISÃO BIBLIOGRÁFICA**

### **2.1 Petróleo**

A palavra petróleo tem sua origem no latim a qual significa óleo de pedra (*petra* = pedra; *oleum* = óleo), provavelmente, devido à sua descoberta na superfície através de fendas no solo, vindo de rochas profundas. Sua cor é escura, variando entre o negro e o castanho-claro. É uma substância menos densa que a água, viscosa, inflamável e com cheiro característico (CAMPOS, 2005; THOMAS, 2001). A formação do petróleo depende de condições especiais existentes em alguns lugares da Terra, como fundos de cavernas, mares e rios, onde há pouco ou nenhum oxigênio, inibindo a decomposição de organismos que morreram no local ou que para lá foram

transportados e originaram um lodo, que foi sendo filtrado pelas rochas porosas do fundo dos mares até se transformar no líquido que conhecemos (NEIVA, 1986). Jazidas petrolíferas são rochas sedimentares que contêm um número elevado de poros, onde fica acumulado o petróleo.

As rochas sedimentares são derivadas de detritos de outras já pré-existentes e, devido ao intemperismo, sofrem alterações constantes gerando um material que, quando transportado pela água ou pelo vento, se deposita como sedimento. Através então de uma compactação ou cimentação durante milhões de anos, há modificações para a formação de uma rocha sedimentar (CAMPOS, 2005). As rochas sedimentares são fundamentais para o processo de formação do petróleo, pois uma quantidade grande matéria orgânica de origem vegetal e animal, sedimentadas no fundo de oceanos, foram sendo soterradas durante a constituição dessas rochas. Parte dessa matéria orgânica foi mantida conservada pelas condições amenas de temperatura e pressão, o que originou os chamados fósseis. No entanto, à pressão e temperatura elevadas (quando influenciadas pelo aquecimento provocado pelas camadas mais baixas da Terra), a tendência é que esses restos orgânicos passem pelo processo de litificação, tornando-se líquidos. Assim, conforme as condições de armazenamento, esse material se acumula e se transforma em petróleo (RIBEIRO, 2014).

Historicamente, o petróleo foi usado por muitos séculos na Mesopotâmia, Egito, Pérsia e China para aquecimento, iluminação e construção de estradas (SHELL INTERNATIONAL PETROLEUM COMPANY LIMITED, 1983) e no Novo Mundo, os índios pré-colombianos o utilizavam na impermeabilização dos potes de cerâmica. Em geral, o petróleo aproveitado pelas civilizações antigas era o que não estava preso por rochas impermeáveis e migrava livremente pelo subsolo, alcançando a superfície. A busca crescente na sociedade moderna data de 1859, quando o coronel Drake, na Pensilvânia, descobriu um poço com 21 metros de profundidade que conseguia extrair 2 m<sup>3</sup> de petróleo por dia. Assim, quando descobriram que a destilação do petróleo substituíria com grande margem de lucro o querosene amplamente utilizado na iluminação, antes obtido a partir do carvão vegetal e do óleo de baleia, iniciou-se a exploração comercial e a era do petróleo. Poucos anos depois, com a invenção dos motores à gasolina (1861) e à diesel (1893), os lucros desta atividade extrativa aumentaram exponencialmente (THOMAS, 2001). Atualmente, principalmente a partir da guerra de 1914-18 onde foi utilizado como insumo estratégico, é, de todos os recursos naturais, o mais cobiçado. No Brasil, a história do petróleo se inicia em 1858

com o Marquês de Olinda cedendo o direito de extração a José Barros Pimentel na Bahia com a finalidade da produção de querosene. No entanto, a partir de 1953, no governo Vargas, com o monopólio estatal da Petrobras é que se deu início as pesquisas e desenvolvimento do petróleo brasileiro (THOMAS, 2001).

Nos dias atuais, os combustíveis derivados do petróleo fornecem aproximadamente 60 % da oferta mundial total de energia (BRITISH PETROLEUM, 2018). Gasolina, querosene e óleo diesel fornecem combustível para automóveis, tratores, caminhões, aeronaves e navios. O óleo combustível e o gás natural são usados para aquecer residências e prédios comerciais, bem como gerar eletricidade. Os produtos petrolíferos são os materiais básicos utilizados na produção de fibras sintéticas para vestuário e em plásticos, tintas, fertilizantes, inseticidas, sabões e borracha sintética. Assim, o uso do petróleo como fonte de matéria-prima na fabricação é fundamental para o funcionamento da indústria moderna e, com isso, apesar do apelo ambiental e do desenvolvimento tecnológico voltado aos combustíveis renováveis, a sociedade ainda é fortemente dependente de combustíveis derivados de petróleo (BRITISH PETROLEUM, 2018).

### 2.1.2 Química do petróleo

O petróleo apresenta características variadas conforme as condições geológicas em que foi formado e pode se apresentar nas formas gasosa (gás natural), líquida (óleo cru) ou sólida (xisto). Ele constitui-se principalmente de misturas de hidrocarbonetos (sendo que o carbono representa entre 83% a 86% do total da sua massa e o hidrogênio entre 11% a 13%), mas também contém moléculas orgânicas com heteroátomos (oxigênio, nitrogênio e enxofre) e organometálicas em menores proporções (SPEIGHT, 2006; ANTUNES, 2007). Para a sua classificação, são consideradas três características:

- **Base:** Considera os tipos de hidrocarbonetos predominantes. Eles podem ser de base parafínica, onde predominam os hidrocarbonetos saturados de fórmula  $C_nH_{2n+2}$  e seu resíduo consiste em uma substância cerosa. São leves, fluidos e preferíveis para a produção de querosene de aviação (QAV), óleo diesel e lubrificantes de boa qualidade. Na base naftênica, predominam hidrocarbonetos cíclicos saturados e seu resíduo é asfáltico. Eles apresentam

baixo teor de enxofre e são indicados na produção de gasolina, nafta petroquímica, QAV e lubrificantes. Os de base aromática, onde há hidrocarbonetos cíclicos não saturados, como benzeno e tolueno, são os mais pesados e os que possuem maior teor de enxofre. Sua utilização se dá na produção de derivados que são empregados na indústria petroquímica e na produção de solventes e asfalto (ANTUNES, 2007; MORAIS, 2013). Há também classificações resultantes de combinações entre estas principais (MORAIS, 2013);

- **Densidade:** Utiliza o grau API (*American Petroleum Society* - API), definido por:

$$^{\circ}\text{API} = \frac{141,5}{\rho} - 131,5 \quad (\text{equação 1})$$

onde  $\rho$  é a densidade relativa medida. São denominados leves quando possuem acima de 30° API, médios quando possuem entre 21 a 30° API e pesados quando possuem abaixo de 21° API. Há ainda os ultrapesados que possuem menos de 10° API. Os óleos leves são os mais desejados e prestigiados, pois permitem uma maior produção de gasolina e gás liquefeito do petróleo (GLP), os chamados derivados leves, sem grandes custos e investimentos nas refinarias;

- **Teor de enxofre:** Quando apresenta teor de enxofre menor que 0,5% de sua base, o petróleo é denominado “doce”, e, acima deste valor, é denominado “azedo”. O tipo doce é preferível devido a danos ao meio ambiente e desgaste dos equipamentos nas refinarias provocados pelo enxofre (ANTUNES, 2007).

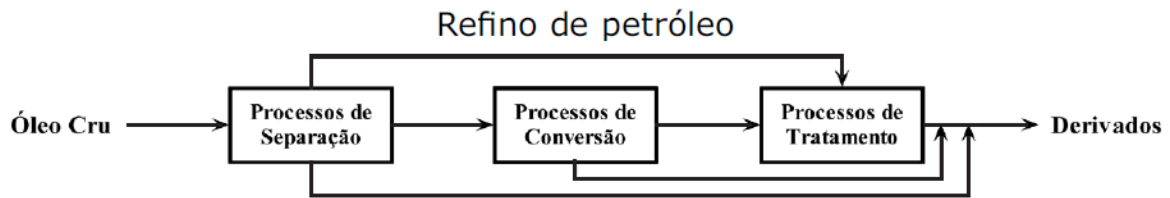
O petróleo brasileiro possui uma média de 24,36 °API e 0,54% em massa de enxofre, sendo o maior e menor valor de densidade, respectivamente, encontrados em Condensado de Merluza na Bacia de Santos no estado de São Paulo (58,30 °API) e Fazenda Belém na Bacia Potiguar no estado do Ceará (12,70 °API). No total de reservas, 45% do petróleo brasileiro é pesado, 34% é intermediário e 6% é leve e 15% das reservas são formadas por gás (MORAIS, 2013). Ao decorrer dos anos, o grau

API dos petróleos brasileiros foi caindo, à medida que crescia a produção da Bacia de Campos, cujos óleos são mais pesados (SZKLO; ULLER, 2008; MORAIS, 2013).

Existe uma tendência no aumento do grau API médio do petróleo brasileiro desde 2006, onde foram descobertas as acumulações na camada geológica do Pré-sal, que são constituídas de petróleos mais leves do que os dos campos do Pós-sal da Bacia de Campos. Quanto ao valor de porcentagem em massa de enxofre, o maior valor é encontrado na corrente Peregrino na Bacia de Campos no estado do Rio de Janeiro (1,80%) e o menor valor é encontrado no Condensado do Mexilhão na Bacia de Santos no estado de São Paulo (0,01%) (MORAIS, 2013). O conhecimento destas propriedades é fundamental para a definição dos processos a serem priorizados no refino, e, conseqüentemente, aumentar a eficiência e maximizar os produtos obtidos de cada tipo de petróleo. Devido às características diferenciarem entre os campos produtores, não existe uma forma única de refino que seja aplicada a todas as refinarias e tampouco todos os derivados podem ser produzidos com qualidade e de forma economicamente viável a partir de qualquer petróleo (MARIANO, 2001).

### 2.1.3 Refino do petróleo

Uma refinaria de petróleo possui como objetivo principal a produção de combustíveis de maior valor econômico e insumos petroquímicos (ABADIE, 2002; CORRÊA, 2003;). De acordo com Antunes (2007), o esquema básico do refino do petróleo se dá de acordo com a Figura 1, onde os três principais processos (separação, conversão e tratamento) são realizados com a finalidade de produzir os derivados.



**Figura 1** - Principais processos de refino do petróleo para a produção de derivados: Separação, conversão e tratamento.

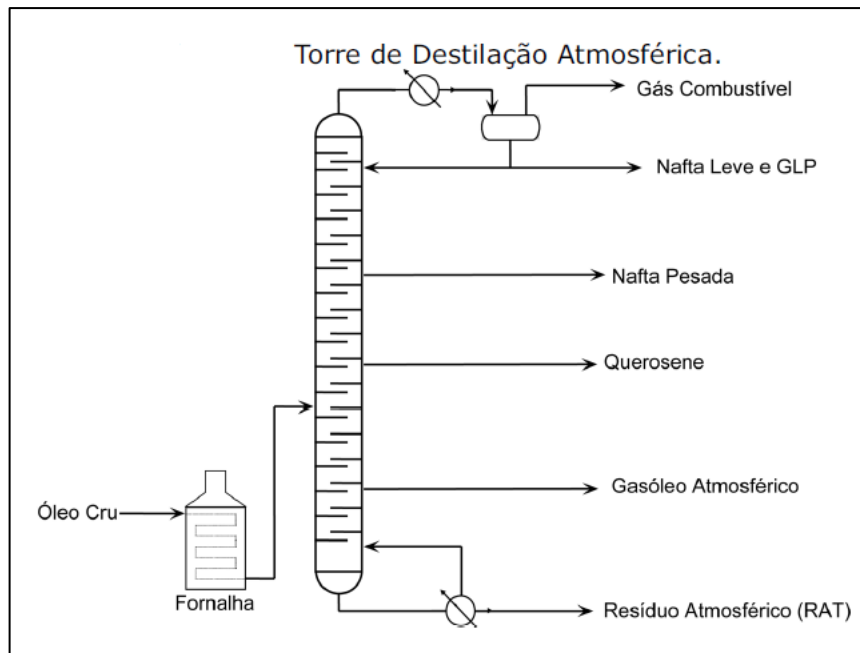
**Fonte:** Antunes (2007).

A separação ou fracionamento, de natureza física, tem por objetivo a separação por diferença de volatilidade em frações com propriedades físicas comuns e distintas entre si do petróleo bruto. A conversão, de natureza química, destina-se a transformar as frações pesadas obtidas na fase de separação, de baixo valor comercial, para a obtenção de derivados leves e médios de alto valor comercial; e, por fim, o tratamento, de natureza química, visa à eliminação de impurezas para a obtenção de derivados com as especificações exigidas para a comercialização (SPEIGHT, 2006; MORAIS, 2013).

Após uma prévia etapa de dessalgação para retirada de sedimentos como sais, areia, metais e traços inorgânicos, há então a separação. A destilação é o processo primordial e fundamental de refino e o único onde o petróleo é a substância de entrada. Ela consiste no fornecimento de calor para gerar vapores e líquidos de composições diferentes entre si e da original. A torre de destilação atmosférica (Figura 2) é onde o petróleo cru é pré-aquecido em um forno a temperaturas entre 300-400 graus Celsius e em seguida, é enviado para uma coluna de destilação vertical, à uma pressão ligeiramente superior à atmosférica, e então é realizada a separação primária dos diversos componentes através de vaporização e condensação, de acordo com as faixas de temperatura de ebulição de cada composto (ANTUNES, 2007).

As frações separadas são coletadas em bandejas ou pratos, dispostos horizontalmente no interior da coluna de destilação. Os vapores que se dirigem para o alto da coluna, onde as temperaturas são menores, condensam-se, permitindo que as frações mais leves sejam coletadas nos pratos internos mais altos e enviadas por tubulações para permutadores de calor. O querosene e o óleo diesel são retirados em pontos laterais da coluna de destilação, de acordo com as temperaturas de destilação dessas e de outras frações. A fração pesada, ou resíduo de fundo (Resíduo

Atmosférico - RAT), de elevado peso molecular, não se vaporiza e normalmente é direcionada para a destilação a vácuo (MORAIS, 2013; ABADIE, 2002; SZKLO e ULLER, 2008).



**Figura 2** - Torre de destilação atmosférica com as frações obtidas do petróleo após a separação física (gás combustível, GLP, nafta, querosene, gasóleo atmosférico e resíduo atmosférico).

**Fonte:** Antunes (2007).

As frações obtidas na unidade de destilação, conhecidas como as sete frações básicas, são (ANTUNES, 2007; DIETEZ, 2014):

- **Gás Combustível:** geralmente é um produto final utilizado na própria refinaria, sendo queimado nas caldeiras e nos fornos. É composto pela faixa de  $C_1$  e  $C_2$ ;
- **-GLP:** pode ser tanto um produto final, sendo armazenado em esferas ou um produto intermediário com destino a unidade de lavagem cáustica. É composto por  $C_3$  e  $C_4$ ;
- **Nafta (Corte 20 a 220 °C):** podem ser produtos finais (nafta petroquímica, gasolina ou solventes) armazenados em tanques ou produtos intermediários que podem passar pelo processo de reforma catalítica a fim de gerar gasolina de melhor qualidade;

- **Querosene (Corte 150 a 300 °C):** pode ser produto final (QAV ou iluminação) ou pode ser intermediário e passar pela unidade de hidrotreamento (HDT), onde pode tanto aumentar a produção de óleo diesel quanto melhorar a viscosidade do óleo combustível;
- **Gasóleo Atmosférico (Corte 100 a 400 °C):** podem já constituírem os produtos finais (óleo diesel armazenado em tanques) ou produtos intermediários, passando por HDT onde resultam produtos com teores reduzidos de impurezas e de compostos indesejáveis (enxofre, oxigênio, nitrogênio e traços de metais) gerando assim óleo diesel com melhor estabilidade e adequado às especificações exigidas pelos mercados de produtos finais;
- **Resíduo Atmosférico:** pode ser produto final (asfalto ou óleo combustível, se for diluído com correntes de menor viscosidade). Como intermediário, pode passar por unidades de tratamento de desasfaltação e gerar solventes.

Então, após o processo de separação por destilação, nos processos de conversão, as frações mais pesadas provenientes da destilação do petróleo, como os gasóleos, através de processo de natureza química e pela ação de pressão e calor, são convertidas em derivados mais leves, tais como naftas, gasolina, óleo diesel e GLP. Nesta conversão, os processos químicos podem alterar a constituição molecular de uma dada fração para a obtenção de moléculas de melhor qualidade, sem, contudo, transformá-la em outra fração. A conversão é realizada por meio dos processos de craqueamento térmico e craqueamento catalítico e, por meio destes, moléculas pesadas dos hidrocarbonetos são quebradas em moléculas menores ou são efetivados rearranjos moleculares.

Posteriormente, os processos de tratamento retiram os elementos químicos causadores de poluição e corrosão, como enxofre, nitrogênio e traços de metais, dos derivados obtidos nas etapas anteriores da refinação. O tratamento promove ainda a separação dos materiais asfálticos (de baixo valor) e corrige o odor dos produtos, adequando-os aos padrões de qualidade para atender às exigências da



comercialização. Os processos de tratamento convencionais, que não operam sob condições severas, são o cáustico simples, o Tratamento Cáustico Regenerativo (TCR) e o Tratamento Bender. Quando há necessidade de condições mais severas para a remoção de impurezas do querosene, do diesel e de frações pesadas são utilizados processos que utilizam o hidrogênio na presença de catalisador, tais como o HDT, que melhoram a qualidade dos derivados tratados.

O processamento do petróleo resulta em aproximadamente de 34% em óleo diesel, 23% de gasolina e 19% de óleo combustível. Este perfil do refino depende das características do petróleo, do parque de refino e da demanda do mercado consumidor de derivados (ANTUNES, 2007). Segundo o Anuário Estatístico da Agência Nacional do Petróleo, Gás Natural e Biocombustível (ANP) de 2018, as refinarias brasileiras em 2017 geraram no total 36,8% de óleo diesel, 25,1% de gasolina e 10,6% de óleo combustível (ANP, 2018a).

#### 2.1.4 Óleo diesel

Conhecido como a fração do petróleo que abrange a faixa de  $C_8$  a  $C_{28}$  (SPEIGHT, 2006), o diesel tem sua produção, no Brasil e no mundo, baseada principalmente no processo de destilação a partir de cortes de gasóleos atmosféricos, seguido de hidrodessulfurização (HDS – tipo de hidrotreatamento para retirada de enxofre). Essa produção é completada por frações de outros processos, como o craqueamento retardado e o craqueamento catalítico, estabilizados por hidrotreatamento. De forma geral, o hidrotreatamento severo leva a uma ligeira redução da faixa de ebulição, da densidade, da lubrificidade, da viscosidade e ao aumento da estabilidade à oxidação e do número de cetano (FARAH, 2012).

Sua importância se deu a partir de 1893, quando Rudolf Diesel desenvolveu um motor de combustão interna com ignição por compressão. Os primeiros motores a diesel eram utilizados em navios, trens e plantas industriais. Na década de 1930, foram empregados também em caminhões e ônibus. Hoje, são utilizados nas mais diversas aplicações, abrangendo transporte, geração de energia, construção, máquinas agrícolas, mineração à céu aberto, aquecimento (DOS SANTOS, 2009).

Os hidrocarbonetos parafínicos são os constituintes do óleo diesel que apresentam as melhores características de combustão, enquanto os aromáticos são os menos desejáveis por apresentarem baixa qualidade de ignição no motor diesel.

Apesar da ótima qualidade de ignição, os parafínicos possuem maior facilidade de cristalização a baixas temperaturas, o que pode limitar a presença de hidrocarbonetos desse tipo de maior ponto de ebulição. Os naftênicos estão presentes no óleo diesel em quantidades importantes, entretanto, não trazem aspectos positivos e nem negativos no que diz respeito à qualidade. Os olefínicos, quando presentes mesmo em pequenas quantidades oriundos de frações dos processos de craqueamento catalítico e coqueamento retardado, podem apresentar problemas de estabilidade à oxidação e de ignição, bem como para que sejam reduzidos os contaminantes que apresentam (FARAH, 2012).

A composição média em volumes de hidrocarbonetos no óleo diesel é apresentada na Tabela 2. Além dos hidrocarbonetos, estão presentes no óleo diesel elementos como enxofre, oxigênio e nitrogênio. Os compostos de enxofre causam corrosão e contribuem para o aumento de emissão de materiais particulados. Os compostos de oxigênio conferem acidez e corrosividade, contribuindo também para a instabilidade termo-oxidativa. Os compostos de nitrogênio provocam instabilidade no combustível por reações de degradação (FARAH, 2012).

**Tabela 2** - Composição de hidrocarbonetos presentes no óleo diesel. (Adaptado de Farah, 2012).

<b>Hidrocarbonetos</b>	<b>Médio (% vol)</b>	<b>Máximo (% vol)</b>	<b>Mínimo (% vol)</b>
<b>Parafínicos</b>	30	62	15
<b>Naftênicos</b>	45	71	24
<b>Aromáticos Totais</b>	20	45	6
<b>Monoaromáticos</b>	10	18	3
<b>Diaromáticos</b>	12	23	0,5
<b>Triaromáticos</b>	5	14	0,3
<b>Poliaromáticos</b>	0,8	2	0,2

Fonte: Farah (2012).

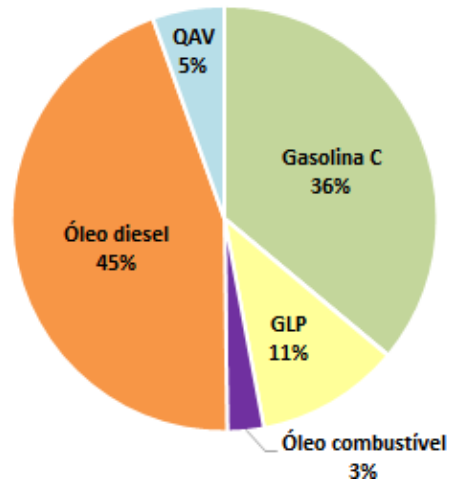
Existem também aditivos os quais são adicionados em pequenas quantidades ao óleo diesel cuja finalidade consiste em melhorar o desempenho ou corrigir alguma deficiência do combustível (FARAH, 2012), a saber:

- **Biocida:** utilizado para inibir o crescimento de micro-organismos que produzem metabólitos ácidos, que ocasionariam obstrução de filtros e danos aos injetores. Podem ser solúveis em óleo ou água, sendo que os solúveis em óleo permanecem durante todo o trajeto do diesel até a chegada no motor, enquanto os solúveis em água atuam somente até que a água seja removida;

- Antiespumante: usado para diminuir a formação de espuma no abastecimento do tanque tornando este processo mais rápido e reduz a tensão superficial das bolhas de óleo que entram em contato com o ar. São substâncias à base de silicone que ficam dispersas no óleo devido à sua insolubilidade;
- Melhorador do número de cetano: facilitador da autoignição, são substâncias do tipo alquilnitrato, que ao se decomporem em altas temperaturas, geram radicais livres que colaboram na ignição;
- Melhorador da lubricidade: forma uma película protetora que reduz o contato entre superfícies metálicas;
- Melhorador de escoamento a frio: empregados para reduzir os efeitos de obstrução e de restrição do fluxo. Reagem com os cristais de hidrocarbonetos formados no diesel modificando o tamanho, a forma e o grau de aglomeração destes cristais.

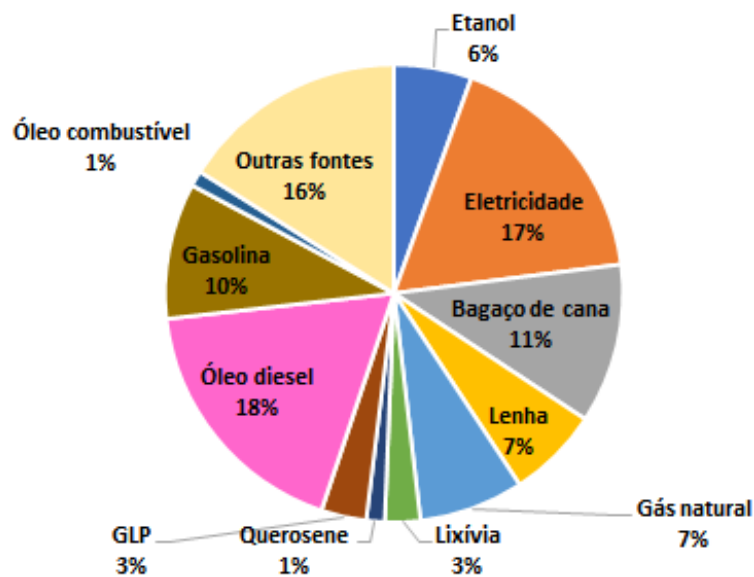
#### 2.1.4.1 A importância do óleo diesel no Brasil

O óleo diesel é o derivado de petróleo mais comercializado no Brasil, sendo que em 2017 sua venda representou 44,7% em volume de barril de petróleo processado (Figura 3), com um crescimento de 0,91 % em relação a 2016. A região Sudeste é responsável por comercializar 41% do diesel total brasileiro, sendo a maior participação nacional nas vendas (ANP, 2018a).



**Figura 3** - Vendas, pelas distribuidoras, dos derivados combustíveis de petróleo (m<sup>3</sup>) em 2017.

**Fonte:** Boletim ANP (2018a).

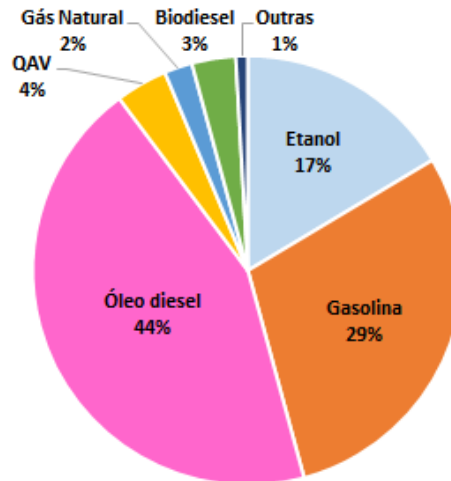


**Figura 4** - Participação de fontes energéticas no consumo total no Brasil em 2017.

**Fonte:** EPE (2018).

Em geral, o óleo diesel é a fonte de energia mais consumida no país, representando 18,1% do total (Figura 4). Segundo o Balanço Energético Nacional de 2018, referente ao ano de 2017, o segmento de transporte, com uma expansão de 1,9 milhões de toneladas equivalentes de petróleo (tep), liderou o avanço da demanda energética em 2017. Isto ocorreu principalmente em virtude do aumento de 2,7% do consumo de óleo diesel, consequência da maior atividade do setor de transporte de carga. Neste setor, o diesel representa 44% da matriz energética (Figura 5). O Balanço

Energético Nacional também apontou que 83,8% do consumo final do óleo diesel corresponde ao setor de transporte (EPE, 2018).



**Figura 5** - Participação dos combustíveis no setor rodoviário em 2017.

**Fonte:** EPE (2018).

Embora o Brasil seja autossuficiente na produção de petróleo desde 2006, ainda há importação de óleo diesel, principalmente dos Estados Unidos, devido à alta demanda que ultrapassa a capacidade de produção das refinarias brasileiras (SZKLO e ULLER, 2008; ANP, 2018a). O Anuário da Indústria Automobilística Brasileira referente ao ano de 2017 mostra que houve um aumento na produção e venda de veículos e máquinas agrícolas com motores do ciclo diesel, o que sugere que haverá consequentemente um aumento da demanda do combustível nos próximos anos (ANFAVEA, 2018).

#### 2.1.4.2 Especificações do óleo diesel rodoviário brasileiro

Devido à sua ampla utilização e grande importância no setor rodoviário, a garantia da qualidade do óleo diesel é imprescindível para haver um desempenho satisfatório e funcionamento adequado do motor. Alguns requisitos visam o atendimento de todas as exigências dos motores, tais como partida rápida e facilitada em qualquer clima, aquecimento rápido do motor, boas características de aceleração e emissão de poluentes em níveis aceitáveis (DOS SANTOS, 2009). Estes requisitos

e seus valores (Tabela 3) são especificados pela ANP, estando presentes na Resolução ANP número 50, de 2013 e apresentando complementações e algumas modificações ocasionadas pela Resolução ANP número 69, de 2014 (ANP, 2013, 2014).

**Tabela 3** - Especificações do óleo diesel rodoviário brasileiro de acordo com a Resolução ANP nº 50/2013, já adicionadas as modificações ocasionadas pela Resolução ANP nº 69/2014.

CARACTERÍSTICA	UNIDADE	LIMITE		MÉTODO	
		TIPO A e B		ABNT NBR	ASTM/EN
		S10	S500		
Aspecto	-	Límpido e isento de impurezas		14954	D4176
Cor ASTM, máx.	-	3,0		14483	D1500 D6045
Teor de biodiesel	% volume	*		15568	EN 14078
Enxofre total, máx.	mg/kg	10,0	-	-	D2622 D5453 D7039 D7212 D7220
		-	500	14533	D2622 D4294 D5453 D7039 D7220
Destilação					
10% vol., recuperados, mín.	°C	180,0	Anotar	9619	D86
50% vol., recuperados		245,0 a 295,0	245,0 a 310,0		
85% vol., recuperados, máx.		-	360,0		
90% vol., recuperados		-	Anotar		
95% vol., recuperados, máx		370,0	-		
Massa específica a 20°C	kg/m³	815,0 a 850,0	815,0 a 865,0	7148 14065	D1298 D4052
Ponto de fulgor, mín.	°C	38,0		7974 14598	D56 D93 D3828 D7094
Viscosidade Cinemática a 40°C	mm²/s	2,0 a 4,5	2,0 a 5,0	10441	D445
Ponto de entupimento de filtro a frio, máx.	°C			14747	D6371
Número de cetano, mín. ou Número de cetano derivado (NCD), mín.	-	48	42	-	D613 D6890 D7170
Resíduo de carbono Ramsbot-tom no resíduo dos 10% finais da destilação, máx.	% massa	0,25		14318	D524
Cinzas, máx.	% massa	0,010		9842	D482
Corrosividade ao cobre, 3h a 50°C, máx	-	1		14359	D130
Teor de Água, máx.	mg/kg	200	500	-	D6304 EN ISO 12937
Contaminação total, máx.	mg/kg	24	-	-	EN 12662
Água e sedimentos, máx.	% volume	-	0,05	-	D2709
Hidrocarbonetos policíclicos aromáticos, máx.	% massa	11	-	-	D5186 D6591 EN 12916
Estabilidade à oxidação, máx.	mg/100mL	2,5	-	-	D2274 D5304
Índice de acidez	mg KOH/g	Anotar	-	14248	D664

**Nota:** \* Porcentual estabelecido pela legislação vigente. O teor de biodiesel atualmente é 10% (Lei nº 13263/2016).

### 2.1.4.3 Qualidade do óleo diesel

A ANP é responsável pela regulamentação, contratação e fiscalização das atividades no mercado de petróleo. Em 1998, foi criado o Programa de Monitoramento da Qualidade dos Combustíveis (PMQC), que, por meio de contratos com universidades e centros de pesquisa mediante licitações públicas, tem como objetivo o levantamento de dados dos principais combustíveis líquidos comercializados e identificar eventuais focos de não conformidade. Atualmente, o programa contempla os estados de Alagoas, Sergipe, Pernambuco, Amapá, Pará, Bahia, Ceará, Tocantins, Goiás, Maranhão, Minas Gerais, Paraíba, Rio Grande do Norte, Rio de Janeiro, Espírito Santo, São Paulo, Santa Catarina, Mato Grosso do Sul, Paraná e Rio Grande do Sul (ANP, 2018b). Diariamente, são efetivadas coletas e análises físico-químicas de amostras de gasolina, etanol hidratado combustível e óleo diesel em postos revendedores escolhidos por sorteio, e, mensalmente, os relatórios desses dados são divulgados pela ANP (LIMA, 2012; ANP, 2018b).

Conforme o crescimento do PMQC, os índices de não conformidade dos combustíveis foi diminuindo, demonstrando a utilidade do programa aliado às fiscalizações no panorama de qualidade de combustíveis comercializados. O quadro atual consta de 14 instituições distribuídas ao longo do território nacional. O estado de São Paulo conta com duas destas instituições, a Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP e o Instituto de Pesquisas Tecnológicas de São Paulo – IPT-SP (ANP, 2018b). Na UNESP, dentro do Instituto de Química no *campus* de Araraquara, há o Centro de Monitoramento e Pesquisa da Qualidade de Combustíveis, Biocombustíveis, Petróleo e Derivados (CEMPEQC), que consiste em um conjunto de laboratórios com capacitação técnica para realização de ensaios e pesquisa científica nesta área. O CEMPEQC é responsável pela coleta e análise de combustíveis de 408 municípios no estado de São Paulo e 79 no estado do Mato Grosso do Sul, resultando em um total de 3627 postos.

Somente os principais parâmetros físico-químicos dos combustíveis são analisados pelos laboratórios credenciados ao PMQC e no óleo diesel esses parâmetros são: aspecto, cor ASTM, massa específica, enxofre total, teor de biodiesel,

ponto de fulgor e temperatura de ebulição das frações 10%, 50%, 85% e 90% destiladas. Estes parâmetros serão brevemente discutidos a seguir.

➤ Aspecto

É um teste visual, indicador de qualidade e de possível contaminação. O óleo diesel deve possuir aspecto límpido e isento de materiais suspensos (água, poeira, ferrugem, etc.), que quando presentes comprometem o bom funcionamento do motor e podem reduzir a vida útil dos filtros de combustível dos veículos (CNT, 2012).

➤ Cor ASTM

O óleo diesel usualmente apresenta coloração de incolor a amarelada, podendo apresentar ligeira alteração para tons amarronzados e alaranjados devido ao biodiesel nele contido. A cor ASTM especifica uma faixa de tonalidades como um indicativo para verificar contaminação ou degradação. O limite disposto na resolução é requerido antes da adição obrigatória de 20 mg/L de corante vermelho ao diesel S500, que contém no máximo 500 mg/kg de enxofre, a fim de diferenciá-lo do S10, que possui no máximo 10 mg/kg de enxofre.

O ensaio deve ser realizado de acordo com as normas ASTM D1500 e D6045 e a norma ABNT NBR 14483.

➤ Massa específica

É definida como a relação entre a massa e o volume do combustível. O valor da massa específica estar adequado é importante para o bom funcionamento do motor, que é projetado para operar em uma determinada faixa de massa específica devido a bomba injetora dosar o volume que é injetado. Variações na massa de combustível injetado dificultam a obtenção de uma mistura ar/combustível balanceada. Por isso, valores abaixo do especificado podem favorecer a perda de



potência do motor e aumento do consumo de combustível e valores acima tendem a aumentar a emissão de poluentes.

Sua medida é realizada através do uso de densímetro segundo as normas ABNT NBR 7148 e 14065 e ASTM D4052 e pelo método do hidrômetro pela norma ASTM D1298 (ASTM)

➤ Enxofre total

O óleo diesel é classificado conforme o teor de enxofre contido, que pode ser máximo de 10 mg/kg denominado S10 ou máximo de 500 mg/kg, denominado S500.

O enxofre é um elemento químico indesejável para o meio ambiente e também para os motores diesel. Durante a combustão, o trióxido de enxofre, ao se juntar à água, forma o ácido sulfúrico, que corrói partes metálicas do motor, como mancais e guias de válvulas, por exemplo. Se a concentração de enxofre for alta, as emissões de material particulado também serão elevadas, assim como as emissões de poluentes primários como  $\text{SO}_2$  e  $\text{SO}_3$ , acarretando grandes prejuízos à saúde humana. Os óxidos de enxofre, produzidos no processo de queima do enxofre, como no caso da combustão dos veículos a diesel, também são irritantes e tóxicos para os seres humanos (CNT, 2012).

O teor de enxofre é determinado por espectrometria de fluorescência de raios X segundo as normas ASTM D2622, D7039, D7212, D7220 e por fluorescência ultravioleta pela norma ASTM D5453.

➤ Destilação

A destilação indica contaminações por frações mais leves, como querosene e gasolina ou mais pesadas, como óleo lubrificante. Um combustível com valores de temperaturas de destilação de acordo das especificações contribui com o bom desempenho do motor, minimiza a formação de depósitos no sistema de combustão e reduz a emissão de poluentes. Combustíveis com volatilidade muito baixa ou muito alta tendem a reduzir a potência do motor e a economia de combustível (ANP, 2018a). A volatilidade é medida segundo as normas (ABNT NBR 9619, ASTM D86). por um procedimento onde a amostra de diesel é destilada e a temperatura do vapor é

registrada para cada porcentagem recuperada. Quatro pontos são monitorados, nos quais 10%, 50%, 85% e 90% do volume da amostra destilada são recuperados.

➤ Teor de biodiesel

Conforme a legislação vigente, de março de 2017 a fevereiro de 2018, o teor de biodiesel adicionado ao óleo diesel era 8% (v/v). A partir de março de 2018 este valor passou a ser 10%. A medida de teor de biodiesel deve ser efetivada de acordo com a norma ABNT NBR 15568 ou com a norma EN 14078, que fazem esta determinação utilizando espectroscopia na região do infravermelho médio.

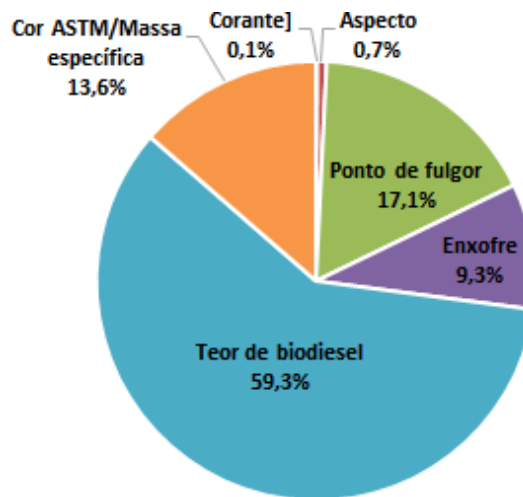
➤ Ponto de fulgor

Por definição, ponto de fulgor é a temperatura na qual o óleo diesel vaporiza em quantidade suficiente para formar uma mistura explosiva junto ao ar que pode se inflamar quando submetido a uma fonte de ignição. Este parâmetro está relacionado com o manuseio e a segurança do produto e indica possível contaminação com frações mais leves, como querosene e gasolina. Sua determinação é feita segundo as normas ABNT NBR 7974 e ASTM D56 (método do vaso fechado TAG), NBR 14598 e ASTM D93 (analisador *Pensky-Martens* de vaso fechado) e as normas ASTM D3828 (método do vaso fechado em pequena escala) e D7094 (método do vaso fechado continuamente modificado).

#### 2.1.4.4 Não conformidades no óleo diesel

Segundo o Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis de 2018 com dados referentes ao ano de 2017, no que diz respeito ao óleo diesel, foram observadas 3,3% não conformidades entre as 32555 amostras analisadas, das quais 59,3% relativas ao teor de biodiesel; 13,6% a cor ASTM e massa específica a 20 °C; 17,1% a ponto de fulgor; 9,3% concentração de enxofre no combustível; 0,1% corante e 0,7% aspecto (Figura 6) (ANP, 2018a). É importante salientar que a maior incidência de não conformidade ser em relação ao parâmetro biodiesel é devido ao descumprimento ao de teor de biodiesel (% v/v) adicionado determinada pela legislação vigente (De 6% passou a 8% no mês de março de 2017)

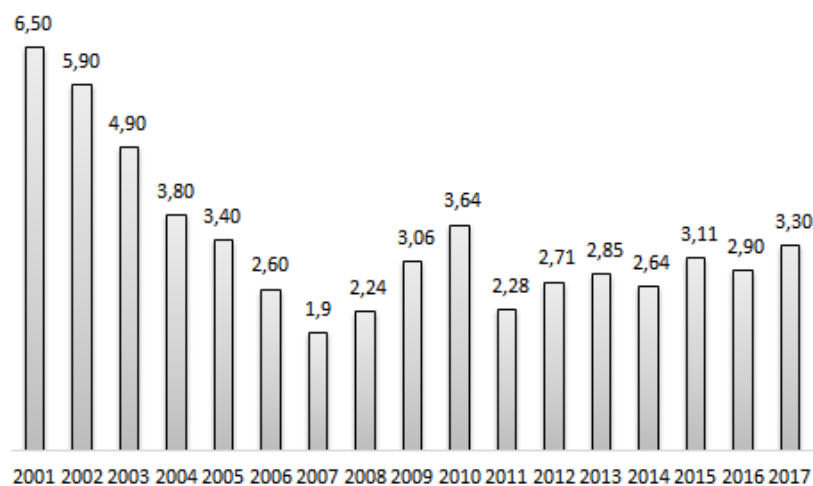
e não há impactos negativos causados por este quesito em relação à qualidade do combustível.



**Figura 6** - Não conformidades observadas no óleo diesel em 2017.

**Fonte:** ANP (2018a) – adaptado pela autora.

Desde a implantação do PMQC, os índices de não conformidade do óleo diesel foram sendo reduzidos com o passar dos anos (Figura 7). Entre 2001 e 2007, essa queda foi bastante nítida, contudo, houve um aumento entre os anos de 2008 e 2010 provocado pela alteração da especificação do óleo diesel com a adição obrigatória de biodiesel e aumento de seu valor na mistura.



**Figura 7** - Índice de não conformidade do óleo diesel encontrado no Brasil ao longo dos anos.

**Fonte:** ANP (2018a) – adaptado pela autora.

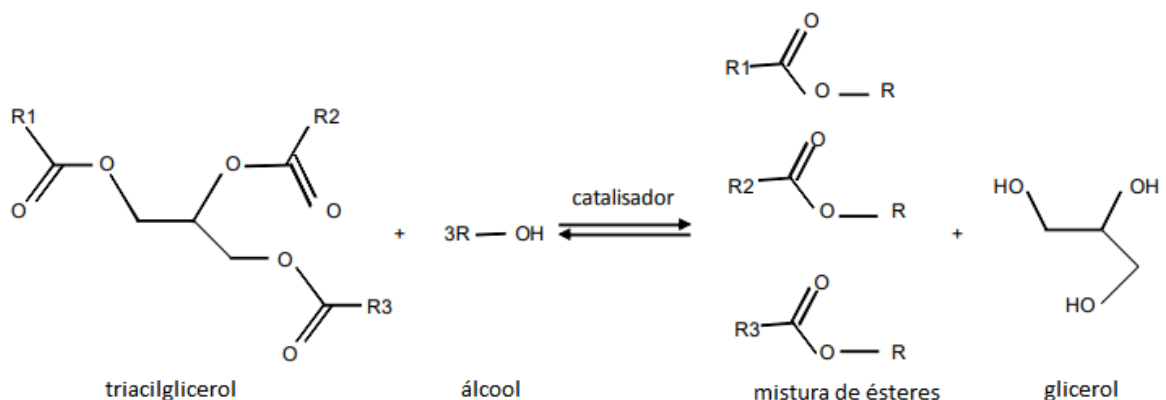
Conforme as novas regulamentações sobre aumento do teor de biodiesel adicionado ao diesel foram surgindo, novamente os índices de não conformidades anuais alavancavam. No entanto, comparado aos primeiros anos de monitoramento da qualidade de combustíveis, estes índices estão consideravelmente menores. Entre os combustíveis monitorados, o diesel é o que apresenta maior índice de amostras não conformes, seguido pelo etanol com 1,8% e a gasolina com 1,5% (ANP, 2018a).

### 2.1.5 Biodiesel

Por definição, de acordo com a Lei n.º 11097/2005, biodiesel é o biocombustível derivado de biomassa renovável, para uso em motores a combustão interna por compressão ou, conforme regulamento, para a geração de outro tipo de energia, que possa substituir, parcial ou totalmente, combustíveis de origem fóssil (BRASIL, 2014a)

O biodiesel consiste de uma mistura de ésteres alquílicos de ácidos graxos resultantes da transformação química de óleos ou gorduras (triacilglicerídeos) oriundos de fontes biológicas renováveis (PIANOVISKI JUNIOR, 2002). Pode ser obtido através dos processos de craqueamento térmico, esterificação ou transesterificação, contudo a tecnologia para produção de biodiesel predominante no mundo é o processo de transesterificação (Figura 8).

Na transesterificação, os triacilglicerídeos reagem com álcool metílico ou etílico na presença de um catalisador ácido ou básico, produzindo uma mistura de ésteres alquílicos de ácidos graxos e glicerol (GASPARINI, 2010).



**Figura 8** - Reação geral de transesterificação.

**Fonte:** Gasparini (2010) – adaptado pela autora.

Existem culturas que se destacam pelo grande potencial de produção de biodiesel através de extração dos óleos, tais como: soja, dendê, girassol, algodão, linhaça, côco, palma, amendoim, colza/canola e babaçu. Além dos óleos vegetais, também são utilizados como fonte de triacilglicerídeos o sebo bovino, a banha suína e óleos e gorduras residuais (OGR) (BUENO, 2007; SLUSZZ; MACHADO, 2006).

A viabilidade de cada oleaginosa para produção do biocombustível dependerá da sua competitividade técnica, socioeconômica e ambiental, além de aspectos como o teor de óleo, produtividade por unidade de área, adequação a diferentes sistemas produtivos, sazonalidade (relacionada ao ciclo de vida da planta), e distribuição territorial de acordo com as condições edafoclimáticas (RAMOS, 2003). Em 2017, as matérias-primas mais empregadas na produção de biodiesel foram o óleo de soja (71,6%), gordura animal, incluindo bovina, suína e de frango (16,8%) e o óleo de algodão (0,3%). Outros materiais graxos, tais como óleo de palma, óleo de amendoim, óleo de girassol, óleo de nabo-forrageiro, óleo de sésamo e óleo de fritura usado, somaram 11,3% (ANP, 2018c).

#### 2.1.5.1 O biodiesel na matriz energética brasileira

A possibilidade de uso de óleos vegetais no motor diesel existe desde a sua invenção, pois Rudolf Diesel utilizou óleo de amendoim em sua demonstração pública. No entanto, o rumo que tomou a história e o desenvolvimento da indústria petrolífera, bem como de seus interesses, deram ao motor de ciclo diesel, como destino, ser alimentado com derivados do petróleo. Os ésteres começaram a ser pesquisados na década de 40, e a primeira patente em termos mundiais sobre o referenciado biodiesel – PI 8007957, tem como autor o Prof. Expedito Parente, professor da Universidade Federal do Ceará, concedida em 1983, com pesquisas que remontam às décadas de 1970 e 1980. Esta patente já teve os privilégios em domínio público em 1993 por falta de uso (HOLANDA, 2004).

Na década de 70, devido ao alto preço do petróleo, o Brasil criou programas de incentivo ao desenvolvimento de combustíveis oriundos de biomassa, o PROÁLCOOL foi o pioneiro, em 1975, incentivando a substituição da gasolina por etanol produzido através de cana-de-açúcar. Em 1980, foi criado o programa de incentivo ao biodiesel conhecido como PRODIESEL. Muitos avanços foram alcançados, mas devido à diminuição dos preços do petróleo e aos altos custos de implantação e produção dos

programas, não houve a inserção e o desenvolvimento efetivo. Atualmente, devido à uma necessidade de maior valor ambiental que econômico, tornou-se necessário retomar as pesquisas e estudos e para a implantação do biodiesel (BUENO, 2007).

A ANP autorizou a realização de testes de combustíveis não especificados através da Portaria ANP nº180/1998. Em consequência das pesquisas realizadas com o biodiesel, em 2004, o programa foi reformulado e passou a ser chamado Programa Nacional de Produção e Uso do Biodiesel (PNPB) (ANP, 1998). A Lei 11097/05 foi a responsável pela introdução do biodiesel na matriz energética brasileira e com ela o PNPB iniciou os testes em motores para validar o uso de mistura diesel/biodiesel, denominadas BX (onde o X corresponde ao valor da porcentagem de biodiesel) a fim de assegurar a garantia de qualidade de veículos e equipamentos (BRASIL, 2014a).

O biodiesel é adicionado ao óleo diesel com o propósito de reduzir a dependência por produtos do petróleo e por razões ambientais, como a diminuição de emissão de gases poluentes, além de motivos socioeconômicos como expansão do mercado de oleaginosas, incentivo a agricultura familiar e aumento do volume de combustível no mercado (MATTEI, 2008). Ele é a opção mais aceita para a substituição do óleo diesel, devido a não necessidade de modificação mecânica dos motores de ciclo diesel atuais. Este biocombustível também aceita misturas com o óleo diesel em qualquer proporção (BUENO, 2007).

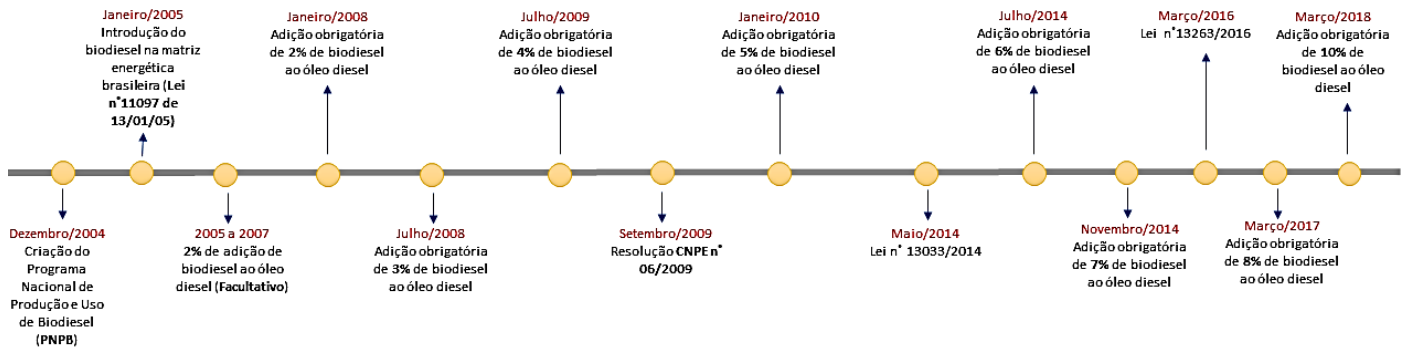
A Lei 11.097/05 prevê que o percentual de biodiesel adicionado ao diesel seja elevado gradualmente ao longo dos anos. Em 2005, a adição de biodiesel ao diesel foi permitida na proporção de 2:98 (v/v), chamada mistura B2, e passou a ser obrigatória em 2008. Assim, era previsto que o teor de biodiesel aumentasse para 5% apenas em 2012, porém, o Conselho Nacional de Política Energética (CNPE) reduziu os prazos para atendimento do percentual mínimo obrigatório e o percentual de 5% de biodiesel passou a ser obrigatório em 2010 (BRASIL, 2014a). A Lei nº 13.033/2014 passou ao CNPE a atribuição de definir diretrizes para comercialização e uso de biodiesel e estabelecer, em caráter autorizativo, quantidade superior ao percentual de adição obrigatória fixado em lei específica. Adicionalmente, a referida Lei alterou o percentual obrigatório de biodiesel ao óleo diesel de 5% para 6%, a partir de 1º de julho de 2014 e 7% a partir de 1º de novembro de 2014 (BRASIL, 2014b). Em 21 de setembro de 2015, o CNPE publicou a Resolução nº 3, para autorizar e definir as diretrizes para comercialização e uso voluntário de biodiesel, autorizando assim, em quantidade superior ao percentual de sua adição obrigatória ao óleo diesel,

observados os limites máximos de adição de biodiesel ao óleo diesel, em volume: 20% em frotas cativas ou consumidores rodoviários atendidos por ponto de abastecimento; 30% no transporte ferroviário; 30% no uso agrícola e industrial e 100% no uso experimental (BRASIL, 2015).

A ANP realizou a revisão da Resolução ANP nº 2/2011, tornando a especificação de B6 a B20 para BX a B30, sendo o “X” o teor logo acima do teor obrigatório determinado pelo CNPE. Com isso, foi publicada a Resolução ANP nº 30/2016, que trata do uso experimental ou específico de biodiesel ou de sua mistura com óleo diesel, em quantidade superior ao percentual de adição de biodiesel obrigatória, conforme autorizado pelo CNPE. De acordo com esta resolução, os testes governamentais em curso não necessitam atender os limites nela contidos, contudo, em outras situações de uso experimental, específico ou de eventos, o distribuidor de combustíveis deve atender à especificação tratada no Regulamento Técnico ANP nº 2/2016, parte integrante de dita resolução apenas quando o combustível BX, com teor inferior ou igual a B30, for utilizado em frotas cativas ou consumidores rodoviários supridos por Ponto de Abastecimento (ANP, 2015, 2016).

A Lei nº 13.263/2016 alterou a Lei nº 13.033, de 24 de setembro de 2014, para dispor sobre os percentuais de adição de biodiesel ao óleo diesel comercializado no território nacional (BRASIL, 2016). A partir de março de 2017 o teor de biodiesel (% v/v) passou a ser de 8% e a partir de março de 2018, seria de 9%, entretanto, o CNPE definiu, conforme a Resolução nº 23, de 09 de novembro de 2017, que a partir de março de 2018, o percentual de adição obrigatória de biodiesel no diesel seria de 10%, que é o valor adicionado até os dias atuais e o CNPE definiu em 29/10/18, pela Resolução nº 16, o cronograma para evolução do B10 (10% de biodiesel adicionado ao diesel) até o B15, com foco na redução do preço do combustível para o consumidor e ampliação da participação de energia renovável na matriz energética (BRASIL, 2017, 2018). Portanto, a partir de junho de 2019, todo o diesel comercializado no território nacional será B11, isto é, terá a mistura de 11% do biocombustível. A adição evoluirá 1% ao ano, chegando a B15 em 2023 (BRASIL, 2018; UBRABIO, 2018).

A Figura 9 elucida a evolução do uso de biodiesel no panorama brasileiro. A Resolução ANP nº 45, de 25 de agosto de 2014 dispõe sobre a especificação do biodiesel contida no Regulamento Técnico ANP nº 3 de 2014 e as obrigações quanto ao controle da qualidade a serem atendidas pelos diversos agentes econômicos que comercializam o produto em todo o território nacional (ANP, 2014c; BRASIL, 2018).



**Figura 9 -** Evolução do uso de biodiesel no Brasil.

**Fonte:** BRASIL (2018) – adaptado pela autora.

## 2.2 Cromatografia

### 2.2.1 Um breve histórico

A etimologia da palavra cromatografia, como foi primeiramente empregada, é definida pelos termos de origem grega a partir das palavras *chroma* e *graphe*, que significam cor e escrever, respectivamente, cujo sentido seria algo como “escrita das cores” (SAKODYNSKII, 1993; COLLINS, 2006; COLLINS, 2009; ETTRE, 2008a; ETTRE 2008b; PACHECO *et al.*, 2015).

Encontramos na literatura o fato de ser atribuído ao botânico russo Mikhael Semenovich Twsett, a primeira descrição e/ou “descoberta” da técnica de cromatografia por meio de seus reconhecidos estudos nos anos de 1903 e 1906, sobre a separação e identificação de diferentes componentes em pigmentos de plantas por meio de um método de separação baseado na adsorção, mesmo tendo Twsett revelado que o processo de separação que ocorria nestes experimentos não tinha relação de dependência alguma com a cor dos compostos, mas sim dependia essencialmente das interações das substâncias contidas nos extratos de clorofilas com o material presente no interior da coluna, os quais eram separados após a passagem de diferentes solventes orgânicos pela coluna (COLLINS, 2006; COLLINS, 2009; ETTRE, 2008a, 2008b, 2008c).

Os primeiros relatos os quais fazem menção aos primeiros processos de separação de substâncias inorgânicas e/ou orgânicas os quais posteriormente auxiliaram direta e/ou indiretamente na construção dos fundamentos da cromatografia



como conhecemos atualmente, são mais antigos do que parece e envolvem muitos personagens do ramo científico e industrial que precederam Tswett e que ficaram esquecidos no tempo sem nenhum reconhecimento pelas suas contribuições (ABRAHAM, 2006; ETTRE, 2008b; COLLINS, 2006; COLLINS, 2009; PACHECO *et al.*, 2015).

No entanto, fato é que foi Tswett quem efetivamente introduziu o termo cromatografia em seu trabalho publicado em 1906 (TSWETT, 1906a). Outro aspecto muito relevante que estes trabalhos publicados em 1906 (TSWETT, 1906a, 1906b) trouxeram foi a inserção da palavra cromatograma, a qual foi empregada para descrever um desenho das bandas separadas na coluna, onde Tswett visualiza-as dentro da própria coluna de vidro. Outro fato curioso que norteia a história da “invenção” da cromatografia foi a própria origem das palavras utilizadas por Tswett (ETTRE, 2008c).

O uso da palavra grega *chroma* que Tswett empregou, embora tenha sido cuidadosamente esclarecido por ele que o mecanismo de separação o qual descrevia não dependia da cor, teria sido na verdade “pensado” em idioma russo (цвет) cuja fonética é *tsvet*, de modo que seu significado é o mesmo que *chroma*, ou seja, cor. Por isso, acredita-se que Tswett teria nomeado sua “nova” técnica como “Escrita de Tswett” ou Tswettgraphy (ABRAHAM, 2006; ETTRE, 2008c).

Portanto, como Tswett apenas nomeou a técnica, não fundamentando de maneira adequada a origem da definição do termo, bem como não apresentando explicações que permitissem boa compreensão na época, a cromatografia tal como foi concebida por ele ficou ignorada por quase 30 anos (COLLINS, 2006; PACHECO *et al.*, 2015).

Somente em 1931, quando Kuhn e Lederer reformularam e aprimoraram, com conceitos bem consolidados, os métodos e técnicas de separação baseadas na cromatografia de adsorção em coluna, remodelando e ampliando os experimentos de Tswett, foi que se considerou de fato o marco inicial da cromatografia moderna. Nestes trabalhos, Kuhn e Lederer conseguiram separar e identificar as xantofilas de gema de ovo, usando como fase móvel éter de petróleo e como fase estacionária carbonato de cálcio pulverizado (ETTRE, 2001; COLLINS, 2006; COLLINS, 2009).

### 2.2.2 Cromatografia gasosa: aspectos gerais, parâmetros cromatográficos e instrumentação

Passaram-se quase 50 anos da divulgação da técnica de cromatografia elaborada por Tswett para ocorrer de fato o desenvolvimento da cromatografia em fase gasosa (CG). Porém, não demorou nem uma década para que o aprimoramento da cromatografia em fase gasosa acontecesse de maneira vertiginosa, sobretudo, em 1960 a partir dos avanços na instrumentação (COLLINS; BRAGA; BONATO, 2006; LANÇAS, 2017).

Na cronologia dos fatos, visando contornar o problema da identificação de pequenas quantidades de aminoácidos hidrolisados em proteínas, Martin e Synge em 1941, propuseram um método de separação microanalítico baseado no fenômeno de partição. Este trabalho de Martin e Synge representa outro marco importantíssimo na história da cromatografia, pois retrata o surgimento da cromatografia por partição a qual foi considerada como precursora da cromatografia em papel e cromatografia líquido-líquido, bem como anteciparam o surgimento de técnicas pressurizadas de cromatografia as quais posteriormente vieram a dar surgimento a cromatografia gasosa. Logo, por esse retumbante feito, Martin e Synge são condecorados com o prêmio Nobel de Química em 1952 (BARTLE; MYERS, 2002; COLLINS; BRAGA; BONATO, 2006).

Outros importantes trabalhos foram realizados por Martin os quais contribuíram efetivamente para o desenvolvimento da cromatografia. Neste contexto, Martin, Consden e Gordon reformularam a cromatografia em papel a partir cromatografia de partição líquido-líquido, e juntamente com James, Martin outra vez firmou seu nome na história da cromatografia pela invenção da cromatografia gás-líquido em uma reunião realizada pela Sociedade de Bioquímica no ano de 1950 em Oxford. Martin e James reportaram neste evento a separação de ácidos graxos voláteis por cromatografia de partição utilizando como fase móvel gás nitrogênio, e uma fase estacionária composta de óleo de silicone e ácido esteárico suportados em terra de diatomáceas (BARTLE; MYERS, 2002; LANÇAS, 2017).

Apesar de ser atribuído a Martin e James o posto de inventores da cromatografia gás-líquido, Hesse, Cramers e Leclercq já haviam descrito um tipo de cromatografia gás-líquido de adsorção em 1941, bem como técnica semelhante havia sido relatada em 1950 por Cremer e Prior, onde apresentaram de maneira completa

e detalhada um cromatógrafo a gás (SMOLKOVÁ-KEULEMANSOVÁ, 2000; VAITSMAN; CIENFUEGOS, 2000)

Avanço exponencial na cromatografia em fase gasosa ocorreu no fim da década de 50 e início de 60, após o desenvolvimento quase simultâneo de colunas capilares por Golay e aperfeiçoamento de detectores termiônicos realizado por Harley e Pretorius, McWilliams e Dewar, que aumentou a sensibilidade da técnica por um fator de  $10^3$ - $10^4$  além de permitir a detecção em escala  $10^9$  g (BARTLE; MYERS, 2002; LANÇAS, 2017; VAITSMAN; CIENFUEGOS, 2000)

A excelente sensibilidade aliada ao elevado poder de resolução garantiu também o sucesso do uso de CG, uma vez que para uma corrida convencional por CG, dependendo da substância a ser analisada e do detector acoplado ao sistema cromatográfico consegue-se uma quantificação da ordem de pictogramas. Outro fator que agrega potencial à CG reside nas pequenas quantidades de amostras necessárias para realizar uma análise frente outras técnicas analíticas (COLLINS; BRAGA; BONATO, 2006; LANÇAS, 2017).

Por outro lado, assim como na maioria das técnicas analíticas, desvantagens também estão presentes na CG. Neste sentido, sem dúvida um dos maiores inconvenientes presentes na CG convencional reside na necessidade de preparação das amostras, em alguns casos, antes da análise propriamente dita, a qual pode ser onerosa demandando bastante tempo e custo (analíticas (COLLINS; BRAGA; BONATO, 2006). Embora seja possível analisar amostras líquidas e sólidas por CG, estas devem ser devidamente solubilizadas em solventes voláteis, bem como seus componentes serem estáveis e apresentem uma pressão de vapor na temperatura operacional do sistema, além de interagir com o material da coluna (adsorvente sólido ou filme suportado) (BRAITHWAITE; SMITH, 1999). Se a amostra não for volátil, as técnicas de derivatização ou pirólise acoplada à cromatografia gasosa (Pi-CG) podem ser utilizadas. Esta última técnica é uma modificação na CG clássica em que uma amostra não volátil é pirolisada antes de entrar na coluna. Assim, os produtos de decomposição são separados na coluna, e, posteriormente, são qualitativamente e quantitativamente identificados de maneira convencional (SILVÉRIO; BARBOSA, 2008).

Outro aspecto negativo refere-se ao fato de que a CG não pode ser considerada uma técnica qualitativa de grande eficiência, pois para se ter certeza das informações qualitativas extraídas de uma análise cromatográfica gasosa exige-se técnicas

complementares para auxiliar na detecção de diferentes substâncias na amostra (COLLINS; BRAGA; BONATO, 2006).

Em um processo cromatográfico em fase gasosa um gás inerte (gás de arraste) que não deve reagir com a amostra, fase estacionária ou superfícies do instrumento de modo apenas a “carregar” os componentes da amostra é empregado como fase móvel e um líquido ou um sólido pode ser utilizado como fase estacionária. Dependendo do estado físico da fase estacionária, a cromatografia gasosa pode ser subdividida em dois tipos: cromatografia de gás-líquido (CGL) e cromatografia gás-sólido (CGS). Assim, na CGS a fase estacionária é um sólido (adsorvente) enquanto que na CGL é um líquido aplicado como um filme fino sobre um suporte sólido (BRAITHWAITE, A.; SMITH, 1999; COLLINS; BRAGA; BONATO, 2006; LANÇAS, 2017).

#### 2.2.2.1 Modalidades e definições de cromatografia gasosa rápida

A rapidez de uma separação cromatográfica realizado por um sistema cromatográfico poder ser representada por distintas formas. Assim, se o tempo gasto para realizar uma separação de espécies componentes de uma amostra complexa é de alguns minutos, por outro lado uma separação de uma amostra simples pode ocorrer em milissegundos (CRAMERS *et al.*, 1999; KORYTÁR *et al.*, 2002). Portanto, o conceito que define a rapidez ou velocidade do tempo de separação não se limita apenas na ideia de tempo total (duração) de análise. Tal conceito é mais amplo e deve considerar, além do tempo necessário para que separação ocorra, outros parâmetros cromatográficos (SEQUINEL *et al.*, 2010).

Em um esforço para quantificar o aumento na velocidade de análise em CG foi realizado por Dagan e Amirav (1995). Estes autores sugeriram a terminologia: cromatografia gasosa rápida (CG-R), cromatografia gasosa muito rápida (CG-MR) e cromatografia gasosa ultrarrápida (CG-UR), para descreverem as modalidades de cromatografia realizadas em velocidades de separação consideravelmente maiores do que a CG convencional. Para diferenciar essas novas modalidades de CG Dagan e Amirav (1998) basearam-se em um termo denominado fator de aumento de velocidade (*speed enhancement factor* - SEF), o qual é definido como o produto da redução do comprimento da coluna pelo aumento da velocidade linear do gás de arraste em relação a separação cromatográfica realizada em CG convencional.

O SEF não representa necessariamente uma condição exata de redução do tempo de análise. Contudo, normaliza as separações em função da CG convencional (CG-C) a qual, segundo Dagan e Amirav, é caracterizada pela uso de colunas de 30 m com diâmetro interno reduzido, de modo que o fluxo do gás de arraste na coluna seja de 1 mL/min (correspondente ao He com velocidade linear de 34 cm/s) (SEQUINEL *et al.*, 2010).

Usando essa descrição, os SEFs de 5–30, 30–400 e 400–4000 correspondem a CG-R, CG-MR e CG-UR, respectivamente. Na prática, os SEFs são normalmente alcançados diminuindo o comprimento da coluna ou o diâmetro interno da coluna e/ou aumentando as pressões de gás de entrada, mantendo a resolução cromatográfica suficiente (DAGAN; AMIRAV, 1998).

Uma definição baseada na largura à meia altura do pico ( $w_{1/2}$ ) e no tempo total de análise foi proposta por Van Deursen e colaboradores para interpretar o aumento de velocidade em CG. Segundo os autores, a largura do pico fornece uma noção do grau de separação por unidade de tempo para cada componente da amostra, de modo que o tempo das separações realizadas por CG-UR, ficariam limitadas ao intervalo dos subsegundos, enquanto que a largura dos picos estaria restritos a aproximadamente 10 ms (van DEURSEN *et al.*, 2000).

Uma maneira semelhante a Van Deursen e colaboradores foi proposta por Blumberg e Klee para avaliar o aumento de velocidade. Assim, Blumberg e Klee categorizaram as diferentes modalidades de CG de forma bem objetiva, baseado apenas na largura dos picos cromatográficos (BLUMBERG; KLEE, 2000; BICCHI *et al.*, 2004). Por outro lado, Magni e colaboradores formularam uma maneira mais precisa que vai encontro as anteriores mencionadas aqui até o momento. Desse modo, de acordo com autores, parâmetros tais como: dimensões da coluna e a taxa de aquecimento, também devem ser considerados para distinguir as modalidades de CG baseado no conceito de aumento de velocidade (BICCHI *et al.*, 2004; MAGNI *et al.*, 2002, SEQUINEL *et al.*, 2010).

A Tabela 4 mostra uma relação entre alguns parâmetros cromatográficos e classificação dos diferentes tipos de CG em função do tempo de análise ou aumento da velocidade separação, conforme a definição proposta por Magni e colaboradores (MAGNI *et al.*, 2002). Como pode ser observado na Tabela 4, a CG-UR é caracterizada utilizando colunas com dimensões bem reduzidas e altas taxas de aquecimento, o que resulta em picos com largura entre 0,05 e 0,2s.

**Tabela 4** - Principais parâmetros que classificam os principais tipos de cromatografia gasosa de acordo com Magni e colaboradores.

<b>PARAMENTOS</b>	<b>CG-UR</b>	<b>CG-R</b>	<b>CG-CC</b> (coluna curta)	<b>CG</b> (convencional)
<b>Comprimento da coluna (m)</b>	2 a 10	5 a 15	5	25 a 30
<b>Diâmetro interno (mm)</b>	0,05 a 0,10	0,10 a 0,25	0,25	0,25 a 0,32
<b>Tempo de análise (s)</b>	≤ 1	≤ 10	3 a 15	10 a 60
<b>Taxa de aquecimento (°C/min)</b>	≥ 60	15 a 60	5 a 40	1 a 10
<b>Largura do pico (s)</b>	0,05 a 0,2	0,5 a 2	1 a 5	1 a 10

Fonte: Sequinel *et al.* (2010).

Embora a terminologia CG-UR e CG-R sejam empregadas com frequência na literatura, de fato, não há uma posição definitiva estabelecida tanto pela comunidade reguladora quanto acadêmica a respeito dessas classificações. Logo, essa certa displicência entre os órgãos competentes em estabelecerem de maneira definitiva, termos que classifiquem as diferentes modalidades de CG em função do aumento de velocidade de separação, acaba gerando discordâncias permitindo com que usuários comuns e pesquisadores utilizem da subjetividade para classificarem o tipo de modalidade utilizada em seus laboratórios, sobretudo em relação à CG-UR (SEQUINEL *et al.*, 2010).

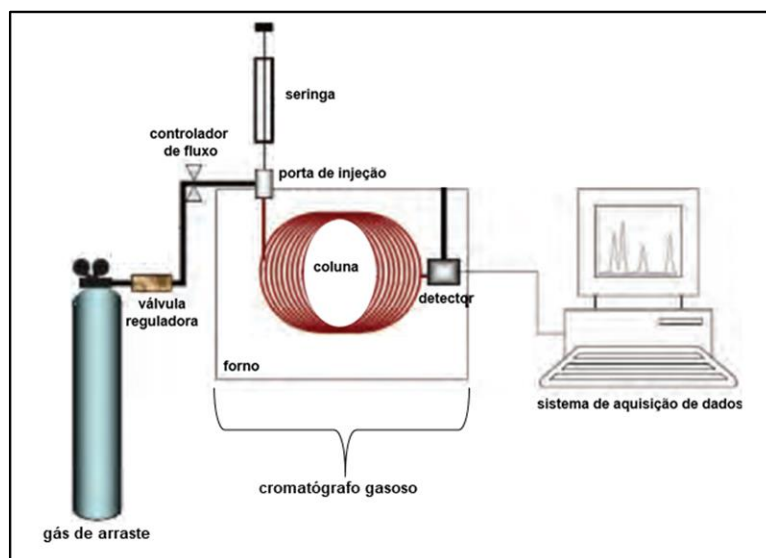
#### 2.2.2.2 Instrumentação

O alcance da notoriedade da cromatografia gasosa como uma das técnicas de separação mais conhecida e empregada no mundo, como vimos a seção 2.2.2, passou pelos intensos avanços na evolução de sua instrumentação a partir da década de 50, sobretudo no que se refere ao desenvolvimento das colunas capilares, das novas tecnologias de injeção de amostra e dos novos e aprimorados sistemas de detecção.

A instrumentação básica presente em CG contém características comuns a praticamente todos os tipos de cromatografia em coluna. Dessa forma, os componentes fundamentais que integram um sistema de CG são: (i) linha de fornecimento e controle do fluxo do gás de arraste (válvulas), (ii) dispositivo de introdução de amostras (injetores ou amostradores automáticos), (iii) coluna, (iv) forno da coluna (forno de aquecimento); (v) detector e (vi) sistema de amplificação do sinal e de processamento dos dados (integradores e computadores) (COLLINS; BRAGA;

BONATO, 2006; HINSHAW,, 2004; FELIZZOLA; COSTA; VASCONCELOS, 2014; LANÇAS, 2017).

Um esquema básico de um sistema cromatográfico gasoso com seus principais componentes está ilustrado na Figura 10. Resumidamente, o processo de separação em um cromatógrafo a gás descrito na Fig. 10 ocorre da seguinte forma: um sistema de válvulas e controladores de fluxo instalados na linha de gás faz fluir constantemente um gás inerte ( $H_2$  ou Ar) através de todo “caminho” do sistema cromatográfico iniciando-se pela porta de injeção, coluna e por fim detector. Os controladores de fluxo mantêm uma taxa de fluxo sistematicamente controlada para que os tempos de retenção sejam reprodutíveis e os ruídos do detector sejam mínimos. A amostra é introduzida no cromatógrafo, normalmente com uso de uma micro-seringa na porta de injeção que contém um sistema de aquecimento programado. Após injetada, a amostra é volatilizada e transportada pelo gás de arraste para coluna, que contém uma fase estacionária adequada. No interior da coluna os componentes da amostra se distribuem de maneira distinta de acordo com suas interações relativas pelas fases móvel e estacionaria. Ao deixarem a coluna, o gás de arraste e os componentes da amostra passam por um detector onde serão identificados. No detector, a quantidade da amostra que chega é registrada e transformada em um sinal elétrico. Este sinal é então amplificado e transformado em um cromatograma pelo sistema de aquisição e integração dos dados.



**Figura 10** - Esquema dos principais componentes de um sistema cromatográfico a gás.

**Fonte:** Felizzola; Costa; Vasconcelos (2014).

Um esforço enorme é feito para obter análises extraordinariamente rápidas em CG, que geralmente exige uma instrumentação especial disponível somente em equipamentos mais modernos. Entretanto, dependendo da velocidade de separação pretendida, adaptações em equipamentos convencionais acompanhado por simples modificações em alguns parâmetros cromatográficos, podem garantir separações mais rápidas do que em CG convencional (MCNAIR; MILLER, 2009).

Desta forma, visando uma redução de duas a quatro vezes no tempo total necessário para realizar uma análise cromatográfica, o uso de colunas de dimensões reduzidas, assistido por pequenos ajustes na condição de análise de alguns parâmetros cromatográficos, por exemplo, aumento na taxa de aquecimento e otimização da velocidade do gás de arraste, podem satisfazer tal condição. Em contrapartida, a exigência de separações as quais proporcionam uma redução maior do que 10 vezes no tempo total de análise, geralmente necessitam de adaptações mais elaboradas nos cromatógrafos a gás convencionais (SEQUINEL *et al.*, 2010).

Para este segundo caso, será necessário adaptações instrumentais as quais satisfaça as seguintes características (MCNAIR; MILLER, 2009):

- aquecedores de coluna (em oposição aos fornos convencionais) que podem fornecer taxas de programação muito rápidas;
- injetores e divisores de entrada que facilitam pequenas amostras em zonas muito estreitas;
- sistema de detecção estável com tempo de resposta muito rápida;

Portanto, para atender essas características solicitadas em separações ultrarrápidas, amostradores automáticos de alta velocidade, sistemas para rápido aquecimento da coluna e detectores com altas taxas de aquisição de dados são itens obrigatórios que devem constar na instrumentação a ser utilizada (DORMAN *et al.*, 2008; EICEMAN *et al.*, 2004; HINSHAW, 2001).

➤ Amostradores e sistema de injeção de amostra

Embora a cromatografia, de modo geral, esteja sempre em constante aprimoramento instrumental, em se tratando de introdução de amostra em CG, não há diferenças operacionais expressivas, que não possam ser solucionadas, entre os tipos de amostradores e injetores já desenvolvidos em CG convencional que não



possam ser utilizados na CG-UR. Contudo, em CG-UR, há uma preferência pelos amostradores automáticos de alta velocidade, pois geralmente, diminuem o tempo necessário para completar um ciclo analítico, isto é, o tempo total entre uma injeção e a próxima, incluindo o resfriamento da coluna e tempo de amostragem (SEQUINEL *et al.*, 2010).

Da maioria dos injetores disponíveis empregados em CG convencional, praticamente todos podem ser usados em CG-UR sem grandes restrições. Desta forma, os mais frequentemente utilizados são *Split/Splitless*, *Programmable Temperature Vaporizer* – PTV (DAGAN; AMIRAV, 1998; DAVIS; MAKAROV; HUGHES, 1999; BICCHI *et al.*, 2004; THERMO SCIENTIFIC, 2007; KRKOSOVÁ *et al.*, 2008) e o *Cold On-Column* (POVOLO; PELIZZOLA; CONTARINI, 2008). Entretanto, para o sistema de injeção direto na coluna (*Cold On-Column*) há uma limitação que restringe o uso de colunas de analítica, pois devido ao diâmetro interno destas, faz-se necessário uma conexão por meio de uma pré-coluna com diâmetro interno maior (MATISOVÁ; DOMOTOROVÁ, 2003).

Em se tratando de amostradores, o sistema com amostrador em *headspace* é útil e sem grandes restrições em CG-UR. Por outro lado, amostradores automáticos convencionais, deverão ser adaptados para que possam injetar volumes em escala manométrica, visto que em CG-UR, são usadas colunas capilares com menor capacidade de processamento de amostra. Tal adaptação pode ser alcançada com a utilização de seringas com êmbolo na agulha (*plunger-in-needle syringe*) (LUONG, GRAS; JENNINGS, 2007).

### ➤ Colunas

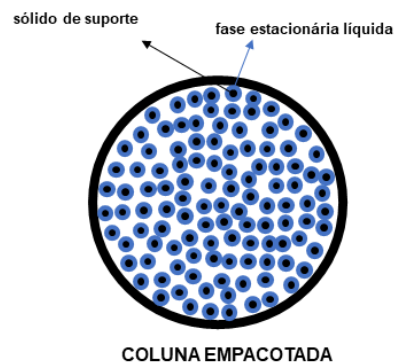
A coluna pode ser considerada o componente mais importante não só de um instrumento da CG, mas de toda técnica cromatográfica, pois é no seu interior, que contém uma fase estacionária, que de fato ocorre a separação dos componentes em uma mistura. Basicamente, as colunas podem ser classificadas em duas categorias mais amplas, as colunas empacotadas ou recheadas e as capilares (BRAITHWAITE; SMITH, 1999; COLLINS; BRAGA; BONATO, 2006; LANÇAS, 2017).

Dependendo da finalidade, as colunas empacotadas ainda podem ser divididas em dois tipos: as empregadas para escala preparativas e as analíticas. Como as dimensões da coluna estão diretamente relacionadas com sua eficiência, e,

consequente, sua finalidade de uso, em termos de dimensão, as colunas empacotadas analíticas exibem um diâmetro interno (DI) de 1 a 4 mm e comprimento de 1 a 3 m. Já as preparativas, possuem um DI de 5 a 100 mm o que lhes permitem maior capacidade de ingestão de amostra (COLLINS; BRAGA; BONATO, 2006).

A Figura 11 ilustra uma representação gráfica da seção transversal de uma típica coluna empacotada. As colunas capilares, também conhecidas como colunas tubulares abertas (*Open Tubular* - OT), oferecem uma série de vantagens em relação às colunas empacotadas, das quais destacam-se: separações com maior resolução, tempo de análise reduzido e normalmente maior sensibilidade (GROB, 2004).

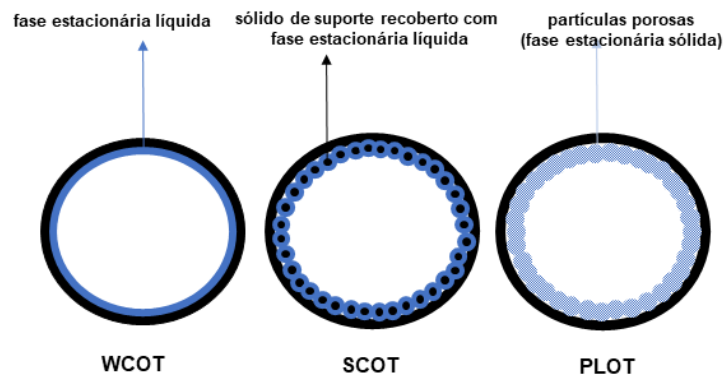
Como sua estrutura tubular é aberta, sua resistência ao fluxo do gás de arraste e pressão interna são menores do que as colunas empacotadas, portanto, permite separações muito mais eficiente de misturas de amostras complexas ou com ponto de ebulição alto sem que ocorra a degradação da coluna (BRAITHWAITE; SMITH, 1999; COLLINS; BRAGA; BONATO, 2006).



**Figura 11** - Representação esquemática de uma coluna empacotada.

**Fonte:** Mcnair; Miller (2009) – adaptado pela autora.

As Colunas OT são colunas cromatográficas simples com dimensões capilares (DI: 0,10-0,53 mm e comprimento: 10-100 m) que não são preenchidas totalmente, tal como as colunas empacotadas. Assim, as colunas OT capilares podem ser categorizadas de acordo com o método de revestimento interno de suas paredes com a fase estacionária. Basicamente, existem três tipos mais usados de colunas capilares, WCOT, SCOT e PLOT, como ilustra a Figura 12 (MCNAIR; MILLER, 2009).



**Figura 12** - Representação esquemática dos principais tipos de colunas capilares de tubo aberto.

**Fonte:** Mcnair; Miller (2009) – adaptado pela autora.

As colunas WCOT (*Wall -Coated Open Tubular Column -WCOT*), inventadas e patenteadas por Marchel Golay em 1858, possuem um filme fino (fase estacionária líquida) que reveste a parede interna, por isso são mais bem conhecidas como colunas capilares com parede recoberta (*Wall-Coated Open Tubular Column - WCOT*) (COLLINS; BRAGA; BONATO, 2006; MCNAIR; MILLER, 2009) Colunas WCOT de tamanhos iguais, mas com diâmetro interno maior, exibem tempos de eluição aproximadamente duas vezes mais rápidos do que colunas mais estreitas e com mesma espessura da camada de filme de fase estacionária devido à maior razão de fase (área de superfície por metro) das colunas de diâmetro largo (BRAITHWAITE; SMITH, 1999).

Dois outros tipos de coluna capilar são ilustrados na Fig. 12, a coluna tubular aberta com suporte recoberto (*Support-Coated Open Tubular Column – SCOT*) e a coluna tubular aberta com camada porosa (*Porous Layer Open Tubular Column – PLOT*).

As colunas SCOT consistem em tubos de vidro com diâmetro interno médio de 0,5 mm com tamanhos semelhantes das colunas WCOT. A parede interna é revestida com uma camada de material sólido de suporte especial fino, por exemplo, terra de diatomáceas, a qual é recoberta por um filme fino de fase estacionária (HORVATH; ETTRE; PURCELL, 1974). De modo geral, as colunas SCOT dispõem de uma área de superfície maior e um filme efetivo mais espesso de fase estacionária, conseqüentemente, uma maior capacidade de amostra do que as colunas WCOT, mas com menos eficiência geral (MCNAIR; MILLER, 2009). No entanto, encontram-

se disponíveis colunas de com diâmetro interno de 0,3 mm ( $\mu$ SCOT) com eficiências que se aproximam das colunas WCOT (BRAITHWAITE; SMITH, 1999).

As colunas PLOT assemelham-se bastante às colunas SCOT, com o diferencial de possuírem apenas um adsorvente sólido (micropartículas porosas de vidro, sílica, zircônia, etc.) como a própria fase estacionária. As colunas PLOT são adequadas para a análise de gases leves e outros compostos voláteis, tais como criptônio, néon, argônio, oxigênio, nitrogênio e xenônio WCOT (BRAITHWAITE; SMITH, 1999).

O primeiro passo a se considerar em um sistema cromatográfico rápido é o uso de colunas com dimensões reduzidas, revestida por uma fase móvel na forma de um filme fino, ou seja, as colunas OT (BICCHI *et al.*, 2004; CRAMERS; LECLERCQ, 1999; MASTOVSKA; LEHOTAY, 2003; MATISOVÁ; DOMOTOROVÁ, 2003; van DEURSEN *et al.*, 2000) Neste sentido, no âmbito da CG-UR, colunas capilares com comprimento entre 2 a 10 m, diâmetro interno variando de 0,05 a 0,1 mm e uma espessura a partir de 0,05  $\mu$ m para o filme, conforme mostrado na Tabela 4, são características, do ponto de vista da coluna, que atendem as condições para a realização de análises ultrarrápidas.

Deste modo, o ajuste destas características determinará o progresso da separação cromatográfica, uma vez que o diâmetro interno é diretamente relacionado a eficiência e a capacidade da coluna, ou seja, número de pratos de teóricos e quantidade de amostra armazenada pela coluna sem que haja distorções ou perda da resolução dos picos, respectivamente (SEQUINEL *et al.*, 2010).

Entretanto, um ajuste simultâneo na direção de máximo desempenho entre estes dois parâmetros (eficiência e capacidade da coluna), dependerá da necessidade e condições analíticas, pois a otimização de um acarreta no prejuízo do outro. Assim, surge como alternativa interessante para compensar a perda de resolução em decorrência da diminuição do comprimento da coluna, a utilização de colunas de diâmetro estreito (*narrow bore columns*), pois proporciona aumento considerável no número de pratos teóricos (CRAMERS; LECLERCQ, 1999; HINSAHW, 2001; Van SACKER, 1992).

Para se ter uma ideia, colunas WCOT de diâmetro estreito (*narrow bore WCOT columns*: d.i.  $\sim$ 0,25  $\mu$ m) exibem maior eficiência (até 6000 pratos/m), enquanto as colunas WCOT de diâmetro largo (*wide bore WCOT columns*: d.i.  $\sim$ 0,50 mm) possuem cerca de metade desta eficiência. Por outro lado, em termos da espessura da camada de filme fino de fase estacionária, as colunas WCOT de diâmetro largo proporcionam

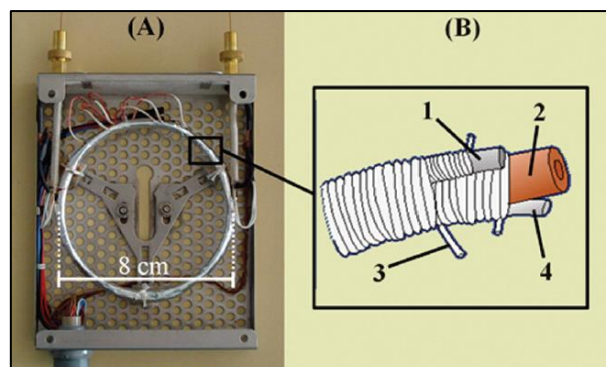
tempos de eluição aproximadamente duas vezes mais rápidos que colunas WCOT de diâmetro estreito, para uma mesma espessura de camada depositada de fase estacionária, devido à maior razão de fase (área de superfície por metro de comprimento) das colunas de diâmetro largo (BRAITHWAITE; SMITH, 1999).

➤ Sistema de aquecimento da coluna

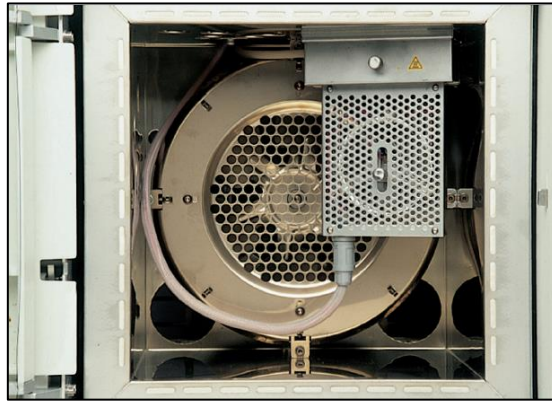
O controle da temperatura do forno em CG convencional deve ser controlado com precisão e exatidão, pois sua operação é crucial para manter a separação reproduzível pela coluna. A variação da temperatura afeta fortemente os fatores cinéticos e termodinâmicos relacionados a dinâmica de separação que ocorre entre fase móvel-analito-fase estacionária, e conseqüentemente altera também alguns parâmetros cromatográficos, tais como seletividade da fase estacionária e o espalhamento das zonas cromatográficas, que acarretarão em flutuações no desvio padrão dos tempos de retenção (PHILLIPS; JAIN, 1995; van DEURSEN *et al.*, 1999).

No regime das separações ultrarrápidas, a taxa de aquecimento exigida, como pode ser observada na Tabela 4, é de aproximadamente 60 °C ou mais. Portanto, para atender à exigência destas altas taxas de aquecimento em CG-UR, o modo de aquecimento por programação de temperatura deverá utilizar um sistema aquecimento/resfriamento que responda mais rapidamente e de forma precisa após a programação de temperatura (GROP, 2004).

Uma alternativa muito atraente a qual supre as condições de aquecimento em CG-UR está no módulo ultrarrápido (*Ultra Fast Module - UFM*) mostrado na Figura 13.



**Figura 13** - Dispositivo UFM. (A) Fotografia de um UFM da empresa Thermo Fischer Scientific Inc. ®. (B) Esquema dos componentes que integram o UFM: 1- elemento aquecedor, 2-conluna, 3-fibra cerâmica e 4-sensor de temperatura.



**Figura 14** - Fotografia de um disposto UFM instalado em forno de uma cromatografia a gás convencional.

**Fonte:** *Thermo Scientific* (2010).

Esse dispositivo, também conhecido como gaiola, é um tipo de sistema dedicado de aquecimento que integra uma coluna de diâmetro estreito (*narrow bore*), um elemento aquecedor e um sensor de temperatura, que estão revestidos por um material isolante (Teflon® ou fibra cerâmica) (BICCHI *et al.*, 2004; BICCHI *et al.*, 2005; THERMO SCIENTIFIC, 2007, 2010).

Outras vantagens do sistema de aquecimento dedicado (direto) da coluna proporcionado pela instalação do sistema UFM no forno (ver Figura 14) residem na ausência da necessidade do aquecimento do forno, do aproveitamento das interfaces de injetor e detector já instaladas no equipamento convencional e rápido resfriamento da coluna (EHRMANN *et al.*, 1996; BICCHI *et al.*, 2004; BICCHI *et al.*, 2005).

Comparado com o forno convencional CG com circulação de ar, o dispositivo UFM possui uma programação de temperatura mais rápida que permite a aplicação de taxas de aquecimento de até 1200 °C/min, rápido resfriamento da coluna (350 a 50 °C em torno de 1 min) e baixa potência consumo (200 W) (DALLÜGE *et al.*, 2002; SLOAN; MUSTACICH; ECKENRODE, 2001; van DEURSEN *et al.*, 1999; van DEURSEN *et al.*, 2001; THERMO SCIENTIFIC, 2010).

#### ➤ Sistema de detecção

O sistema de detecção em CG, de modo geral, fornece um sinal em resposta aos efluentes da coluna. Assim, um fluxo de quantidades finitas de compostos químicos na fase gasosa chega ao detector em bandas discretas. Estas bandas têm

um tempo de residência relativamente curto no detector, de modo que em alguns casos em menos de um segundo. Contudo, o detector deve responder rapidamente à presença dos componentes separados em fluxo, para depois retornar ao seu estado original e estar pronto para registrar o próximo composto eluído (BRAITHWAITE; SMITH, 1999; GROB, 2004).

Os detectores de modo geral, a não ser alguns específicos, não identificam de forma direta os componentes da amostra ao deixarem a coluna. A maioria dos disponíveis comercialmente é do tipo diferencial, ou seja, gera um sinal zero (linha de base) quando apenas o gás de arraste está circulando no sistema, de modo que quando um componente da amostra é eluído e detectado, o sinal produzido será proporcional à concentração ou massa desse componente. Por outro lado, os detectores de integração fornecem um sinal contínuo o qual é proporcional à quantidade de substâncias que foram eluídas, produzindo um sinal escalonado (BEVAN; THORBURN, 1965; SCOTT, 1996).

A maioria dos sistemas cromatográficos em fase gasosa utiliza detectores proporcionais ou diferenciais, dentre os quais destacam-se: detector de ionização de chama (*Flame Ionization Detector*- FID), o detector de condutividade térmica (*Thermal Conductivity Detector* - TCD) e o detector de captura de elétrons (*Electron Capture Detector* - ECD) (EICEMAN, 2006).

No entanto, há atualmente disponível no mercado e na comunidade científica mais de 60 diferentes tipos de detectores empregados em CG, assim como sistemas detecção próprios em CG que utilizam os fundamentos teórico-práticos e metodologia instrumental de técnicas espectroscópicas na forma acoplada a CG, tais como a espectrometria de massa, espectroscopia de infravermelho e espectroscopia de emissão atômica com plasma indutivamente acoplado (MCNAIR, H. M.; MILLER, 2009).

Com vimos até aqui em CG-UR, a velocidade requerida nas separações ultrarrápidas demanda um sistema cromatográfico capaz de fornecer condições mais dinâmicas desde o momento de introdução da amostra, passando pela coluna e por fim pelo detector. Neste sentido, de nada adianta possuir um equipamento com amostradores e injetores automáticos, colunas *narrow bore*, sistema de aquecimento dedicado (UFM) e condições otimizadas de análises, se o sistema de detecção não responder suficientemente rápido as espécies componentes da amostra que eluem rapidamente da coluna.

Para que o tempo de resposta do detector não comprometa a eficiência da separação da coluna em CG-UR, é necessário dispor no equipamento detectores que operem em alta frequência de amostragem (>200 Hz) para que seu sistema de aquisição de dados seja capaz de processar bandas cromatográficas com largura de 50 ms na linha de base, pois quanto maior a frequência de aquisição, mais rapidamente o detector responderá as alterações de sinal relativas as moléculas do analito ao eluirem da coluna (HINSHAW, 2003; KORYTÁR *et al.*, 2002).

Os detectores mais frequentemente empregados para comporem o sistema cromatográfico em CG-UR são, destacadamente, os de ionização em chama e espectrômetro de massa (EM), sendo que o FID é mais comumente encontrado nos modernos cromatógrafos os quais operam em frequência de aquisição de até 300 Hz (BICCHI *et al.*, 2004) (BICCHI *et al.*, 2005).

Para sistemas sem varredura total da amostra (*Non-Scanning*), a opção pelo uso de detectores de massa, como os espectrômetros de massas com analisadores por tempo de voo (EM-ATV), são notadamente mais vantajosos mesmo diante alguns analisadores quadrupolo (DALLÜGE *et al.*, 2002) (MASTOVSKA; LEHOTAY, 2003). Um sistema CG-UR acoplado ao detector EM-ATV é capaz de processar em torno de 500 espectros completos por segundo com uma exatidão na detecção de picos cuja largura está na ordem de milissegundos, sem provocar distorção dos dados espectrais (KORYTÁR *et al.*, 2002) (MASTOVSKA; LEHOTAY, 2003).

Por outro lado, essa elevada taxa de aquisição de dados característica do detector EM-ATV, vem acompanhada por uma queda da sensibilidade. Assim, visando minimizar esta condição, sistemas SMB (*Sonic Molecular Beam* – SMB) presentes nos espectrômetros de massas resolvem este problema de perda da sensibilidade dos sistemas ATV em decorrência das altas taxas de aquisição (SEQUINEL *et al.*, 2010). O desempenho em CG-UR melhora com emprego do sistema SMB, pois a intensidade do sinal do íon molecular é amplificada pela queda de energia térmica do sistema em decorrência da geração do feixe supersônico (DAGAN; AMIRAV, 1995; QUINTELLA, 1996; DAVIS; MAKAROV; HUGHES, ;1999).

A possibilidade do emprego de outros detectores, tais como o nitrogênio-fósforo, fotométrico de chama, fotoionização, captura de elétrons e condutividade térmica é limitada pela frequência de aquisição na qual operam e seus designs internos, os quais normalmente são células com volume grande, tornando-se, por isso, inviáveis para realização de separações ultrarrápidas (SEQUINEL *et al.*, 2010).



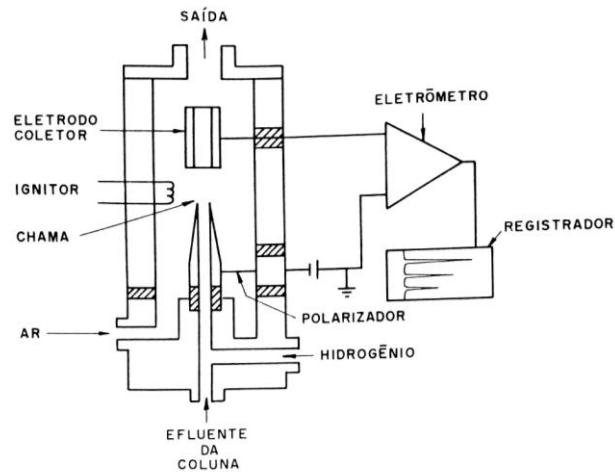
Apesar do grande número de detectores disponíveis em CG, de um de modo geral, o que é considerado como um dos pilares da técnica, de longe, é o FID (EICEMAN, 2006) (SCOTT, 1996), além de ser muito útil em CG-UR (SEQUINEL *et al.*, 2010).

O detector de ionização de chama foi introduzido em 1958 por Harley, Nel e Petrorius, McWilliam e Dewa, sendo desenvolvido especificamente para ser empregado na CG (HARLEY; NEL; PETRORIUS, 1958; McWILLIAM; DEWA, 1958).

Desde então, o FID tornou-se o detector mais usado para CG por apresentar características como: (i) responde a praticamente todos os compostos orgânicos com elevada sensibilidade ( $10^{-11}$  g/s); (ii) a resposta do detector não é afetada por mudanças ambientais (fluxo, pressão ou temperatura) modestas; (iii) sob operação normal não sofre grandes influências das impurezas comuns presentes no gás de arraste, tais como  $\text{CO}_2$  e água; (iv) ampla faixa de linearidade da ordem de  $10^7$  (BRAITHWAITE; SMITH, 1999; GROB, 2004; SCOTT, 1996).

O princípio de funcionamento do FID baseia-se na reação de combustão de uma mistura gasosa (gás de arraste + impurezas) por uma chama ( $\text{H}_2 + \text{Ar}$ ) que altera a condutividade elétrica de dessa mistura gasosa quando uma quantidade de partículas e/ou compostos (amostra volatilizada) estão presentes na combustão (COLLINS; BRAGA; BONATO, 2006; LANÇAS, 2017).

Um design típico do FID é ilustrado na Figura 15. O efluente da coluna é misturado com hidrogênio e levado a uma pequena ponta do queimador que é circundada por um alto fluxo de ar para suportar a combustão. Um dispositivo de ignição é fornecido para “acender” remotamente a chama. Um eletrodo coletor é polarizado em torno de +300 V em relação à ponta da chama. Assim, à medida que o gás de arraste e demais moléculas de impurezas presentes nessa mistura são queimados na chama, gera-se uma quantidade muito pequena de íons que são coletados e formam uma pequena corrente ( $\sim 10^{-14}$  A) a qual é registrada como linha de base no cromatograma (LANÇAS, 2017).

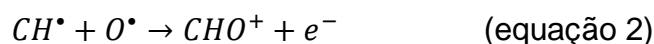


**Figura 15-** Representação esquemática de um detector FID.

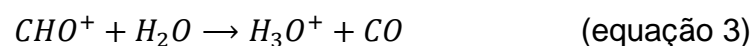
**Fonte:** Collins; Braga; Bonato (2006).

Entretanto, se compostos orgânicos voláteis são eluídos da coluna e entram na chama, uma série de reações complexas ocorrem (fragmentação térmica, quimioionização, reações de radicais livres, etc.) para produzir espécies carregadas,  $\text{CO}_2$  e  $\text{H}_2\text{O}$ . Uma chama de hidrogênio puro contém radicais ( $\text{H}^\bullet$ ,  $\text{O}^\bullet$ ,  $\text{OH}^\bullet$  e  $\text{HO}^\bullet$ ) e espécies excitadas, mas nenhum íon. Assim, quando moléculas orgânicas são transportadas até a chama, a ionização ocorre devido a alta reatividade destas espécies, assim como de maneira direta via ionização térmica, sendo que a quantidade de espécies ionizadas é proporcional ao número de átomos de carbono presentes e, portanto, ao número de moléculas (BRAITHWAITE; SMITH, 1999; GROB, 2004; MCNAIR; MILLER, 2009).

Os principais processos de ionização envolvem inicialmente a formação de espécies  $\text{CH}^\bullet$  a partir das moléculas orgânicas que reagem imediatamente com os radicais de oxigênio presente no ar (comburente), a partir da reação:



As espécies  $\text{CHO}^+$  reagem rapidamente com a água produzida na chama para gerar íons hidrônio e monóxido de carbono (GROB, 2004).



O processo de ionização descrito pelas equações 2 e 3 ocorre aproximadamente uma vez a cada  $10^5$  átomos de carbono introduzidos na chama. Portanto, a resposta do FID essencialmente é proporcional ao número de átomos de carbono, em vez da massa ou número de moléculas (GROB, 2014).

Por fim, a corrente observada devida à maior concentração das espécies carregadas presentes na chama em decorrência da eluição de compostos orgânicos da amostra é da ordem de  $10^{-8}$  a  $10^{-12}$  A, a qual será proporcional ao número de átomos de carbono na chama. Dessa forma, este aumento de espécies carregadas ocasionará uma perturbação na corrente e, conseqüentemente, no sinal base por esta gerado relacionado apenas a passagem do gás de arraste. Assim, esta perturbação produzirá um sinal que será amplificado e por um eletrômetro com alta impedância de entrada para produzir um sinal de saída tipicamente em uma faixa de 0 a 10 mV ou 0 a 1 V, permitindo que um registrador gere um cromatograma (BRAITHWAITE; SMITH, 1999; COLLINS; BRAGA; BONATO, 2006; LANÇAS, 2017; MCNAIR; MILLER, 2009).

O FID responde a todos os compostos orgânicos que queimam na chama ( $H_2$  + Ar), pois o sinal é aproximadamente proporcional a quantidade de carbono. Desse modo, para estimar a resposta relativa de qualquer composto, o número efetivo de carbono (*Effective Carbon Number* – ECN) foi introduzido o FID (SIMON, 1985; SCANLON; WILLIS, 1985). Esse número explica o fato de que a resposta do FID diminui na presença de certos heteroátomos, tais como oxigênio ou nitrogênio. O cálculo do ECN é feito atribuindo valores aos grupos de átomos específicos, em relação à resposta de um material de referência, para o qual o ECN é simplesmente o número de carbonos na molécula, que geralmente uma n-parafina (SCANLON; WILLIS, 1985).

Por outro lado, compostos que não contêm carbono orgânico não sofrem combustão na chama, e por isso não são detectados pelo FID. Neste sentido, a água, produto da combustão de orgânicos na chama, assim como inorgânicos ( $H_2$ ,  $O_2$ ,  $N_2$ ,  $SiCl_4$ ,  $SiF_4$ ,  $H_2S$ ,  $SO_2$ ,  $COS$ ,  $CS_2$ ,  $NH_3$ ,  $NO$ ,  $NO_2$ ,  $N_2O$ ,  $CO$ ,  $CO_2$ , Ar, Kr, Ne, Xe,) de modo geral, não são detectados pelo FID. Portanto, a ausência de um pico relativo a presença de moléculas de água permite que seja usado nas análises por FID, amostras que contenham água em suas composições, uma vez que não interfere no cromatograma (COLLINS; BRAGA; BONATO, 2006; LANÇAS, 2017; MCNAIR; MILLER, 2009).

O FID não é indicado em escala preparativa por ser um detector destrutivo e limitado a análise da maioria dos inorgânicos mesmo que voláteis. Por outro lado, devido à sua excelente seletividade por compostos orgânicos, é capaz de identificar traços da ordem de  $10^9$  g ou menos, apresentando de 100 a 1000 vezes poder de detecção em comparação ao detector de condutividade térmica (LANÇAS, 2017).

### 2.2.3 Validação de métodos cromatográficos

Métodos cromatográficos são comumente utilizados para a análise quantitativa e qualitativa, em escala preparativa ou analítica, em vários setores da indústria (farmacêutica, petroquímica, agroquímica, alimentícia, etc.), em órgãos e agências de acreditação (ANVISA, INMETRO, IUPAC, ISO, ICH, EPA, etc.) e no âmbito acadêmico-científico. Embora as finalidades entres estes entes sejam diferentes, o que se têm em comum é o fato de que os métodos cromatográficos por eles desenvolvidos e/ou empregados sejam reproduzíveis quando usados por outros analistas, em outros equipamentos equivalentes e em dias ou locais distintos (RIBANI *et al.*, 2004).

Para garantir que um novo método analítico gere informações confiáveis e interpretáveis sobre a amostra em análise, este deve ser submetido à uma avaliação denominada validação, o qual contém uma série de parâmetros associados ao desempenho analítico do método proposto (ANVISA, 2017).

A validação de um método analítico é o processo pelo qual se estabelece, através da realização de estudos de laboratório, se as características de desempenho do método atendem aos requisitos para as aplicações analíticas pretendidas. Portanto, a validação é um passo importante na determinação da confiabilidade e reprodutibilidade do método, pois é capaz de confirmar que o método pretendido é adequado para ser conduzido em um sistema particular (LEITE, 2008).

Os parâmetros de validação aplicados aos métodos analíticos de modo geral têm sido definidos por diversas instituições nacionais e internacionais, públicas e privadas. Assim, algumas definições e terminologias para estes parâmetros possuem significados ligeiramente distintos em decorrência da especificidade e finalidade destas organizações. No entanto, algumas terminologias, mesmo que distintas, possuem significados equivalentes para o avaliar o desempenho do método proposto. Portanto, diante disso, os parâmetros analíticos frequentemente usados para

validação de métodos de separação, tal como os cromatográficos, são: seletividade, linearidade e alcance, precisão, limite de detecção, limite de quantificação e robustez (GROB, 2004; RIBANI *et al.*, 2004; INMETRO, 2018).

➤ Linearidade

Os métodos são descritos como lineares quando há uma relação diretamente proporcional entre a resposta do sinal medido (área ou altura do pico) e a concentração ou massa do analito na amostra, na faixa de concentrações de analito de interesse. Essa correlação entre sinal medido e concentração ou massa do analito a ser determinada, é obtida empiricamente por algum tipo de transformação ou modelo matemático a partir de medidas realizadas para valores conhecidos de massa ou concentração do analito de interesse, ou seja, a partir da curva analítica (GROB, 2004).

O ajuste geralmente utilizado para a determinação da linearidade baseia-se no método dos Mínimos Quadrados Ordinários, ou simplesmente na curva de regressão linear, a partir das respostas instrumentais medidas (ou área de pico ou altura) como uma função do aumento da concentração do analito (BAUMANN; WAETZIG; 1997; ERMER; LANDY, 2007).

Orienta-se que seja tomando pelo menos cinco concentrações diferentes na curva, de modo que pontos adicionais de concentração tendem a ser incluídos na região de menor da concentração. A avaliação da linearidade pelo modelo de regressão linear é representada pelo coeficiente de correlação ( $r$ ) que define o quão próximo uma determinada função se encaixa em um conjunto particular de dados experimentais. Assim, valores de  $r$  próximos de 0,999 são pretendidos e considerados aceitáveis (GROB, 2004).

➤ Seletividade e especificidade

Um método é considerado "específico" se fornecer uma resposta para apenas um único analito, enquanto o termo "seletivo" significa que o método fornece respostas para um número de espécies químicas que podem ser distinguidas umas das outras. A seletividade também implica a capacidade de separar o analito de produtos de

degradação, metabólitos e drogas coadministradas (BRESSOLE; BROMET-PETIT; AUDRAN, 1996)

Entretanto, os termos especificidade e seletividade são frequentemente usados de forma intercambiável, gerando confusão desnecessária. Portanto, a IUPAC recomenda o uso apenas do termo seletividade (RIBANI *et al.*, 2004).

Em métodos cromatográficos, a seletividade pode ser comprovada pela existência de uma boa separação entre os componentes do analito e os outros componentes de eluição relacionados à matriz, impurezas e produtos de degradação da coluna). Para tanto, considera-se um método cromatográfico seletivo aquele que apresenta uma resolução do analito, a partir dos outros componentes, superior a 1,5-2,0 (YUWONO; INDRAYANTO, 2005)

#### ➤ Exatidão

A exatidão de um método analítico é dada pela concordância na qual um o valor medido se aproxima do valor referência (verdadeiro). Dependendo da finalidade da análise e dos órgãos que regulamentam o laboratório responsável pela análise, haverá uma abordagem adequada para o cálculo da exatidão do método (INMETRO, 2018). Neste sentido, a ANVISA e a ICH (*International Conference on Harmonization*) recomenda a realização de no mínimo nove determinações envolvendo um mínimo de três diferentes níveis de concentração, ou seja, ensaios em triplicata para três níveis de concentração (baixa, média e alta). Para estes órgãos, a exatidão é expressa como a diferença porcentual entre as médias e o valor considerado verdadeiro (RIBANI *et al.*, 2004). Assim, o cálculo da exatidão pode ser obtido a partir da seguinte equação:

$$Exatidão = \frac{C_m}{C_t} \times 100 \quad (\text{equação 4})$$

onde  $C_m$  é o valor médio da concentração obtido experimentalmente e a  $C_t$  representa a concentração teórica (INMETRO, 2018).

➤ Precisão

Representa o grau de concordância entre uma série de medidas independentes obtidas repetidamente de uma mesma amostra ou amostras similares sob condições especificadas (INMETRO, 2018). Em validação de métodos, por dispor de um número de medições pequeno (< 20), a avaliação da precisão é calculada a partir da estimativa do desvio padrão relativo ( $s$ ) (equação 5) e não pelo desvio padrão absoluto ( $\sigma$ ), uma vez que este último requer um número significativo de determinações.

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} \quad (\text{equação 5})$$

onde  $\bar{x}$  representa a média para um pequeno número de medições,  $x_i$  é o valor individual de uma medição e  $n$  corresponde ao número de medições (RIBANI *et al.*, 2004).

Uma segunda maneira de se calcular a exatidão de um método é por meio da estimativa do desvio padrão relativo (RSD – *Relative Standard Deviation*) (equação 6) ou também conhecido coeficiente de variação (CV).

$$RSD (\%) = \frac{s}{\bar{x}} \times 100 \quad (\text{equação 6})$$

De acordo com a complexidade da amostra, considera-se aceito valores de RSD de até 20%. Porém, usualmente, métodos propostos para quantificar compostos em macro quantidade exige-se um RSD em torno de 1 a 2% (HUBER, 1998).

Portanto, a finalidade da precisão consiste em garantir que as medidas não foram efetuadas aleatoriamente e que podem ser repetidas e reproduzidas (LEITE, 2008). Logo, em validação de métodos, a precisão pode ser dividida em três categorias: repetibilidade, precisão intermediária e reprodutibilidade (RIBANI *et al.*, 2004; GROB, 2004).

○ Repetibilidade

A repetibilidade, também conhecida como precisão intralaboratorial, expressa a precisão obtida a partir de sucessivas medidas de um mesmo método sob mesmas

condições de operação (mesmo procedimento, mesmo analista, mesmo aparato instrumental usado sob as mesmas condições e mesmo local) durante um curto período de tempo (LEITE, 2008).

Para cálculo da precisão expressa em termos da repetitividade, o INMETRO estabelece sete ou mais repetições para determinar a estimativa do desvio padrão, enquanto a ANVISA e ICH recomenda um mínimo de nove repetições, considerando o intervalo linear do método, isto é, três concentrações em triplicata, ou um mínimo de seis determinações a uma concentração próxima do valor referência (RIBANI *et al.*, 2004).

- Precisão intermediária

Expressa a variação obtida quando a análise é realizada em um mesmo laboratório por distintos analistas, usando equipamentos, reagentes e colunas diferentes durante dias aleatórios. A finalidade da precisão intermediária consiste em apurar sistematicamente se em um mesmo laboratório o método proposto produz os mesmos resultados (INMETRO, 2018).

Em termos práticos, número de ensaios necessários para se avaliar a precisão intermediária segue a mesma recomendação estabelecida pela ANVISA e ICH para a repetitividade (RIBANI *et al.*, 2004).

- Reprodutibilidade

A reprodutibilidade, conhecida com análise interlaboratorial, representa a precisão obtida a partir de resultados medidos em laboratórios diferentes, para um mesmo procedimento, cuja finalidade é verificar se o método proposto reproduz os mesmos resultados em diferentes instalações. Este tipo de precisão é empregado em regime de colaboração entre laboratórios para padronização de procedimentos analíticos (INMETRO, 2018).

A IUPAC estabelece um número igual a oito ensaios realizados por diferentes laboratórios para que se possa considerar um método reprodutível (RIBANI *et al.*, 2004).

Os estudos que estabelecem normas para avaliar a precisão de um método analítico em termos da reprodutibilidade utilizam ferramentas estatísticas para



observar a dispersão dos valores obtidos entre os diferentes laboratórios. Neste sentido, a partir de uma relação proposta por Horwitz e colaboradores, denominada de curva de Horwitz (equação 7), é possível expressar uma relação que interprete a dispersividade de valores em um ensaio interlaboratorial por meio de uma dependência entre o desvio padrão relativo e a concentração.

$$RSD (\%) = 2^{(1-0.5\log C)} \quad (\text{equação 7})$$

onde  $C$  é dada pela relação de gramas do constituinte por grama da amostra. Segundo Horwitz, é considerado um nível tolerável valores até 5% no desvio padrão relativo (LEITE, 2008).

➤ Limites e detectabilidade

A maneira com que uma técnica responde a presença de uma quantidade mínima de uma espécie componente ou analito presente em uma amostra é definida de detectabilidade. Neste sentido, deriva-se desta definição o conceito de limite, o qual também está associado ao modo de representar certa quantidade de amostra em condições instrumentais estabelecidas. O limite pode ser categorizado em três tipos: instrumental, detecção e quantificação (GROB, 2004; LEITE, 2008).

○ Limite de detecção (LD)

O limite de detecção representa a menor concentração de um analito que pode ser detectado, mas não determinado quantitativamente, dentro de um nível de confiabilidade de precisão aceitável sob condições analíticas estabelecidas. A IUPAC recomenda um nível de confiabilidade de 95% para a estimativa do LD (RIBANI *et al.*, 2005)

A maneira mais adotada para a estimativa do LD consiste no método baseado em parâmetros da curva analítica, em que o LD corresponde a concentração que produziria um valor sinal de três vezes maior que o ruído médio da linha de base do sistema analítico em estabilidade, medido a partir do branco (matriz isenta do composto de interesse). Na prática o LD pode ser obtido a partir da seguinte expressão:

$$LD = 3X \frac{D}{S} \quad (\text{equação 8})$$

onde  $s$  é a estimativa do desvio padrão do branco, ou a estimativa do desvio padrão do coeficiente linear da reta ajustada pela regressão, e  $S$  representa o coeficiente angular (*slope*) da reta da curva analítica (calibração) (INMETRO, 2018).

- Limite de quantificação (LQ)

O limite de quantificação é definido como a menor concentração que pode ser medida com acurácia e precisão aceitáveis sob condições analíticas estabelecidas (LEITE, 2008).

Assim como para o LD, a estimativa do LQ estatisticamente mais confiável consiste no método baseado nos parâmetros da curva analítica, uma vez que corrige alguns problemas relacionados as alturas e área dos picos presente no método sinal-ruído, pois para concentrações muito baixas (traços), para que o analito seja quantificado com precisão aceitável, exige-se alturas de pico de 10 a 20 vezes maior do que o ruído da linha de base (GROB, 2004).

Portanto, considerando o método dos parâmetros da curva analítica, o LQ pode ser calculado a partir da seguinte relação:

$$LQ = 10X \frac{S}{S} \quad (\text{equação 9})$$

Embora esse método forneça melhores resultados ao nível de traços, em altas concentrações, geralmente estima valores de LQ acima dos reais (INMETRO, 2018).

- Intervalo

O intervalo de um método analítico compreende o intervalo entre os limites de quantificação superior e inferior obtidos com precisão, exatidão e linearidade usando o método analítico. Contudo, o intervalo aceitável será definido como o intervalo de

concentração sobre o qual a linearidade, exatidão e precisão são aceitáveis (GROB, 2004).

➤ Robustez

A robustez pode ser compreendida como a capacidade de um método analítico em não ter sua resposta ou sinal afetado pela presença de pequenas variações nos seus parâmetros. Em CG tais parâmetros normalmente são: temperatura e volume de injeção, temperatura da coluna, fluxo do gás de arraste, temperatura do detector, etc. Assim, ao detectar quais variáveis influenciam, sob pequenas variações, fortemente a resposta do método, estes deverão ser controlados dentro de limites estabelecidos a fim de evitar prejuízos quando o método for transferido, ou seja, quando o método for submetido a ensaios conduzidos por diferentes analistas, equipamentos e/ou laboratórios. Para tanto, seleciona-se um número de variáveis as quais serão investigadas dentro de um intervalo o qual ultrapasse ligeiramente a variação esperada quando o método é transferido (VANDER-HEYDEN; NIJHUIS; SMEYERS-VERBEKE, 2001; DEJAEGHER; VANDER-HEYDEN, 2007).

Portanto, os testes de robustez têm grande utilidade para identificar e controlar possíveis fontes de variabilidade em uma ou em várias respostas, por exemplo, de um método analítico em desenvolvimento.

Em cromatografia, a utilização de planejamento experimental tem sido muito útil na identificação de tais variáveis, em planejamento experimental denominadas de fatores, a partir da execução de testes de robustez. Geralmente, o planejamento experimental em CG é empregado a fim de otimizar as condições de análise verificando se as potenciais fontes de variabilidade (os fatores), as quais podem sofrer alterações, são contínuas (fatores quantitativos) ou discretas (fatores qualitativos), bem como uma mistura destas (DEJAEGHER; VANDER-HEYDEN, 2007).

Existem diversas metodologias de planejamento experimental, de modo que a escolha de um método adequado dependerá essencialmente do número de fatores ou versões (se as variáveis são quantitativas (fatores) ou qualitativas (versões)) e níveis (valores que os fatores/valores podem assumir conforme a metodologia de planejamento aplicada) a serem avaliados. Normalmente, os planejamentos são realizados assumindo dois níveis os quais são capazes de avaliar um número razoavelmente grande de fatores, considerando um número reduzido de experimentos

(BARROS NETO; SCARMINIO; BRUNS, 1995; VANDER-HEYDEN, NIJHUIS e SMEYERS-VERBEKE, 2001).

A escolha do método que executará o planejamento não deve ser feita ao acaso, pois a atividade estatística mais importante não é a própria análise de dados desempenhada pelo método escolhido, mas sim o planejamento experimental no qual os dados devem ser organizados. Uma vez que isso não é realizado, o resultado do planejamento acarretará apenas em um amontoado de dados improdutíveis sendo impossível extrair quaisquer conclusões (BARROS NETO; SCARMINIO; BRUNS, 1995).

Na próxima seção, quimiometria, voltaremos a dissertar a respeito de planejamento experimental com mais detalhes.

### **2.3 Quimiometria**

O aprimoramento das técnicas instrumentais de análise química, sobretudo, a partir da inovação de modernos sistemas de aquisição de dados processados por microcomputadores, permitiu um nível de sofisticação instrumental capaz de extrair quantidades cada vez maiores de informação analítica. No entanto, esta enorme quantidade de dados muitas vezes torna-se laborioso encontrar alguma relação entre as variáveis em análise, exigindo-se, assim, tratamentos de dados mais complexos do ponto de vista matemático e estatístico para que seja possível fornecer uma relação coerente entre os sinais obtidos com os resultados desejados (FERREIRA, 2015; KRAMER, 1998; MALINOWSKI, 2002; SOUZA; POPPI, 2012).

Desta forma, a disponibilidade de enormes conjuntos de dados produzidos por cada técnica representa ao mesmo tempo uma grande oportunidade e um desafio para o pesquisador moderno. A oportunidade está no enorme conteúdo de informações “implícitas” nos dados coletados, enquanto o desafio está na extração de informações significativas para alcançar o objetivo desejado (FERREIRA *et al.*, 1999; FERREIRA, 2015).

Para dar conta e sentido ao enorme conjunto de dados gerados, paralelamente ao intenso avanço tecnológico das técnicas analíticas, a evolução na engenharia de *software* e *hardware* tornou possível o desenvolvimento de novos e mais eficientes métodos de tratamento de dados baseados em abordagens teórico-computacional, pois a capacidade de processamento destes novos computadores atingiram níveis

jamais explorados nas abordagens de análise de dados experimentais na área da química (KRAMER, 1998; MALINOWSKI, 2002; DASZYKOWSKI; WALCZAK, 2006).

Portanto, diante deste cenário de retumbante aperfeiçoamento tecnológico tanto por parte das sofisticadas técnicas analíticas quanto das avançadas ferramentas teórico-computacionais, foi que surgiu a Quimiometria atuando como “ferramenta útil” no tratamento de grandes conjuntos de dados experimentais, por meio de seus inúmeros métodos de organização, planejamento e análise de dados fundamentados em princípios matemáticos e estatísticos (MALINOWSKI, 2002; DASZYKOWSKI; WALCZAK, 2006; SOUZA; POPPI, 2012).

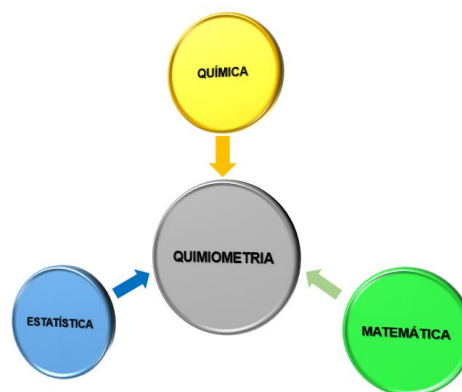
Vale ressaltar que o surgimento, de fato, da quimiometria ocorreu no início dos anos 70, mais precisamente em 1971 pelo químico orgânico Svante Wold, porém, sua utilidade como metodologia capaz de analisar de maneira útil e confiável grandes conjuntos de dados, só foi consolidada após o expressivo aperfeiçoamento da informática e, concomitantemente, à medida que novos métodos teórico-computacionais baseados em conceitos matemáticos e estatísticos estavam disponíveis (BARROS NETO; SCARMINIO; BRUNS, 2006; FERREIRA, 2015).

Cabe ressaltar que não é exclusividade ou tão somente um problema da química, estar diante de situações que produzem um grande número de dados, mas sim de todas as áreas científicas e técnicas as quais manipulam medidas e também utilizam conceitos matemáticos e estatísticos para desenvolverem seus próprios métodos de organização e análise dos dados. Desta forma, apesar de que a fundamentação teórica da quimiometria e de todas as outras áreas que aplicam seus respectivos métodos de tratamento de dados, estar baseado em uma relação interdisciplinar com a matemática e estatística, cada uma destas áreas, inclusive a quimiometria, possuem especificidades que geralmente não farão muito sentido se aplicadas em outra área (FERREIRA, 2015).

Em consequência desta natureza interdisciplinar do conceito de quimiometria a qual foi sendo estabelecido desde a ideia inicial proposta por Wold em 1971, Chertién (2003) formulou uma definição moderna que se aproxima mais das técnicas de análise química, a qual a considera como uma ciência que gerencia e processa informações necessariamente de natureza química. Atualmente, há uma concepção bem aceita pela comunidade científica em geral de que a quimiometria contempla conceitos de três grandes ciências exatas: Química, Matemática e Estatística. Neste sentido, a

Figura 16 ilustra uma representação esquemática que abrange esta ideia (FERREIRA, 2015).

Portanto, um entendimento global sobre quimiometria nos conduz a uma definição a qual implica que esta seja uma disciplina da química que utiliza métodos matemáticos e estatísticos para organizar de forma otimizada experimentos, com o objetivo de extrair o máximo de informação química relevante a partir da análise de conjunto de dados disponíveis (FERREIRA *et al.*, 1999).



**Figura 16** - Relação entre as disciplinas que definem a quimiometria.

**Fonte:** Elaborado pela autora.

Segundo definição da IUPAC, quimiometria consiste no emprego da estatística para analisar os dados de natureza química coletados a partir de uma técnica instrumental química qualquer e o planejamento de experimentos químicos e simulações (IUPAC, 2014).

Por muitos anos que precederam a consolidação, de fato, da quimiometria, experimentos cromatográficos eram realizados a partir da abordagem univariada de análise de dados, a qual investiga apenas um fator (variável medida experimentalmente) de maneira sistemática para um conjunto de amostras. Neste tipo de abordagem, todos os outros fatores são mantidos em um nível constante (DECLERCK, *et al.* 2018).

Embora, a análise univariada seja há tempos muito empregada na resolução de vários problemas químicos, sua utilização tornou-se limitada para alguns casos de natureza complexa em que se deseja obter uma correlação entre um grande número de variáveis para um elevado conjunto de dados, cujo resultado aponte para uma

solução que não seja puramente matemática e/ou estatística sem uma aplicação prática (FERREIRA *et al.*, 1999; FERREIRA, 2015).

Por outro lado, os métodos quimiométricos aplicam uma abordagem multivariada baseada em dois pontos centrais: o planejamento experimental, em que se elege as condições experimentais ótimas para uma determinada investigação; e a análise de dados multivariados, que extrai destes o máximo possível de informação química relevante (MALINOWSKI, 2002; RENCHER; 2002; FERREIRA, 2015). Desta forma, problemas químicos complexos podem ser resolvidos por meio da análise multivariada, pois esta, considera uma distribuição de múltiplas variáveis de maneira simultânea o que torna possível a obtenção de uma quantidade maior e mais significativa de informação do que a análise univariada, visto que esta, por analisar individualmente cada variável, negligencia a correlação entre elas (BEEBE; PELL; SEASHOLTZ, 1998; GEMPERLINE; 2006).

Apesar de que os métodos quimiométricos sejam considerados gerais, isto é, podem ser aplicados a qualquer tipo de experiência analítica da química, bem como a qualquer tipo de sinal instrumental, existem problemas associados a tipos específicos de sinais instrumentais ou a técnicas analíticas específicas que necessitam de tratamento especial. Logo, somente após o conhecimento do sistema a ser estudado é que métodos quimiométricos bem adaptados poderão ser escolhidos para efetuarem a análise (DASZYKOWSKI; WALCZAK, 2006).

Diante de todo exposto acima sobre a quimiometria, fazemos o seguinte questionamento: qual(ais) o(s) método(s) quimiométrico(s) específico(s) para cromatografia. Pois, ainda que algumas abordagens quimiométricas aplicadas em determinada técnica e/ou experimento seja inapropriado, a aplicação em cromatografia são inúmeras, visto que, basta mencionar os incontáveis métodos de compactação, visualização, calibração e classificação de dados, deram suporte com êxito para as técnicas cromatográficas por muitos anos (DASZYKOWSKI; WALCZAK, 2006).

Não podemos esquecer que a quimiometria é relativamente nova e com grande potencial de ampliação e desenvolvimento, de modo que muitos métodos que fazem parte de seu arsenal atual compõem três principais áreas de atuação: planejamento de experimentos, reconhecimentos de padrões e calibração multivariada (SOUZA; POPPI, 2012). Nas próximas seções, atentaremos para todos esses métodos acima

citados, na ordem em que foram utilizados neste trabalho: planejamento de experimentos, reconhecimento de padrões e método de calibração multivariada.

Por fim, antes de apresentar os conceitos de cada método quimiométrico, reportaremos algumas etapas que precedem o tratamento dos dados que são cruciais para a obtenção de resultados confiáveis, isto é, a etapa preparação ou processamento prévio dos dados (organização e pré-tratamento) e também a etapa de planejamento de experimentos. Segundo Ferreira (2015, p. 29) para que os dados coletados apresentem utilidade, a maneira com a qual foram adquiridos deve ser adequada para evitar o desperdício de tempo e recursos, uma vez que um número elevado de dados coletados aleatoriamente não garante a qualidade do tratamento, tampouco, resultados mais eficientes.

### 2.3.1 Processamento prévio dos dados: organização dos dados

A função da etapa de organização é de fundamental importância em qualquer tipo de análise, pois ao ser realizada torna-se possível obter, previamente, uma noção da qualidade dos dados para que, então, se avalie a necessidade de efetuar algum tratamento ou transformação dos dados e partir destes, avaliar qual método será mais apropriado para realizar a análise de fato dos dados (MOITA NETO; MOITA, 1998; FERREIRA, 2015).

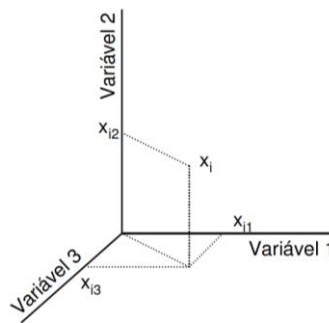
Ao medir um sinal, registramos a magnitude da resposta do equipamento em função de uma variável correspondente às características instrumentais da técnica utilizada, ou seja, em função da frequência, comprimento de onda, número de onda, energia de ligação, tempo, etc. Na cromatografia, o sinal do detector fornece uma resposta no domínio do tempo cuja variável em questão é o tempo de retenção o qual possui uma relação indireta a respeito de uma propriedade ou informação química do sistema em estudo, por exemplo, quantificação de uma determinada espécie componente na amostra (FERREIRA, 2015).

Portanto, após coletados os dados, estes serão organizados em um arranjo matricial representado pela matriz  $X$ , formada por  $m \times n$  elementos, onde as linhas correspondem as  $n$  amostras e as colunas aos  $m$  valores das medidas (variáveis) como mostra a equação 10 (MOITA NETO; MOITA, 1998).



$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n1} & \dots & X_{nm} \end{bmatrix} \quad (\text{equação 10})$$

Cada amostra na matriz  $\mathbf{X}$  corresponde a um vetor-linha, analogamente, cada variável a um vetor-coluna, de modo que é possível construir uma representação gráfica da matriz  $\mathbf{X}$ , onde cada vetor linha representa um ponto no espaço  $m$ -dimensional, ou seja, os eixos das coordenadas são descritos pelas  $m$ -variáveis, como ilustrado na Figura 17 (PARREIRA, 2003).



**Figura 17** - Representação vetor-linha da amostra  $x$  nos eixos das variáveis: 1, 2, e 3;

**Fonte:** Parreira (2003).

Embora o arranjo como são dispostos os dados na forma de gráfico permita obter informações mais eficientes e de uma maneira mais evidente do que na forma da matriz  $\mathbf{X}$ , a possibilidade de associar apenas três variáveis na construção gráfica torna-se um fator limitante, pois não é possível analisar simultaneamente a influência de todas as variáveis contidas no sistema. Entretanto, métodos estatísticos de análise multivariada são úteis para solucionar esta limitação (PARREIRA, 2003).

Em uma análise multivariada realizada da maneira correta, a organização dos dados deve ser avaliada a partir da investigação do método de medida utilizado, o modo como foram registrados, a confiabilidade dos dados e a representatividade das amostras disponíveis para análise. Desta forma, o conjunto de dados a ser analisado deve conter informações úteis relativas a um grupo de amostras suficientemente representativas, para que se possua informações adequadas e relevantes as quais serão utilizadas no tratamento de dados o qual se pretende aplicar (FERREIRA, *et al.* 1999).

Não menos importante do que a avaliação da maneira como se organiza os dados para análise, o pré-tratamento é uma etapa prévia à análise quimiométrica que visa minimizar as variações indesejáveis que podem influenciar os resultados finais, quando estas não foram removidas durante a aquisição de dados e/ou nem extintas naturalmente no decorrer da análise (FERREIRA, 2015). Neste sentido, o pré-tratamento dos dados é realizado por um conjunto de procedimentos baseado em manipulações matemáticas classificado em dois tipos principais: transformação, quando aplicados às amostras (linhas da matriz  $\mathbf{X}$ ) e pré-processamento, quando aplicados às variáveis (colunas da matriz  $\mathbf{X}$ ) (FERREIRA *et al.*, 1999).

### 2.3.1.1 Processamento prévio dos dados: transformação

A etapa de transformação dos dados utiliza de recursos matemáticos para remover fontes de variações aleatórias (ruídos experimentais) e informações sistemáticas indesejáveis (desvios da linha e base) que prejudicam a obtenção de informações relevantes do sistema em estudo (FERREIRA, 2015).

Uma necessidade chave para se ter um conjunto de dados de boa qualidade é obter um sinal tão informativo quanto possível após remover o ruído de um conjunto de dados. Um método importante para revelar a relação sinal-ruído envolve suavização ou alisamento (*smoothing*), cujo princípio é que o ruído será suavizado usando funções matemáticas suaves, enquanto o sinal permanecerá. Entretanto, demasiada suavização acarretará na redução de intensidade e resolução do sinal, por outro lado pouca suavização, o ruído permanece. A função de suavização ideal depende da largura de pico (em pontos de dados), bem como características de ruído (BRERETON, 2003).

Portanto, a redução das variações aleatórias, como os ruídos experimentais, pode ser efetuada por meio de técnicas de alisamento (*smoothing*). Dentre as técnicas de alisamento, as mais utilizadas são: alisamento pela média e alisamento pela média móvel (FERREIRA, 2015).

O emprego do alisamento pela média aplica-se quando se deseja reduzir o número total de  $m$ -variáveis ( $J$ ) de cada  $n$  amostra no sistema em estudo. O processo de alisamento basicamente envolve a escolha e o deslocamento de uma janela móvel ao longo de todo registo analítico (espectro, cromatograma, etc.). O primeiro passo consiste na seleção de uma janela de abertura  $(2m + 1)$  contém  $(2m + 1)$  variáveis

caracterizadas por  $-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m$  onde o centro da janela está em zero. Vale lembrar que o “ $m$ ” utilizado aqui na expressão da janela de abertura não corresponde ao mesmo  $m$  o qual representa cada de variável que forma a matriz  $\mathbf{X}$ . Assim, o resultado da curva alisada, representada por  $x_{alis}$  (equação 11), é obtida por meio do cálculo simples da média aritmética das repostas das  $(2m + 1)$  variáveis da respectiva janela. Este processo se repetirá para as próximas janelas até que todo cromatograma seja alisado. Ao final do processo, além do próprio alisamento (smoothing) do gráfico há uma diminuição no número de variáveis o qual inicialmente era  $J$  e passou a  $J/(2m + 1)$  (FERREIRA 2015).

$$x_{alis} = \sum_{k=-m}^{+m} c_k x_k = \frac{1}{(2m+1)} \sum_{k=-m}^{+m} x_k \quad (\text{equação 11})$$

O processo de alisamento realizado pela média móvel ocorre similarmente ao da média descrito pela equação 11, diferenciando-se na maneira com a qual a janela move-se ao longo do gráfico a ser alisado, além de manter o número de variáveis inicial  $J$  ao final do processo. Desta forma, no alisamento pela média móvel a janela varre variável por variável ao invés de janela por janela (FERREIRA 2015).

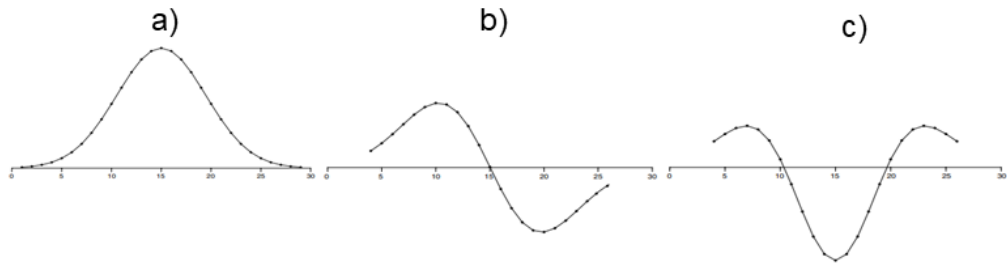
Do ponto de vista das variáveis sistemáticas, as técnicas derivação (primeira e segunda derivada), logaritmo e normalização são frequentemente usadas para corrigir a linha de base. Porém, atentaremos apenas para técnicas: primeira e segunda derivada, logaritmo, normalização, pois, são as mais frequentemente utilizadas em cromatografia (BRERETON, 2003).

➤ Primeira e segunda derivada

Seu princípio é que os pontos de inflexão em picos próximos se tornam pontos de inversão nas derivadas. A primeira (Fig. 18b) e segunda (Fig. 18c) derivadas de um Gaussiano puro (Fig. 18a) são apresentadas na Figura 18.

A primeira derivada é igual a zero no centro do pico e é uma boa maneira de identificar com precisão a posição de um pico amplo (largo). Exibe dois pontos de inflexão. Já a segunda derivada é um ponto de mínimo localizado no centro do pico, intercepta o zero nas posições dos pontos de inflexão para a primeira derivada e exibe

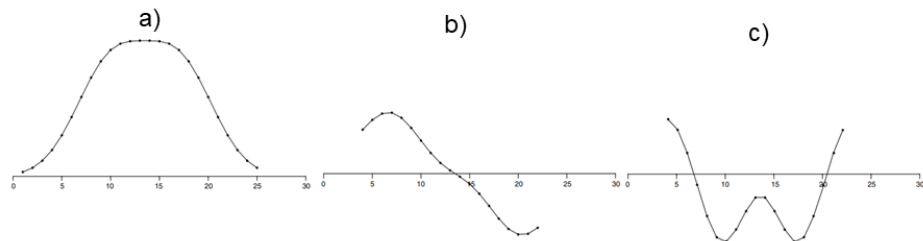
dois pontos de inflexão adicionais e mais distantes do que na primeira derivada. A técnica de derivadas provoca ligeiro estreitamento dos picos (BRERETON, 2003).



**Figura 18** - Primeira (a) e segunda (b) derivada de um Gaussiano puro.

**Fonte:** Brereton (2003).

O uso da técnica de primeira e segunda derivada é mais útil quando há vários picos que se sobrepõem. Neste sentido, A Figura 19 ilustra a primeira e segunda derivada de dois picos próximos que estão sobrepostos. A segunda derivada indica claramente a presença dois picos aponto com precisão suas posições. Apesar de que a aparência da primeira derivada insinue que o pico (Fig. 19a) não é puro, para este caso, não é possível concluir olhando somente para primeira derivada.



**Figura 19** - Um sinal representado por dois picos sobrepostos (a) aplicado a primeira (b) e segunda (c) derivada.

**Fonte:** Brereton (2003).

Embora, a técnica de derivação seja muito empregada para correção da linha de base, no entanto, há duas desvantagens ao empregá-la. A primeira, está relacionada ao processamento computacional intenso, pois um novo cálculo é processado para cada ponto de dados em um espectro ou cromatograma. Em segundo lugar, e mais importante, consiste na amplificação considerável do ruído e, portanto, exigindo baixos níveis de sinal para ruído. Entretanto, estas limitações podem ser superadas usando-se coeficientes de Savitsky-Golay, que envolvem

cálculos rápidos de funções suavizadas de derivada de ordem superior (CERQUEIRA *et al.*, 2000; BRERETON, 2003).

➤ Logaritmo

Basicamente, a finalidade da transformação logarítmica é evidenciar picos de baixa intensidade, frente aos de intensidades maiores, a fim de não perder nenhuma informação química relevante. Outra aplicação muito importante da transformação logarítmica consiste na linearização das funções entre as variáveis do sistema em estudo, sem prejuízos a interpretação real entre elas (FERREIRA, 2015).

➤ Normalização

Nesse tipo de transformação, os valores de cada uma das variáveis (altura do pico, área do pico, etc.) de uma dada amostra são divididos por uma constante, também conhecida como fator de normalização. Os métodos de normalização mais usados são realizados pela norma dos dados, cuja resposta máxima de cada uma das amostras é igual a uma unidade (FERREIRA, 2015).

Geralmente, em cromatografia, a normalização usa a variável área dos picos. Uma das finalidades da normalização é reduzir as interferências provocadas pela dimensão de partículas e diferenças de densidades entre as amostras (BEEBE; PELL; SEASHOLTZ, 1998; BYLESJÖ, CLOAREC; RANTALAINEN, 2009).

### 2.3.1.2 Processamento prévio dos dados: pré-processamento

Na seção anterior (2.3.1.1), reportamos alguns métodos de tratamento prévio dos dados aplicados as amostras (linhas da matriz **X**), agora iremos abordar os métodos de pré-processamento aplicados às variáveis (colunas da matriz **X**) mais comuns empregados na etapa de tratamento prévio de qualquer análise quimiométrica.

A realização do pré-processamento desempenha a função de compatibilização dos dados e/ou até remoção de possíveis fontes de variação, pois, em muitos casos as variáveis agrupadas possuem dimensões ou amplitudes distintas, de modo que necessitam de um recurso matemático para eliminar potenciais fontes de erros,

preservando, assim, as informações estáticas destes dados (MOITA NETO; MOITA, 1998) (FERREIRA 2015).

Portanto, a escolha adequada de um procedimento de pré-tratamento é essencial para o sucesso de qualquer análise quimiométrica. Neste sentido, os procedimentos mais comuns são: centragem de dados na média, escalamento pela variância e o autoescalamento (MATOS *et al.*, 2003; FERREIRA 2015).

➤ Centragem de dados na média

Centrar os dados na média, consiste em calcular o valor médio de cada coluna da matriz de dados (matriz **X**), e subtrair desse valor (média) cada um dos valores da respectiva coluna, como mostra as equações 12 e 13:

$$x_{ij(cm)} = x_{ij} - \bar{x}_j \quad (\text{equação 12})$$

$$\bar{x}_j = \frac{1}{I} \sum_{i=1}^I x_{ij} \quad (\text{equação 13})$$

onde  $x_{ij(cm)}$  é o valor centrado na média para a variável  $j$  na amostra  $i$ ,  $x_{ij}$  é o valor da variável  $j$  na amostra  $i$  e  $\bar{x}_j$  é a média dos valores das mostras na coluna  $j$  para um número  $I$  de amostras (FERREIRA 2015).

Ao final do processo de centragem na média, o eixo do sistema de coordenadas é deslocado para o centroide dos dados. Estatisticamente falando, ao centrar dados na média estamos prevenindo que dados localizados muito afastados do centroide possuam maior influência do que os dados mais próximos. Para sistema cujas variáveis foram medidas com mesma unidade, a centragem na média nos sugere que as variáveis com maior variação são as mais importantes (MASSART *et al.*, 1997).

➤ Escalamento pela variância

Nos processos de escalamento em geral, os dados são divididos por um fator de escala que representa uma medida da dispersão dos dados. Desta forma, no escalamento por variância (equações 14 e 15), cada uma das variáveis (colunas da matriz **X**) é dividida pelo respectivo desvio padrão ( $s_j$ ).

$$x_{ij(v)} = \frac{x_{ij}}{s_j} \quad (\text{equação 14})$$

$$s_j^2 = \frac{1}{I-1} \sum_{i=1}^I (x_{ij} - \bar{x}_j)^2 \quad (\text{equação 15})$$

onde  $x_{ij(v)}$  é o valor escalado pela variância para a variável  $j$  na amostra  $i$ ,  $s_j$  é o desvio padrão dos valores da variável  $j$  calculado a partir da variância ( $s_j^2$ ),  $s_j^2$  é a variância da  $j$ -ésima variável,  $x_{ij}$  é o valor da variável  $j$  na amostra  $i$  e  $\bar{x}_j$  é a média dos valores das amostras na coluna  $j$  calculados pela equação 15 e  $I$  é o número de amostras (FERREIRA 2015).

#### ➤ Autoescalamento

O autoescalamento é realizado a partir da centralização dos dados na média seguido por escalamento pela variância, ou seja, cada elemento de uma coluna da matriz  $\mathbf{X}$  é subtraído pela média da respectiva coluna e dividido pelo resultado do desvio-padrão dessa coluna, como mostra a equação 16:

$$x_{ij(a)} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (\text{equação 16})$$

onde  $x_{ij(a)}$  é o valor autoescalado da variável  $j$  na amostra  $i$ ,  $s_j$  é o desvio padrão dos valores da variável  $j$  calculado a partir da variância ( $s_j^2$ ) e  $\bar{x}_j$  é a média dos valores das amostras na coluna  $j$  calculados pela equação 16 e  $I$  é o número de amostras (FERREIRA 2015). Ao final do autoescalamento, os dados tornam-se adimensionais apresentando o mesmo peso (importância) para todas as variáveis. Logo, é um procedimento utilizado quando as variáveis possuem unidades diferentes ou quando a faixa de variação dos dados é grande (MATOS *et al.*, 2003).

### 2.3.2 Planejamento de experimentos

O planejamento experimental tem como principal objetivo a melhoria da qualidade dos produtos ou processos, avaliando um maior número de variáveis ao mesmo tempo que reduz o número de testes, otimizando assim o uso de recursos (GALDAMEZ, 2002). Através de um planejamento, pode-se determinar quais as variáveis que exercem maior influência no desempenho de determinado processo, e, com isso, reduzir tempo e custos operacionais, além de melhorar o rendimento do processo e as concordâncias entre os valores nominais obtidos e os pretendidos (CALADO; MONTGOMERY, 2003).

Em um experimento, os fatores ou variáveis de entrada atuam sobre o sistema produzindo mudanças nos valores das respostas seguindo uma função matemática desconhecida e o objetivo do planejamento experimental é descobrir essa função (BARROS NETO; SCARMINIO; BRUNS, 2001).

Antes de dar início aos experimentos, é necessário que os critérios e objetivos já estejam bem definidos, então, pode-se seguir com as escolhas das variáveis que serão envolvidas, de sua faixa de variação e níveis para cada uma; da variável que servirá de resposta e do planejamento em si, considerando o número de réplicas e a ordem de realização dos experimentos (CALDADO; MONTGOMERY, 2003).

O planejamento de experimentos é dado por uma otimização multivariada e geralmente emprega o planejamento fatorial completo em uma etapa inicial, denominada triagem, onde as variáveis experimentais que influenciam significativamente a resposta são selecionadas e, numa etapa seguinte, estas variáveis selecionadas são otimizadas por ferramentas estatísticas que geram modelos quadráticos, como, por exemplo, o planejamento composto central e, por fim, é estabelecida uma equação matemática que relaciona as variáveis a serem otimizadas com a resposta analítica através da metodologia de superfície de resposta (ROSA; RODRIGUES; FERREIRA, 2016).

#### 2.3.2.1 Planejamento fatorial

O planejamento fatorial é empregado no estudo de efeitos onde duas ou mais variáveis são consideradas importantes. É classificado como um método do tipo simultâneo, no qual as variáveis que apresentam influências significativas na resposta são avaliadas ao mesmo tempo. Elas são escolhidas e realiza-se experimentos para



todas as combinações possíveis dos níveis de cada variável selecionada (BARROS NETO; SCARMINIO; BRUNS, 1995; PEREIRA-FILHO, 2015).

Quando se estuda o mesmo número de níveis para todos os fatores, o planejamento fatorial completo pode ser representado por  $b^k$ , onde  $b$  é o número de níveis escolhidos e  $k$  é o número de fatores. Geralmente, os planejamentos fatoriais são do tipo  $2^k$  pelo fato de não ser necessário realizar muitos ensaios, entretanto, este número reduzido de níveis não permite uma exploração completa de uma ampla região no espaço das variáveis. Neste caso, este tipo de planejamento é bastante útil para observar tendências importantes para posteriormente realizar uma investigação mais aprofundada dos dados (BARROS NETO; SCARMINIO; BRUNS, 1995; CALADO; MONTGOMERY, 2003).

Nos planejamentos onde as variáveis são exploradas em 2 níveis, normalmente se usa a codificação (+) para o nível alto e (-) para o nível baixo como forma de ilustração, sem interferir na realização dos experimentos e na análise dos resultados (MACEDO, 2007).

Quando há um número elevado de fatores, o uso do planejamento fatorial completo acaba sendo inviável pelo também elevado número de experimentos necessários e, nesses casos, um planejamento fatorial fracionário é indicado. Experimentos fatoriais fracionados são muito úteis quando uma grande quantidade de fatores está sendo estudada e deseja-se identificar quais deles são realmente representativos. Nesta técnica, se utiliza um fator de redução que limita o número de ensaios a um valor mais baixo do que no planejamento completo, mas conduz, possivelmente, às mesmas conclusões que seriam obtidas com o número total de ensaios (BARROS NETO; SCARMINIO; BRUNS, 2001). O número total de observações é  $2^{k-1}$ , onde  $k$  é o número total de fatores e o número 1 no expoente remete à meia-fração. Por exemplo, um planejamento completo  $2^6$  resulta em 64 experimentos, enquanto  $2^{6-1}$  (o mesmo que  $2^5$ ), resulta em metade, ou seja, 32 experimentos, daí a denominação meia-fração (CALADO; MONTGOMERY, 2003).

Após efetivar os experimentos, para identificar quais são as variáveis que realmente exercem influência e requerem maior atenção, utiliza-se a observação dos efeitos, que é calculado como a média da resposta quando dada variável está no nível alto ( $\bar{Y}_+$ ) subtraído da média da resposta quando essa variável está no nível baixo ( $\bar{Y}_-$ ). Em termos prático, para 16 experimentos são calculados 15 efeitos, dos quais: - 4 principais (primeira ordem), sendo um para cada variável avaliada; - 6 secundários

(segunda ordem), onde são calculados os efeitos de interação linear entre as variáveis investigadas; - 4 terciários (terceira ordem), onde é feita a combinação em terceira ordem entre as variáveis, por exemplo, (123; 134...);- 1 de quarta ordem, onde há a interação de todos os efeitos (1234). (PEREIRA-FILHO, 2015).

No cálculo dos efeitos de segunda ordem, os níveis alto e baixo são identificados quando as duas variáveis estão no mesmo nível (-,- ou +,+) e em níveis diferentes (+,- ou -,+), respectivamente (PEREIRA-FILHO, 2015). Nem todos os efeitos ou combinações entre eles serão importantes e há algumas maneiras de avaliar quais realmente apresentam influência na resposta obtida, como exemplo, a análise de um gráfico de Pareto ou de um gráfico normal dos efeitos.

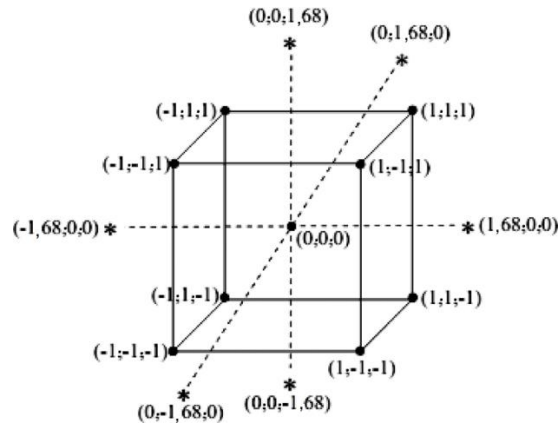
### 2.3.2.2 Planejamento composto central

O planejamento composto central (*Central Composite Design – CCD*) é utilizado para verificar a existência de termos quadráticos no modelo de regressão, isto é, a curvatura de um plano. Este tipo de planejamento consiste de uma parte referente ao planejamento fatorial  $2^k$ , com  $n_F$  ensaios,  $2k$  ensaios axiais ou estrela e  $n_c$  ensaios centrais (CALDADO e MONTGOMERY, 2003). Ele foi desenvolvido por Box e Wilson na década de 1950 como uma alternativa aos planejamentos fatoriais de três níveis que demandavam muitos experimentos para um pequeno número de fatores, mesmo quando se empregavam planejamentos fracionários (NOVAES *et al.*, 2017).

Há dois parâmetros que necessitam ser especificados neste tipo de experimento: a distância do centro do planejamento até os pontos axiais ( $\alpha$ , denominado rotabilidade) e o número de pontos centrais. A codificação das observações se dá por  $(\pm 1)$  para pontos fatoriais,  $(0)$  para ponto central e  $(\pm \alpha)$  para os pontos axiais. O número de experimentos ( $N$ ) para sua aplicação é dado por  $N = 2^k + 2k + PC$ , onde  $k$  é o número de fatores e  $PC$  é o número de réplicas do ponto central (CALADO; MONTGOMERY, 2003).

O CCD é bastante usado em conjunto com a metodologia de superfície de resposta, que será descrita mais a frente, para gerar um modelo quadrático que estime o efeito da curvatura. No CCD, o ponto central é replicado para obter boa estimativa do erro experimental e aumentar a precisão do experimento (NOMAN *et al.*, 2018). A forma geral de um CCD com 3 fatores, como o utilizado neste trabalho, e seus códigos

valores ( $\pm 1$ ), ponto central (0) e pontos axiais ( $\pm \alpha$ ) são ilustrados na Figura 20. Quando se utiliza 3 fatores, o valor codificado de  $\alpha$  é 1,68. Para chegar nesse valor de codificação, foi empregada a equação 17.



**Figura 20** - Pontos experimentais para o planejamento composto central utilizando três fatores.

**Fonte:** Noman *et al.* (2018).

$$\alpha = \sqrt[4]{n_F} \quad (\text{equação 17})$$

onde  $n_F$  é o número de ensaios fatoriais, no caso de um CCD com 3 fatores,  $n_F$  é igual a 8 (BRUNS; SCARMINI; BARROS NETO, 2006). Para 2 e 4 fatores, por exemplo, o valor de  $\alpha$  é 1,41 e 2,00, respectivamente (NOVAES *et al.*, 2017).

Para elaborar os experimentos nos pontos axiais, é preciso encontrar o valor real das variáveis que corresponderão ao valor codificado  $\alpha$  quando testadas. Para isso deve ser utilizada a equação 18, onde  $\alpha$  é chamado  $x_c$  (PEREIRA-FILHO, 2015).

$$x_c = \frac{x_0 - PC}{\frac{\Delta}{2}} \quad (\text{equação 18})$$

onde  $x_c$  e  $x_0$  são as condições codificadas e originais, respectivamente. PC é o ponto central e  $\Delta$  é a diferença entre os valores originais dos níveis +1 e -1.

### 2.3.2.3 Metodologia da superfície de resposta (MSR)

Superfícies de resposta são muito importantes em casos em que se deseja otimizar as variáveis de resposta que sofrem influência de muitas variáveis independentes. A MSR pode ser definida como um método estatístico que utiliza dados quantitativos de um desenho experimental adequado para determinar e simultaneamente solucionar equações multivariadas (NOVAES *et al.*, 2017).

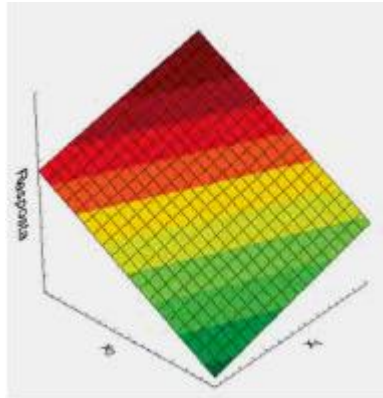
A determinação da superfície de resposta ocorre após a realização de um planejamento experimental que, normalmente, visa determinar quais fatores são os mais importantes para serem investigados e quais não são significativamente afetados pelos resultados experimentais. O primeiro passo para utilizar essa metodologia é a determinação da relação matemática existente entre a variável de resposta e os fatores. Se um polinômio de menor grau, tal como uma reta, se ajustar bem à resposta, então a função será dada por um modelo denominado modelo de primeira ordem, linear ou planar. Se houver curvatura, então um modelo de segunda ordem ou quadrático deverá ser utilizado (BEZERRA *et al.*, 2008).

A modelagem é feita ajustando-se modelos simples (em geral, lineares ou quadráticos) a respostas obtidas com planejamentos fatoriais realizados previamente. O deslocamento se dá sempre ao longo do caminho de máxima inclinação de um determinado modelo, que é a trajetória na qual a resposta varia de forma mais pronunciada (BARROS NETO; SCARMINIO; BRUNS, 2001; BEZERRA *et al.*, 2008).

Primeiramente, então, é investigada a possibilidade de a superfície de resposta ser descrita por polinômios de primeira ordem, ou seja, que ela siga o modelo linear descrito pela equação 19 nos parâmetros:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + \varepsilon \quad (\text{equação 19})$$

onde  $y$  é a resposta,  $x_1$  e  $x_2$  são os fatores experimentais,  $b_0$ ,  $b_1$  e  $b_2$  são os coeficientes do polinômio, e  $\varepsilon$  é o erro experimental (BEZERRA *et al.*, 2008). Se o termo de interação  $b_{12}x_1x_2$  for desprezível, a superfície de resposta é planar, como mostra a Figura 21.



**Figura 21** - Superfície de resposta linear ou planar.

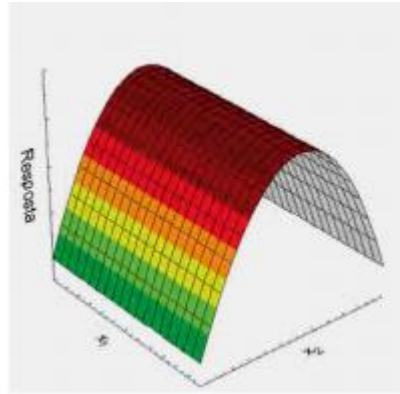
**Fonte:** Bandeira (2004).

Caso o modelo linear não seja suficiente para descrever os dados experimentais em decorrência da existência de uma curvatura na superfície de resposta, os dados gerados por este modelo podem ser úteis, os quais serão usados para indicar a direção na qual o planejamento original deve ser deslocado para que as condições ótimas sejam atingidas. Uma estratégia de encontrar condições operacionais eficientes para encontrar o ponto ótimo deve ser adotada seguindo um caminho de ascendência (máximo) ou descendência (mínimo), e, quando a possível região onde estiver o ponto ótimo for encontrada, deve-se adotar um modelo de segunda ordem, e (BEZERRA *et al.*, 2008; CALADO; MONTGOMERY, 2003; NOVAES *et al.*, 2017). Neste caso, uma função que pode descrever estes processos é um polinômio de segunda ordem, ou modelo quadrático completo, cujas respostas são estimadas pela equação:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{12}x_1x_2 + \varepsilon \quad (\text{equação 20})$$

onde,  $y$  é a resposta,  $x_1$  e  $x_2$  são os fatores experimentais,  $x_1x_2$  é a interação entre os fatores,  $b_0$ ,  $b_1$ ,  $b_2$ ,  $b_{11}$ ,  $b_{22}$  e  $b_{12}$  são os coeficientes, e  $\varepsilon$  é o erro experimental (BEZERRA *et al.*, 2008).

Algumas vezes, a superfície quadrática (Figura 22) obtida não corresponde ao modelo correto para os dados experimentais, mas pode ser utilizada para indicar um ponto ótimo localizado fora da região estudada, mostrando a direção na qual o planejamento deve ser deslocado, se possível (NOVAES *et al.*, 2017).



**Figura 22** - Superfície de resposta quadrática.

**Fonte:** Bandeira (2004).

Os valores de  $b$  dos polinômios podem ser obtidos pelos mínimos quadrados, conforme a equação:

$$b = (X^T X)^{-1} X^T y \quad (\text{equação 21})$$

onde  $b$  é o vetor com os coeficientes,  $X$  é a matriz com os parâmetros experimentais codificados, e  $y$  é a resposta experimental (NOVAES *et al.*, 2017).

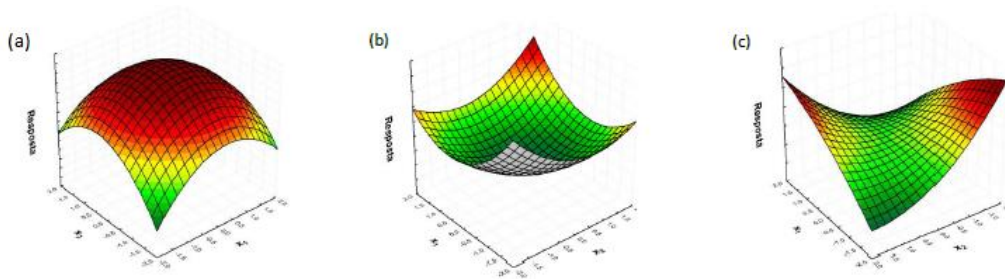
A última etapa consiste em estudar a superfície de resposta na região de interesse a fim de determinar as condições ótimas para os dados modelados através da localização do ponto ótimo, também conhecido como ponto crítico ou estacionário. Esta determinação pode ser realizada a partir da equação da regressão obtida (equação 20). Através dela, é possível obter as derivadas parciais de  $y$  em relação a  $x_1$  ( $\frac{\partial y}{\partial x_1}$ ) e de  $y$  em relação a  $x_2$  ( $\frac{\partial y}{\partial x_2}$ ), equações 22 e 23, respectivamente neste caso onde há 2 fatores, e, para um modelo que tenha mais variáveis, obtém-se os valores de  $x_1, x_2, x_3 \dots x_k$ , conforme o número de fatores presentes (BEZERRA *et al.*, 2008; NOVAES *et al.*, 2017).

$$\frac{\partial y}{\partial x_1} = b_1 + 2b_{11}x_1 + b_{12}x_2 = 0 \quad (\text{equação 22})$$

$$\frac{\partial y}{\partial x_2} = b_2 + 2b_{22}x_2 + b_{12}x_1 = 0 \quad (\text{equação 23})$$

Para modelos quadráticos, o ponto ótimo pode ser caracterizado como máximo, mínimo ou sela (Figura 23) e é definido pelo conjunto de pontos  $(x_1, x_2, \dots, x_k)$  para os quais as derivadas parciais se igualam a zero, conforme a equação:

$$\frac{\partial y}{\partial x_1} = \frac{\partial y}{\partial x_2} = \dots = \frac{\partial y}{\partial x_k} = 0 \quad (\text{equação 24})$$



**Figura 23** - Perfis de superfícies de resposta comumente encontrados: (a) Superfície com ponto de máximo; (b) Superfície com ponto de mínimo; (c) Superfície com ponto de sela.

**Fonte:** Novaes *et al.* (2017).

Quando há três ou mais variáveis, a visualização do gráfico predito pela função ajustada apenas é possível se uma ou mais variáveis forem fixadas em um determinado valor, de forma que o gráfico apresente a resposta *versus* duas variáveis na forma tridimensional. Essa visualização pode ser realizada por meio de gráficos de superfícies de resposta ou através de gráficos de contorno. Em um gráfico de superfície, os valores dos dois fatores são representados nos eixos x e y, enquanto os valores da resposta são no eixo z, gerando uma visão tridimensional que pode exibir um desenho mais nítido da superfície de resposta, enquanto um gráfico de contorno contém os valores dos dois fatores representados nos eixos x e y e os valores da resposta são representados por regiões sombreadas, chamadas contorno. Um gráfico de contorno é similar a um mapa topográfico (SILVA, 2008).

Após a elaboração do modelo e determinação do ponto ótimo, deve haver uma etapa de avaliação, para confirmar se o modelo construído foi realmente adequado ao sistema descrito e, assim, finalizar a otimização multivariada (BARROS NETO; SCARMINIO; BRUNS, 2001).

### 2.3.2.4 Análise de Variância (*Analysis of Variance* – ANOVA)

A análise de variância ou ANOVA é usada para a avaliação numérica da qualidade do ajuste de um modelo, sendo o método mais utilizado atualmente na estatística (BARROS NETO; SCARMINIO; BRUNS, 2001; NOVAES *et al.*, 2017). A Tabela 5 mostra de forma resumida a análise de variância para o ajuste a partir do método dos mínimos quadrados.

A soma quadrática da regressão ( $SQ_r$ ), representa a soma dos desvios das previsões feitas pelo modelo, em relação à média global, já a soma quadrática residual ( $SQ_{res}$ ) corresponde a soma das diferenças entre os valores observados e os valores previstos e a soma quadrática total ( $SQ_t$ ) é a soma dos desvios dos valores observados, em relação à média global (BARROS NETO; SCARMINIO; BRUNS, 2001). A Soma quadrática total pode ser calculada pela equação 25:

$$SQ_t = SQ_r + SQ_{res} \quad (\text{Equação 25})$$

Através da equação 25 é possível concluir que para um modelo bem ajustado, a  $SQ_t$  deve ser aproximadamente igual a  $SQ_r$ . O coeficiente de determinação do modelo ( $R^2$ ) é dado pela equação 26, e o seu valor percentual representa a variação explicada pelo modelo aplicado (BARROS NETO; SCARMINIO; BRUNS, 2001).

$$R^2 = \frac{SQ_r}{SQ_t} \quad (\text{equação 26})$$

A máxima variância explicável pelo modelo é definida pela equação 27, descrita abaixo:

$$\text{Máxima variância explicável} = \frac{SQ_t - SQ_{ep}}{SQ_t} \quad (\text{equação 27})$$

Os valores obtidos para a variância explicável e a máxima variância explicável, juntamente com os valores de  $F$  calculados e  $F$  tabelados para a regressão e a falta de ajuste indicam a qualidade do modelo utilizado (CALADO; MONTGOMERY, 2003; PEREIRA-FILHO, 2015).



Tabela 5 - Descrição da tabela ANOVA para regressão múltipla.

Parâmetro	Soma Quadrática	Graus de liberdade	Média Quadrática	Valor de $F$
Regressão	$SQr = \sum_{i=1}^m \sum_{j=1}^{n_i} (\hat{y}_i - \bar{y})^2$	$p-1$	$MQr = \frac{SQr}{p-1}$	$F = \frac{MQr}{MQres}$
Resíduo	$SQres = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_i)^2$	$n-p$	$MQres = \frac{SQres}{n-p}$	-
Total	$SQt = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$	$n-1$	-	-
Erro puro	$SQep = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_i)^2$	$n-m$	$MQep = \frac{SQep}{n-m}$	-
Falta de ajuste	$SQfaj = \sum_{i=1}^m \sum_{j=1}^{n_i} (\hat{y}_i - \bar{y}_i)^2$	$m-p$	$MQfaj = \frac{SQfaj}{m-p}$	$F = \frac{MQfaj}{MQep}$

**Fonte:** NOVAES *et al.*, 2017. Onde:  $i$  e  $j$  = níveis de experimentação;  $n_i$  = número de repetições no nível  $i$ ;  $m$  = número de níveis distintos da variável independente;  $n = \sum n_i$  = número total de observações;  $p$  = número de parâmetros do modelo;  $\hat{y}_i$  = valor estimado ou previsto pelo modelo ajustado, para a variável resposta em função do nível do fator;  $y_i$  = valor experimental observado na variável resposta ou variável dependente, em função do nível da variável independente;  $\bar{y}_i$  = valor médio de  $y_i$ ;  $F$  = distribuição  $F$  de Snedecor com graus de liberdade das somas de quadrados envolvidos.

Para um modelo ser considerado bem ajustado aos dados experimentais ele deve apresentar uma regressão significativa e uma falta de ajuste não significativa. A maior parte da variação total das observações em torno da média deve ser descrita pela equação de regressão e a variação e o restante ficarão com os resíduos. Além disso, a maior parte da variação relacionada com os resíduos deve ser atribuída ao erro puro e não à falta de ajuste, a qual é diretamente atribuída à qualidade do modelo (NOVAES *et al.*, 2017; BRUNS; SCARMINIO; BARROS NETO, 2006).

### 2.3.3 Análise multivariada

Como mencionado na seção 5, a quimiometria emergiu da química, matemática e estatística introduzindo novos métodos de tratamento de dados de natureza química capazes de lidar com grandes quantidades de informação por meio de análise multivariada dos dados.

Entretanto, em alguns casos, pode ser produtivo isolar cada variável em um sistema e estudá-la separadamente. Mas, em muitos casos, as variáveis estão interligadas de tal forma que, quando analisadas individualmente, produzem pouca informação sobre o sistema (BEEBE; PELL; SEASHOLTZ, 1998; GEMPERLINE, 2006; MALINOWSKI, 2002). Por outro lado, na análise multivariada, as variáveis

podem ser examinadas simultaneamente para acessar as principais características do processo que as produziu. Logo, a abordagem multivariada permite-nos realizar duas ações muito importantes sobre o sistema em estudo: (i) explorar o desempenho de um conjunto grande de variáveis e (ii) determinar o efeito de cada variável na presença das outras (RENCHEER, 2002).

Um conjunto de dados é definido como um determinado número de objetos descritos em função de um determinado número de variáveis. Desta forma, em um conjunto de dados químicos, os objetos são representados por amostras, compostos, espectros ou cromatogramas, enquanto as variáveis, representam as características e/ou propriedades destes objetos, tais como: condutividade, pH, concentração, transmitância ou absorvância para diferentes comprimentos de onda, alturas e áreas dos picos cromatográficos, etc. (BARROS NETO; SCARMINIO; BRUNS 2006).

Na prática, os métodos baseados em análise multivariada de dados químicos extraem informações a partir de múltiplas variáveis de resposta medidas em uma amostra, e ainda repetitivamente medidas para um conjunto grande de amostras, adotando sempre as mesmas condições e os mesmos critérios analíticos (RENCHEER, 2002). Entretanto, em muitos casos, apenas algumas destas múltiplas variáveis medidas contém informações relevantes de natureza química, ao passo que a maioria destas fornecem pouco ou quase nenhuma contribuição que possa ser útil na interpretação dos resultados (CORREIA; FERREIRA, 2007; FERREIRA, 2002; FERREIRA et al., 1999). Assim, o julgamento a respeito de quais variáveis são, de fato, importantes para auxiliar na compreensão dos resultados, é efetuado a partir de critérios subjetivos e objetivos (MOITA NETO; MOITA, 1998).

A escolha dos critérios utilizados para definir o método quimiométrico multivariado que reduzirá o número de variáveis no sistema em estudo, deixando somente as mais relevantes quimicamente, dependerá da intuição química ou da experiência do pesquisador, pois cada método possui suas limitações e potencialidades, de modo que a escolha de um método inapropriado para um determinado sistema pode acarretar na distorção dos resultados (MOITA NETO; MOITA, 1998).

Em termos práticos, de modo geral, uma análise multivariada compreende os seguintes objetivos (BEEBE; PELL; SEASHOLTZ, 1998; GEMPERLINE, 2006; MALINOWSKI, 2002; MASSART *et al.*, 1997; RENCHEER, 2002):

- Redução da matriz original de dados (matriz  $X$ ): decomposição e redução de dados, sem prejuízos as informações relevantes, extraindo variáveis subjacentes lineares independentes (componentes principais) das variáveis originais, até encontrar um pequeno número de variáveis as quais possam explicar toda a variação na matriz de dados estudada ();
- Organização e tratamento prévio: agrupamento e adequação dos dados a partir de metodologias de organização, transformação e pré-processamento dos dados;
- Investigação da dependência entre as variáveis: estudar as relações estruturais entre as variáveis no sistema em análise;
- Previsão: a determinação das relações entre as variáveis deverá ser conduzida a partir da previsão de uma ou mais variáveis baseadas na observação de outras;
- Construção e teste de hipótese: a descrição dos dados frequentemente leva à formulação de hipóteses, as quais são então verificadas por testes de hipóteses, para descrever quantitativamente o valor de uma ou mais variáveis em função do valor de algumas outras variáveis (modelagem) ou tentar classificar os objetos de acordo com os valores das variáveis obtidas (classificação).

O vasto campo da análise multivariada implica diretamente na quantidade enorme de métodos quimiométricos disponíveis, cada qual com finalidades distintas. Vale ressaltar que não é nosso foco aqui abordar todos, logo, dissertaremos na sequência apenas sobre os métodos multivariados mais comumente utilizados em análise multivariada que compreende os seguintes: métodos de reconhecimento de padrões - Análise Hierárquica de Agrupamentos (*Hierarchical Cluster Analysis* - HCA); Análise de Componentes Principais (*Principal Component Analysis* – PCA); algoritmo do k-ésimo Vizinho Mais Próximo (*k-Nearest Neighbor* - kNN) e algoritmo de Modelagem Flexível Independente Por Analogia de Classe (*Soft Modeling of Class Analogy* - SIMCA), e por fim, o método de calibração multivariada - Regressão pelo

Método dos Quadros Mínimos Parciais (*Partial Least Squares* – PLS). Dentre estes citados, o HCA e PLS foram os métodos de análise multivariado utilizado neste trabalho.

### 2.3.3.1 Métodos de reconhecimento de padrões

A busca por semelhanças que se repetem sistematicamente em várias amostras, objetos ou indivíduos formando, assim, um padrão e/ou uma regularidade, é tarefa constante de todo profissional da área científica ou técnica que trabalha com medidas de rotina.

Do ponto de vista da quimiometria, a procura por esse padrão e regularidade em amostras químicas estende-se para um enorme conjunto de dados, de modo que a matriz de dados gerada na maioria dos casos é muito grande, demandando por recursos matemáticos e estáticos cada vez mais sofisticados para dar algum significado químico para o problema. Uma alternativa interessante reside nos métodos quimiométricos de reconhecimento de padrões, pois estes identificam as semelhanças e as diferenças em várias amostras, para, então, agrupá-las e classificá-las de acordo uma tendência observada (FERREIRA 2015).

Os métodos de reconhecimento de padrões podem ser classificados de acordo com sua finalidade, logo, são divididos em dois grupos: os não supervisionados e os supervisionados de reconhecimento de padrões (BRERETON, 2009; FERREIRA 2015). Neste trabalho, foi utilizado o método HCA, do tipo não supervisionado, o qual será melhor abordado no próximo tópico.

#### 2.3.3.1.1 Métodos não supervisionados

Os métodos não supervisionados de reconhecimento de padrões, sem a preocupação de uma regra preditiva, são conhecidos também como métodos de análise exploratória de dados (FERREIRA, 2015). Assim, no reconhecimento não supervisionado, não há necessidade do conhecimento prévio sobre a classificação das amostras, de modo que estas são naturalmente agrupadas a partir da informação contida nos dados experimentais sob análise. Desta forma, as amostras de um mesmo grupo apresentam relações naturais entre si, enquanto as de diferentes grupos são distintas uma das outras. O reconhecimento não supervisionado é capaz também de

detectar a presença de amostras atípica (*outliers*), ou seja, amostras que não possuem o mesmo padrão de comportamento das demais (BRERETON, 2003; CORREIA; FERREIRA, 2007)

Na análise exploratória de dados, o interesse principal consiste em verificar e interpretar a correlação entre variáveis ou amostras. Desta forma, a análise hierárquica de agrupamentos (HCA) e análise de componentes principais (PCA) são dois métodos não supervisionados de reconhecimento de padrões amplamente empregados que permitem a interpretação multivariada de conjuntos de dados complexos a partir de tendências e semelhanças entre os dados observadas por meio de gráficos bi- ou tridimensionais, as quais não seriam perceptíveis por simples observações (MATOS *et al.*, 2003; CORREIA; FERREIRA, 2007; SOUZA; POPPI, 2012; FERREIRA 2015).

#### 2.3.3.1.1.1 Análise hierárquica de agrupamentos (HCA)

O principal objetivo da análise hierárquica de agrupamentos consiste em associar os dados de modo o qual enfatize os agrupamentos de maneira natural. A HCA pode se aplicar ao agrupamento tanto de amostras como de variáveis, de modo que o agrupamento de amostras revela semelhanças entre estas, enquanto o agrupamento de variáveis fornece uma noção de correlação das mesmas (MASSART *et al.*, 1997; MOITA NETO; MOITA, 1998; INFOMETRIX, 2014).

Apesar que as observações dadas nesta seção sejam expressas em termos do agrupamento de amostras, elas se aplicam perfeitamente ao agrupamento de variáveis.

Assim, um agrupamento hierárquico interliga as amostras em agrupamentos com atributos semelhantes entre si, conforme as variáveis selecionadas. Ao contrário dos métodos que agrupam amostras em categorias grupos preestabelecidos, tais como os métodos supervisionados, a HCA procura definir essas categorias em primeiro lugar. O resultado desse agrupamento é apresentado na forma de um diagrama conhecido como dendrograma, facilitando o reconhecimento visual de semelhanças entre amostras ou objetos espacialmente representados por pontos contidos em uma dimensão maior que três (3D), uma vez que gráficos convencionais não conseguem reproduzir graficamente com a mesma riqueza de informação (MOITA NETO; MOITA, 1998; INFOMETRIX, 2014).

Há diversos modos de determinar as semelhanças entre amostras as quais definiram os agrupamentos destas em um espaço multidimensional de um dendograma. Desta forma, uma maneira simples é dada a partir de uma expressão matemática que agrupa os pares de pontos mais próximos, com base na distância euclidiana, substituindo-os por um novo ponto situado em uma região cuja distância até este novo ponto é a metade da distância entre eles. Logo, quando todos os pontos estiverem agrupados em um único ponto, um dendograma pode ser construído, de modo que no seu eixo horizontal encontram-se as amostras e no vertical o índice de similaridade ( $s_{ij}$ ), entre os pontos  $i$  e  $j$ , dado pela seguinte expressão:

$$s_{ij} = 1 - \left( \frac{d_{ij}}{d_{m\acute{a}x}} \right) \quad (\text{equação 28})$$

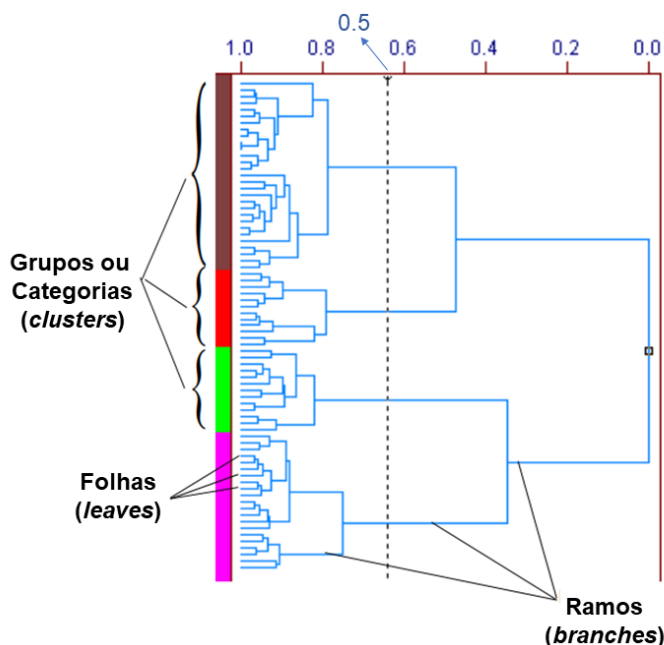
onde  $d_{ij}$  é a distância entre os pontos  $i$  e  $j$  e  $d_{m\acute{a}x}$  corresponde à distância máxima entre qualquer par de pontos (MOITA NETO; MOITA, 1998).

Uma vez determinado as similaridades, o par de amostras mais próximo é interligado dando origem a um agrupamento. De maneira análoga, o segundo par de amostras mais próximo se conecta e, assim, sucessivamente, até que todos os pares estejam interligados formando um único grupo. Assim, estando as amostras interligadas pela proximidade entre elas, o próximo passo consiste em definir a distância entre uma amostra e um grupo ou entre os próprios grupos de amostra (INFOMETRIX, 2014). Para tanto, existe atualmente uma série de técnicas disponíveis, dentre as quais destacam-se a conexão pelo vizinho mais próximo (*nearest-neighbor*), a conexão pela distância média ou centroíde (*average-linkage*) e a conexão pelo vizinho mais distante (*farthest-neighbor*) (SMOLIŃSKI; WALCZAK; EINAX, 2002; ALMEIDA *et al.*, 2007;).

A conexão pelo vizinho mais próximo é realizada em duas etapas., inicialmente, conecta-se duas amostras ou entre uma amostra e um grupo, a partir da menor distâncias entre eles. Na sequência, conexões são estabelecidas a partir da segunda menor distância e, assim, sucessivamente, até que um único agrupamento seja formado. Na conexão pela distância média, as amostras conectam-se aos grupos os quais têm seus respectivos centros mais próximos. O cálculo da distância entre dois grupos neste caso é realizado pela média aritmética ou ponderada. Na técnica pelo vizinho mais distante, as conexões são efetuadas sempre a partir da maior distância

entre as amostras das quais o par de similaridade é agrupado (SHARAF; SMOLIŃSKI, WALCZAK e EINAX, 2002; ALMEIDA *et al.*, 2007).

A Figura 24 ilustra um dendrograma hipotético de 75 amostras no qual apresenta os resultados dos agrupamentos de dados obtidos por HCA na forma de uma estrutura de árvore hierárquica dos dados, onde os comprimentos dos ramos desta árvore correspondem aos graus de similaridade entre as amostras.



**Figura 24** - Dendrograma obtido via HCA composto por 4 grupos a uma similaridade de 0.5 aplicada a análise de 75 amostras.

**Fonte:** INFOMETRIX (2014) – adaptado pela autora.

Os comprimentos dos ramos ou ramificações são proporcionais às distâncias entre grupos vinculados. O término dos ramos, extremidade esquerda do dendrograma, é denominado de folhas, e representa amostras únicas. O comprimento dos ramos que conecta dois grupos, está relacionado à sua similaridade, de modo que quanto mais longo o ramo, menor a semelhança e vice-versa. A similaridade (eixo horizontal no topo do gráfico) assume o valor 1.0 para a uma duplicata exata, e 0.0 indicando uma condição de distância máxima e não similaridade. A linha vertical pontilhada ao longo do dendrograma na Fig. 24 representa um valor de similaridade de aproximadamente 0.65, onde quatro agrupamentos (cores) podem ser distinguidos. É possível notar que os comprimentos dos quatro ramos são maiores (longos) comparados aos ramos que conectam os aglomerados à esquerda. No entanto, se o

ruído fosse adicionado aos dados, os comprimentos das folhas aumentariam e os comprimentos ramos diminuiriam. Geralmente, em dados aleatórios o comprimento das folhas é maior que a maioria das ramificações (INFOMETIX, 2014).

Vale reportar que, para uma HCA onde as variáveis foram obtidas com escalas distintas, um pré-tratamento deverá ser realizado, pois caso este não seja feito, as variáveis com valores numéricos mais elevados terão maior relevância no cálculo que as variáveis com valores numéricos menores. Um dos procedimentos de pré-tratamento muito útil neste caso é o autoescalado, uma vez que este equaliza a importância estatística de todas as variáveis utilizadas (MATOS *et al.*, 2003; INFOMETRIX, 2014)

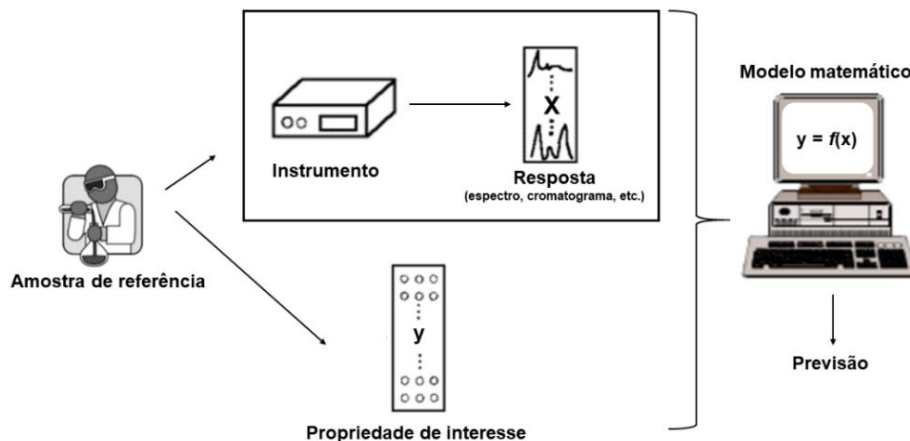
#### 2.3.3.2 Métodos de calibração multivariada

Há muitos fatores que influenciam a resposta obtida para uma substância química em uma amostra, de modo que, geralmente, não é possível obter de maneira direta, por exemplo, a concentração dessa substância a partir da resposta medida. Assim, a relação entre sinal ou resposta e a concentração deverá ser determinada por uma etapa denominada de calibração. O termo calibração pode ser definido como o uso de dados empíricos e conhecimento prévio para determinar informações quantitativas desconhecidas a partir de um número de medidas disponíveis, por meio de uma função matemática que represente um modelo (matemático, funcional, algorítmico, etc.) o qual descreva o comportamento do sistema sob análise (MASSART *et al.*, 1997; NAES *et al.*, 2002; BRERETON, 2003; FERREIRA, 2015).

Cabe esclarecer que com há vários métodos e calibração disponíveis em quimiometria, vamos nos atentar aqui aos métodos que visam a construção de modelos empíricos de propriedades mensuráveis os quais utilizam dados experimentais disponíveis a partir de métodos matemáticos conhecidos como métodos de regressão.

A Figura 25 ilustra um processo de calibração típico o qual visa a construção de modelos matemáticos que relacionam dois conjuntos de dados, a resposta instrumental e a(s) propriedade(s) de interesse.





**Figura 25** - Esquema de um processo de calibração convencional que descreve a relação entre as respostas instrumentais ( $X$ ) e a informação concreta ( $y$ ). As respostas são transformadas em informação química por meio do modelo matemático definido  $y=f(x)$ .

**Fonte:** Ferreira (2015) – adaptado pela autora.

O primeiro conjunto de dados apresentado na Fig. 25 refere-se as repostas instrumentais medidas para uma série de amostras de referência, constituindo-se, assim, o conjunto de calibração. Este conjunto é representado pelo bloco  $X$  na Fig. 25 cujas variáveis são denominadas como preditoras ou variáveis independentes, dependendo do caso. O outro conjunto de dados, conjunto das variáveis dependentes, contém os valores de alguma(s) propriedade(s) de interesse relativa(s) a cada uma das amostras de referência, de modo que se for uma única propriedade o conjunto é representado pelo bloco  $y$ , e caso forem várias por  $Y$ .

Um dos métodos mais simples de calibração é o método de regressão linear o qual está representado pelas equações 29 e 30. Neste caso, a calibração descreve uma função matemática simples que representa a relação entre a variável medida  $x$  e uma variável dependente  $y$  (BRERETON, 2003).

$$y = f(x) \quad \text{(equação 29)}$$

$$y = kx + l \quad \text{(equação 30)}$$

onde  $k$  é coeficiente de regressão e  $l$  representa coeficiente linear da reta dada pela equação 30. No método de regressão simples, a função ( $f(x)$ ) é uma equação de primeiro grau ( $kx + l$ ).

Independentemente do método empregado na calibração, para um dado conjunto de amostras este deverá ser capaz de estabelecer uma expressão matemática que represente um modelo de calibração que determine a(s) propriedade(s) pretendida(s), em função de seus respectivos gráficos respostas (espectros, cromatogramas, etc.), quando uma variação nessa(s) propriedade(s) for relacionada a variações nos respectivos gráficos. Em outras palavras, se conjuntos de dados do bloco **X** (variáveis independentes) e bloco **y** ou **Y** (variáveis dependentes) estabelecerem uma relação direta ou indiretamente entre si por meio de princípios físico-químicos fundamentais (teoria), existirá um modelo teórico-experimental que permitirá quantificar essa correlação (FERREIRA, 2015).

Em termos de análise multivariada, o processo de calibração multivariada torna-se mais laborioso, pois envolve múltiplas variáveis, ou seja, uma matriz de dados de alta dimensionalidade cujo objetivo é construir um modelo de calibração, por exemplo um modelo de regressão, que satisfaça adequadamente as relações das propriedades **Y** em função das variáveis preditoras **X** medidas experimentalmente ou determinadas computacionalmente. Uma vez concluído a modelagem, para finalizar o processo de calibração, o modelo será submetido a validação a partir da otimização de todo sistema (instrumento + modelo matemático) a fim de encontrar a melhor correlação entre as várias disponíveis no conjunto de dados. Ao ser validado, o modelo estará habilitado para fazer previsões das propriedades, as quais foram utilizadas na sua construção e validação, de amostras desconhecidas (FERREIRA *et al.*, 1999).

Dos métodos de calibração multivariada que utilizam algoritmos específicos para construção de um modelo de regressão, o método de regressão por componentes principais (*Principal Component Regression* - PCR) e o método de regressão pelo método dos mínimos quadrados parciais (*Partial Least Square* - PLS), são de longe, os métodos de regressão mais usados em calibração multivariada (FERREIRA *et al.*, 1999; SOUZA e POPPI, 2012).

Embora apresentem muitas características semelhantes, tais como utilizam como ponto de partida a PCA, são métodos robustos e muito eficientes para lidar com ruídos experimentais, colinearidade e não linearidades, o PLS e o PCR possuem suas particularidades (BEEBE; PELL; SEASHOLTZ, 1998). Contudo, o que se pode considerar de diferente e/ou vantajoso entre esses dois métodos, de modo geral, é que o PLS contorna a dificuldade característica do PCR com relação ao modo de

construção das componentes principais, que utiliza unicamente as respostas instrumentais ( $\mathbf{X}$ ) desconsiderando as informações relativas à propriedade ( $\mathbf{y}$ ), por exemplo, concentração (FERREIRA *et al.*, 1999).

Existem diversos algoritmos matemático-estatísticos usados em PLS cuja diferença fundamental está na maneira com a qual estes manipulam e decompõem a matriz do conjunto de dados multivariados para construir o modelo de calibração (FERREIRA *et al.*, 1999). Outra categoria de algoritmo muito utilizada nos métodos de regressão linear é empregada para definir as variáveis e amostras mais representativas do conjunto original de dados para formarem o conjunto de calibração o qual será utilizado para avaliar a capacidade preditiva do modelo de calibração construído pelo PLS (KENNARD; STONE, 1969; SOUZA *et al.*, 2011) (INFOMETRIX, 2014).

Como não é nosso foco abordar todos os algoritmos aqui, abordemos o algoritmo Kennard-Stone (KENNARD; STONE, 1969), pois foi este o algoritmo utilizado neste trabalho para selecionar as amostras e variáveis do conjunto de calibração para ser aplicado ao modelo de calibração obtido pelo método PLS.

#### 2.3.3.2.1 Regressão por mínimos quadrados parciais (PLS)

Sem dúvida, o PLS é o método de calibração multivariada mais popular em quimiometria, embora tenha sido originalmente desenvolvido dentro da área de economia em 1970 por Herman Wold. Seu desenvolvimento na quimiometria aconteceu poucos anos depois por Svante, filho de Wold, que empregava o PLS para modelar dados com variáveis altamente correlacionadas, substituindo o método tradicional dos mínimos quadrados. A versão original do PLS era baseada na obtenção de uma sequência de modelos parciais ajustados pelo método dos mínimos quadrados do qual resultava em dois blocos de matrizes de dados, a matriz  $\mathbf{X}$  (variáveis independentes) e a matriz  $\mathbf{Y}$  (variáveis dependentes) (FERREIRA, 2015).

A base fundamental da maioria dos métodos modernos de tratamento de dados multivariados utiliza princípios da PCA. Contudo, a principal diferença entre estes, em termos de como utilizam a PCA para construir seus respectivos modelos de calibração, está na maneira como cada método, ao decompor a matriz de dados multivariados, elabora uma relação entre as variáveis independentes e dependentes a partir de uma combinação linear utilizando fatores, componentes principais, variáveis

latentes ou autovalores para criar um novo sistema eixo de menor dimensão e maior facilidade de visualização dos dados (FERREIRA *et al.*, 1999).

O PLS é basicamente um método de regressão linear desenvolvido para lidar com regressões de alta dimensão, cujo principal diferencial perante aos métodos similares, tal como PCR, está na elevada capacidade de previsão de uma ou mais propriedades de amostras desconhecidas, pois mantém um compromisso entre a explicação da variância em uma matriz de dados  $\mathbf{X}$  e a previsão da variável dependente  $\mathbf{y}$  ou do bloco de variáveis dependentes  $\mathbf{Y}$ , ao serem decompostas simultaneamente pela estimativa sucessiva das componentes principais, denominadas variáveis latentes (VLs), que capturam com maior eficiência a variância e a correlação entre  $\mathbf{X}$  e  $\mathbf{Y}$  de maneira simultânea durante a construção de um novo eixo de coordenada que melhor represente os dados e seus resíduos (FERREIRA *et al.*, 1999; FERREIRA 2015)

A decomposição da matriz de dados, a qual contém o bloco  $\mathbf{X}$  das variáveis independentes (respostas instrumentais) e o bloco  $\mathbf{Y}$  das variáveis dependentes (propriedade de interesse, por exemplo, concentração), é realizada por meio da análise de fatores. Neste tipo de decomposição, diferentemente ao método PCR, as variáveis independentes e dependentes são decompostas simultaneamente antes que se estabeleça a regressão entre elas, pois a etapa de decomposição simultânea impede a restrição de ortogonalidade (GELADI; KOWALSKI, 1986; KOMSTA; HEYDEN, 2018). Impedimento este crucial para garantir que a nova base formada, ou seja, os novos eixos resultantes da decomposição da matriz de dados, resulta da combinação linear de vetores linearmente independentes, e, por isso, formam um novo espaço vetorial (TRYGG; WOLD, 2002).

Os fatores denominados de variáveis latentes (VLs) no PLS são diferentes dos descritos no PCR. No PCR esses fatores decompõem a matriz  $\mathbf{X}$  utilizando os valores singulares (SVD). No PLS as VLs são definidas de maneira tal que expliquem simultaneamente a variância em  $\mathbf{X}$  e a previsão de  $\mathbf{Y}$ , pois incluem no cálculo das VLs as informações das propriedades de interesse que são mais eficientes para capturar informações relevantes do que as VLs definidas no PCR, para a construção do modelo de calibração (FERREIRA, 2015).

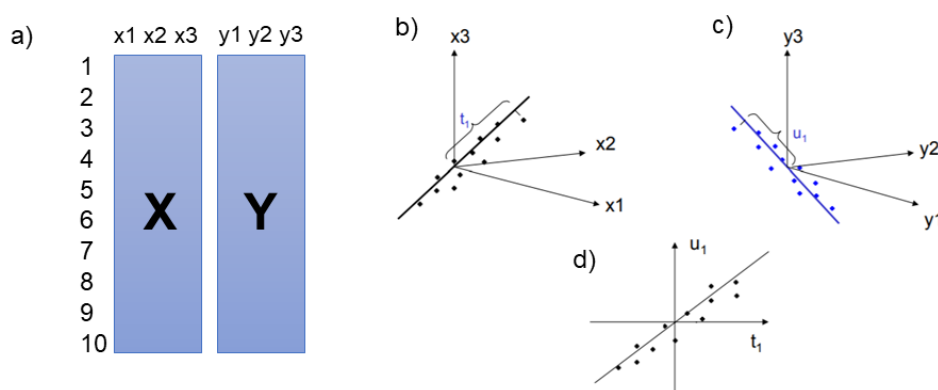
De modo geral, um processo de calibração multivariada pode ser dividido em dois passos: o conjunto de calibração ou treinamento e a validação. No conjunto de calibração, constrói-se um modelo, a partir da decomposição do conjunto da matriz

dos dados multivariados, que melhor descreve a relação entre as variáveis dependentes e independentes dessa matriz. A validação é uma etapa que precede a aplicação do modelo construído, cuja finalidade consiste em testar a capacidade preditiva do mesmo, ou seja, testar se o modelo construído é capaz de prever, por exemplo, a concentração ou qualquer outra propriedade ( $y$ ) ou conjunto de propriedades ( $Y$ ) de amostras desconhecidas (conjunto teste) ou amostras que não foram utilizadas na construção do modelo de calibração. Realizada a validação, será possível verificar se o modelo construído refletirá o comportamento real de um analito de interesse (FERREIRA *et al.*, 1999).

A etapa de construção de modelos por PLS, os blocos de variáveis  $X$  e  $Y$  são decompostas simultaneamente a partir de  $h$  variáveis latentes que capturam a variância e a correlação entre  $X$  e  $Y$ . A matriz  $X$  é formada por  $n \times p$  elementos, sendo  $n$  o número de amostras e  $p$  o número de variáveis independentes, enquanto o bloco  $Y$  pode ser representado por um vetor ( $y$ )  $n \times 1$ , quando o modelo PLS é capaz de prever somente um analito por vez (PLS1), ou uma matriz  $Y$  formada  $n \times k$ , sendo  $k$  o número de analitos, quando o PLS prediz  $k$  analitos simultaneamente (PLS2). Cabe esclarecer que as variações do PLS (PLS1 e PLS2) não representam a quantidade de analitos que o método é capaz de prever, mas sim a ordem do tensor, visto que um vetor é um tensor de primeira ordem e uma matriz é um tensor de segunda ordem (FERREIRA, 2015).

Portanto, o PLS pode ser visto como a extensão de regressão do PCA, uma vez que encaixa simultaneamente dois modelos do tipo PCA, um para o espaço  $X$  e outro para o espaço  $Y$ , de tal forma que as correlações entre  $X$  e  $Y$  são maximizadas. Em contraste com o PCA, o PLS é um método de covariância máxima, uma vez que o objetivo principal do PLS é prever as variáveis  $y$  pertencentes ao bloco  $Y$  a partir das variáveis independentes ( $x$ ) pertencentes ao bloco  $X$  (GELADI; KOWALSKI, 1986; WOLD, SJÖSTRÖM; ERIKSSON, 2001).

A Figura 26 ilustra uma interpretação geométrica da aplicação do método PLS em dois pequenos conjuntos de dados formados por dez objetos (amostras), três variáveis independentes ( $x_1$ ,  $x_2$  e  $x_3$ ) pertencentes ao bloco  $X$  e três variáveis dependentes ( $y_1$ ,  $y_2$  e  $y_3$ ) pertencente ao bloco  $Y$ .



**Figura 26** - Interpretação geométrica de um típico modelo PLS a partir da decomposição simultânea das matrizes de  $\mathbf{X}$  e  $\mathbf{Y}$  pelas variáveis latentes (VLs). a) um conjunto de dados composto por dez objetos (amostra) e três variáveis  $x$  e  $y$ . b) interpretação geométrica da primeira componente VLs no espaço  $\mathbf{X}$ . c) interpretação geométrica da primeira componente VLs no espaço  $\mathbf{Y}$ . d) relação interna entre as componentes VLs.

**Fonte:** Massart *et al.* (1997).

Como em um típico método PCA, o espaço  $\mathbf{X}$  recebe um sistema de coordenadas onde cada variável independente ( $x_1$ ,  $x_2$  e  $x_3$ ) dá origem a um eixo cujo comprimento é correspondente à sua escala (Fig. 26b), e de maneira análoga, o mesmo se aplica para o espaço  $\mathbf{Y}$  (Fig. 26c). Como os conjuntos de dados neste exemplo são formados por três variáveis  $x$  e três variáveis  $y$ , os espaços  $\mathbf{X}$  e  $\mathbf{Y}$  recebem três eixos. A centralização média de ambas variáveis ( $x$  e  $y$ ) então move os pontos para a região central do sistema de coordenadas. Uma função é então ajustada aos dados de maneira que melhor descreva a variação nos espaços  $\mathbf{X}$  e  $\mathbf{Y}$ , e ao mesmo tempo garanta a máxima correlação entre  $\mathbf{X}$  e  $\mathbf{Y}$ . Esta função é uma reta no espaço  $\mathbf{X}$  e  $\mathbf{Y}$  e constitui a primeira componente VLs produzida pela combinação linear da primeira variável latente, e, assim sucessivamente.

O PLS resume-se a decomposição do conjunto de dados multivariados ( $\mathbf{X}$  e  $\mathbf{Y}$ ) em novas variáveis latentes, vetores-colunas (*scores*) e vetores-linha (*loadings*), formando, assim, a base estrutural do modelo PLS. Desta forma, ao projetar cada ponto até a reta e medir a distância entre o ponto central e o ponto de projeção, obtém-se os *scores* para o espaço  $\mathbf{X}$  ( $t_1$ ) e os *loadings* para o espaço  $\mathbf{Y}$  ( $u_1$ ). Deste modo, o modelo de calibração criado por PLS pode ser constituído a partir de uma relação externa e uma relação interna entre as variáveis latentes, onde a relação externa descreve o bloco  $\mathbf{X}$  e  $\mathbf{Y}$  individualmente, enquanto a relação interna conecta os dois blocos em um mesmo sistema de coordenadas (Fig. 26d), ou seja, ao plotar  $u_1 \times t_1$

obtém-se uma relação interna das VLs mostrando a estrutura de correlação entre  $\mathbf{X}$  e  $\mathbf{Y}$  (GELADI; KOWALSKI, 1986; WOLD; SJÖSTRÖM; ERIKSSON, 2001).

As relações externas são dadas pelas seguintes decomposições de matrizes:

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} \quad (\text{equação 31})$$

$$\mathbf{Y} = \mathbf{UC}' + \mathbf{F} \quad (\text{equação 32})$$

onde  $\mathbf{T}$  é a matriz *scores* e  $\mathbf{P}'$  a matriz *loadings* do espaço  $\mathbf{X}$ , e  $\mathbf{U}$  corresponde a matriz *scores* e  $\mathbf{C}'$  a matriz *loadings* do espaço  $\mathbf{Y}$ .  $\mathbf{E}$  e  $\mathbf{F}$  são as matrizes residuais dos espaços  $\mathbf{X}$  e  $\mathbf{Y}$ , respectivamente. Ao plotar os valores de  $\mathbf{t}_1$  de dez amostras por exemplo, pelos valores de  $\mathbf{u}_1$  correspondentes, obtém-se a relação interna das variáveis latentes mostrando a estrutura de correlação entre  $\mathbf{X}$  e  $\mathbf{Y}$  (Fig. 26d) a partir da seguinte relação:

$$\mathbf{U} = \mathbf{BT} + \mathbf{H} \quad (\text{equação 33})$$

onde  $\mathbf{B}$  é uma matriz de identidade e  $\mathbf{H}$  é uma matriz residual (GELADI; KOWALSKI, 1986; WOLD; SJÖSTRÖM; ERIKSSON, 2001).

No PLS, calcula-se um tipo de *loadings* adicionais ( $\mathbf{W}$ ) que expressam a melhor correlação entre  $\mathbf{X}$  e  $\mathbf{Y}$ . Esses *loadings* estão relacionados com os coeficientes de regressão PLS ( $\mathbf{B}_{\text{PLS}}$ ) da seguinte forma:

$$\mathbf{B}_{\text{PLS}} = \mathbf{W}(\mathbf{P}'\mathbf{W})^{-1}\mathbf{C}' \quad (\text{equação 34})$$

Por fim, os coeficientes de regressão PLS mostram a direção e a magnitude da influência de uma variável  $x$  em uma variável  $y$  específica, de modo que a previsão de variáveis  $y$  ( $\hat{\mathbf{Y}}$ ) de novas amostras é dada por:

$$\hat{\mathbf{Y}} = \mathbf{XB}_{\text{PLS}} + \mathbf{F} \quad (\text{equação 35})$$

Depois que a primeira variável latente for calculada, a próxima pode ser calculada nas matrizes dos resíduos contidos nas matrizes  $\mathbf{E}$  e  $\mathbf{F}$  (GELADI; KOWALSKI, 1986; WOLD; SJÖSTRÖM; ERIKSSON, 2001).

Neste processo de construção do modelo de calibração a partir da decomposição das matrizes  $\mathbf{X}$  e  $\mathbf{Y}$  pelas variáveis latentes no qual obtemos relações

internas e externas entre  $\mathbf{X}$  e  $\mathbf{Y}$ , é necessário determinar um número adequado dessas variáveis ou fator  $k$ , uma vez que poucos fatores podem ser insuficientes para modelar adequadamente o sistema, por outro lado muitos fatores podem introduzir ruído no modelo de calibração, resultando em modelo com baixa capacidade preditiva para amostras externas ou desconhecidas (fora do conjunto calibração). Um número ideal de fatores corresponde ao número cujo modelo de calibração possui boa capacidade preditiva para amostras externas (FODOR; KOHL, 1996; FERREIRA *et al.*, 1999).

Uma metodologia muito útil e comumente empregada para determinar o número de variáveis latentes é conhecida como validação cruzada (WOLD, 1978). O método de validação cruzada adota um critério para definir um número de variáveis latentes baseado na avaliação da magnitude dos erros de previsão para um dado modelo de calibração. Essa avaliação é realizada por meio da comparação das previsões de uma propriedade de interesse, por exemplo concentração, previamente conhecidas (FERREIRA *et al.*, 1999).

No processo de avaliação por validação cruzada, o mesmo conjunto de amostras (conjunto original) é usado tanto para construir o modelo quanto para testá-lo, de modo que algumas amostras formarão o conjunto de calibração (construção do modelo) e as que não foram selecionadas para o conjunto de calibração formarão o conjunto teste ou de previsão (testar o modelo). Logo, a maneira como é conduzida a avaliação vai depender de como é feita a escolha das amostras as quais formarão esses dois conjuntos (calibração e teste). Assim, na validação cruzada *leave-one-out*, as amostras vão sendo excluídas do conjunto de calibração e constrói-se um modelo como o primeiro (original) a partir das amostras remanescentes. Usa-se esse modelo para prever a(s) propriedade(s) (variável dependente) das amostras removidas. Este processo se repete até que todas as amostras tenham sido excluídas ao menos uma vez. Calcula-se então o erro de previsão (equação 36) e a partir deste calcula-se a soma dos quadrados dos erros de previsão (*Predicted Residual Error of Sum Squares* - PRESS) (equação 37) e o desvio padrão da soma dos quadrados dos erros de previsão (*Root Mean Square Error of Prediction* - RMSEP) (equação 38) (FERREIRA *et al.*, 1999; VIAENE; HEYDEN, 2018).

$$(y_i - \hat{y}_i) \quad \text{(equação 36)}$$

$$\text{PRESS} = \sum_i (y_i - \hat{y}_i)^2 \quad \text{(equação 37)}$$



$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (\text{equação 38})$$

onde  $n$  é o número de amostras consideradas no conjunto de calibração,  $y_i$  é o valor medido e  $\hat{y}_i$  é o valor previsto da propriedade de interesse em questão para a  $i$ -ésima amostra.

Portanto, o número de fatores  $k$  ou de variáveis latentes mais adequado para o modelo de calibração analisado pela validação cruzada *leave-one-out* será aquele cuja soma dos quadrados dos erros de previsão (PRESS) para todas as amostras seja o menor valor encontrado (NAES *et al.* 2002)

Após determinado o número adequado de VLs por meio da validação cruzada, para finalizar o processo de validação antes da aplicação, de fato, do modelo construído pelo método PLS, é necessário avaliar a qualidade dos modelos de regressão criados (FERREIRA *et al.*, 1999). Para isso, são utilizados os erros padrão de previsão e calibração. No primeiro caso, utiliza-se um conjunto teste (amostras externas) para verificar a capacidade preditiva do modelo de previsão a partir do cálculo do o desvio padrão da soma dos quadrados dos erros de previsão (RMSEP). No segundo caso, para validar o modelo de calibração utiliza-se um conjunto de amostras de referência (validação externa) cujos valores para as variáveis dependentes (propriedade de interesse) são diferentes das amostras do conjunto de calibração empregadas para obter o modelo de calibração na validação cruzada. Neste caso, obtém-se o desvio padrão da soma dos quadrados dos erros de calibração (*Root Mean Square Error of Calibration* RMSEC). Quanto menor o RMSEP, melhor as propriedades preditivas do modelo (VIAENE; HEYDEN, 2018).

#### 2.3.3.2.2 Algoritmo Kennard-Stone (K-S)

Um fator crucial na etapa de calibração multivariada é a maneira como são definidas as amostras que farão parte do conjunto de calibração a partir de conjunto original de amostras muito grande (SALES *et al.*, 2000; DASZYKOWSKI; WALCZAK; MASSART, 2002;). Uma escolha aleatória de um conjunto de amostras para desenvolvimento de modelos de calibração pode ser prejudicial, pois muitas vezes as amostras selecionadas não representam com precisão a variabilidade dos objetos que se deseja avaliar (KANDUC; ZUPAN; MAJCEN, 2003).

Um outro reflexo dessa aleatoriedade é que em situações nas quais as análises são de alto custo e demoradas, a seleção de um elevado número de amostras para comporem o conjunto de calibração pode conter uma quantidade considerável de amostras que não contribuem para o aumento da capacidade preditiva dos modelos de calibração desenvolvidos (SOUZA *et al*, 2011).

A fim de evitar esses problemas é aconselhável que o conjunto de validação disponha de amostras homoganeamente distribuídas dentro da faixa da propriedade de interesse (composição, concentração, pH, temperatura, etc.), sem extrapolar o conjunto de calibração (KENNARD; STONE, 1969; KANDUC; ZUPAN; MAJCEN, 2003).

Neste trabalho, o qual não há controle da composição da matriz analítica (diesel), isto é, não é possível preparar as amostras de acordo com um planejamento prévio, torna-se necessário o emprego de alguma técnica a qual garanta uma seleção sistemática, e não aleatória, das amostras mais representativas no conjunto de calibração. Para tanto, foi usado o algoritmo Kennard-Stone que consiste em uma técnica que visa contemplar todo o espaço multidimensional dos dados de uma maneira mais uniforme para selecionar amostras que representem toda a variabilidade do conjunto de dados do sistema em análise (KENNARD; STONE, 1969).

Desenvolvida em 1969 por Kennard e Stone, o algoritmo K-S, basicamente, utiliza no conceito de distância euclidiana para selecionar um subconjunto significativo de amostras M, para constituir o conjunto de calibração, a partir de um conjunto mais amplo de amostras N, de modo que as amostras definidas para formar esse subconjunto M possam representar a maior variabilidade do conjunto N, assim como garantir uma distribuição homogênea (KENNARD; STONE, 1969).

Para certificar que o subconjunto de amostras N foi estabelecido a partir de critérios que atendam a homogeneidade da distribuição e a representatividade do espaço multidimensional dos dados (espaço  $\mathbf{X}$  das varáveis independentes), o algoritmo K-S adota um procedimento em seu processo de seleção denominado de *stepwise* (passo-a-passo), em que novas seleções de amostras são efetuadas em regiões do espaço distante das amostras já selecionadas (KENNARD e STONE, 1969). Para tanto, o algoritmo K-S usa a distância euclidiana ( $d_{Xab}$ ) entre os vetores  $\mathbf{X}$  para cada par ( $a,b$ ) de amostras da seguinte forma:

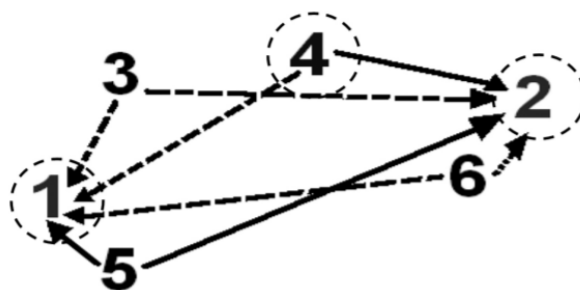
$$d_{Xab} = \left[ \sum_{j=1}^J (\mathbf{X}_a(j) - \mathbf{X}_b(j))^2 \right]^{1/2} \quad (\text{equação 39})$$

onde  $X_j$  são vetores que contêm as variáveis independentes (medida instrumental) no  $j$ -ésimo, por exemplo, comprimento de onda e  $J$  representa o número de comprimentos de onda nos quais as repostas instrumentais são registradas (SOUZA *et al.*, 2011).

O procedimento *stepwise* de seleção de amostras do K-S inicia-se a partir da seleção do par  $(a_1, b_2)$  de amostras para os quais a distância  $d_{xab}(a_1, b_2)$  seja maior. Para cada interação subsequente, o algoritmo K-S selecionará a amostra que apresentar a maior distância em relação a alguma amostra já selecionada. Este processo repetirá até que o número de amostras a ser selecionada, previamente estabelecido pelo analista, seja atingido. Assim, aplicando todos esses passos do procedimento, o algoritmo K-S garante que amostras as quais formarão o conjunto de calibração são representativas do modelo e homoganeamente distribuídas (KENNARD; STONE, 1969; INFOMETRIX, 2014).

Para se ter uma noção menos abstrato a respeito da maneira como o algoritmo K-S realiza seu procedimento de seleção de amostras, A Figura 27 ilustra um esquema simplificado do critério de seleção assumido pelo algoritmo.

O esquema exibido na Fig. 27 mostra que as amostras contemplam um conjunto “mais amplo” de seis amostras, em que foram selecionadas apenas três, de forma ordenada, para formarem o subconjunto de amostras as quais constituirão o conjunto de calibração.



**Figura 27** - Representação esquemática do processo de seleção de amostras pelo algoritmo de Kennard-Stone.

**Fonte:** Souza *et al.* (2011).

Assim, as amostras 1, 2 e 4 foram selecionadas respeitando essa ordem por serem consideradas pelo algoritmo K-S as mais representativas em relação as demais

dentro do espaço amostral considerado, levando em consideração as distâncias entre cada uma delas.

Portanto, a grande vantagem do algoritmo K-S, caso confirmado sua precisão ao final da sua realização, seria a obtenção de respostas analíticas instrumentais (cromatogramas, espectros, voltamogramas, etc.) relativamente mais rápidas e simples que o habitual, acompanhado pela diminuição na quantidade de análises laboratoriais, as quais demandam tempo e são onerosas (SOUZA *et al.*, 2011).

## **3 OBJETIVOS**

### **3.1 Objetivo geral**

O objetivo geral deste trabalho é o desenvolvimento de modelos de calibração multivariadas por PLS para a previsão de parâmetros físico-químicos do óleo diesel comercial B8 utilizando perfis obtidos através de análises por cromatografia gasosa ultrarrápida (CG-UR), visando a obtenção de uma metodologia alternativa confiável, viável, rápida e ambientalmente mais limpa para o monitoramento destes parâmetros.

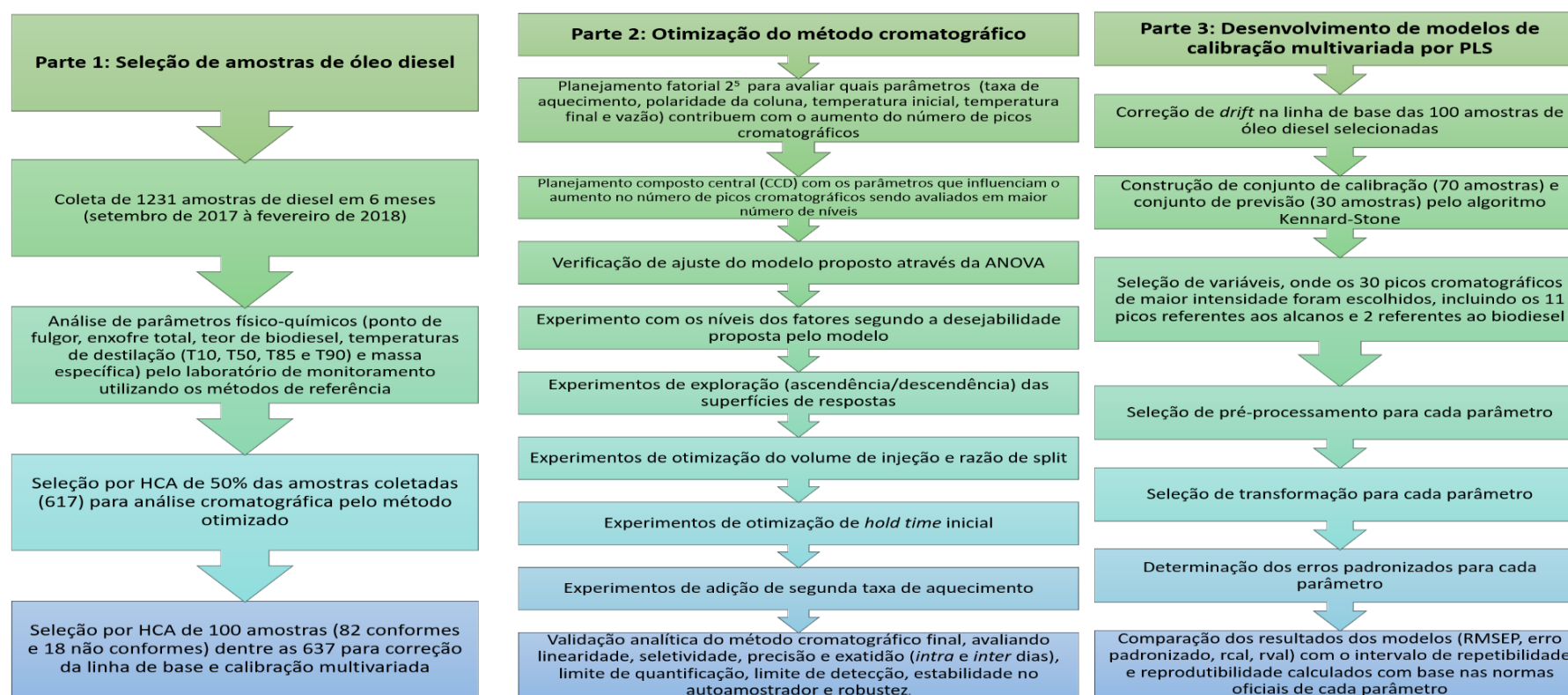
### **3.2 Objetivos específicos**

Com a finalidade de alcançar o objetivo geral, este trabalho tem como objetivos específicos:

- 1) Desenvolvimento de um método cromatográfico empregando CG-UR através de planejamento experimental com o intuito de obter o maior número de picos cromatográficos e uma resolução global adequada;
- 2) Seleção de um conjunto representativo de óleo diesel para efetuar as análises cromatográficas e quimiométricas;
- 3) Tratamento dos dados cromatográficos e criação das matrizes a serem utilizadas nos modelos;
- 4) Seleção de variáveis, pré-processamentos e transformações para obter melhor ajuste nos modelos de calibração multivariada por PLS;
- 5) Previsão dos parâmetros físico-químicos: massa específica, teor de biodiesel, teor de enxofre, ponto de fulgor e temperaturas de destilação em 10%, 50%, 85% e 90% do volume recuperado.

## 4. MATERIAIS E MÉTODOS

Este trabalho foi realizado em três etapas, não exatamente consecutivas: seleção de amostras, otimização do método cromatográfico e desenvolvimento dos modelos de calibração multivariada (Figura 28). Os materiais e métodos utilizados serão descritos por tópicos nas próximas seções.



**Figura 28** - Resumo das etapas do procedimento experimental realizado neste trabalho.

**Fonte:** Elaborado pela autora

#### 4.1 Coleta e armazenamento de amostras de óleo diesel

As amostras de óleo diesel utilizadas neste trabalho foram cedidas pelo CEMPEQC, que realizou toda a etapa de amostragem e coleta na região centro-oeste do estado de São Paulo (SP) e no estado de Mato Grosso do Sul (MS), as quais estão sob sua responsabilidade perante a ANP, de acordo com o contrato n° 1.069/16 (Tabela 6). Foram coletadas ao total 1231 amostras ao longo de 6 meses (setembro de 2017 a fevereiro de 2018), diretamente das bombas de abastecimento e armazenadas em frascos PET de 1L com tampas lacráveis.

Após a coleta, as amostras foram registradas e etiquetadas com códigos alfanuméricos, garantindo o sigilo acerca de suas procedências e foram transportadas em caixas com material absorvente e isolamento térmico, permitindo que seu armazenamento até o laboratório fosse realizado a temperaturas abaixo de 13 °C, a fim de evitar perda de compostos mais voláteis. No laboratório, as amostras foram mantidas em refrigerador a temperaturas de 10 °C até o momento da realização dos ensaios físico-químicos. Após estes ensaios, foram retiradas alíquotas em *vials* com tampas rosqueáveis de 1,5 mL e, novamente mantidas a temperaturas baixas (em *freezers* a - 24 °C) até a análise cromatográfica.

**Tabela 6** - Regiões de amostragem e coleta sob a responsabilidade do CEMPEQC.

Identificação ANP	REGIÃO
22	Lins (SP)
23	Franca (SP)
24	Barretos (SP)
25	São José do Rio Preto (SP)
26	Araçatuba (SP)
27	São Carlos (SP)
28	Araraquara (SP)
29	Marília (SP)
30	Presidente Prudente (SP)
31	Bauru (SP)
44	Campo Grande (MS)
45	Três Lagoas (MS)
46	Corumbá (MS)
47	Dourados (MS)

Fonte: Elaborado pela autora.

## 4.2 Análise dos parâmetros físico-químicos

As amostras de óleo diesel foram analisadas conforme procedimentos estabelecidos na Resolução ANP nº 50. O teor de biodiesel está de acordo com o estabelecido na Lei 13.263/2016, vigente quando as amostras foram obtidas e analisadas (8% v/v). Os parâmetros físico-químicos avaliados nas amostras de óleo diesel foram analisados pelo laboratório de monitoramento do CEMPEQC, que integra a Rede Brasileira de Laboratórios de Ensaio – RBLE desde 2008, através de metodologias acreditadas junto ao INMETRO segundo ABNT NBR ISO/IEC 17025:2005 (CRL:0308). Todos os ensaios foram efetivados seguindo as normas ABNT ou ASTM. A Tabela 7 apresenta as normas e os equipamentos utilizados em cada ensaio.

**Tabela 7** - Normas e equipamentos utilizados para cada parâmetro físico-químico do diesel.

PARÂMETROS	NORMAS	EQUIPAMENTO
Teor de Biodiesel	EN 14078	Espectrômetro de infravermelho FT-IR da Thermo Scientific, modelo Nicolet IR200.
Massa Específica	ABNT NBR 14065	Densímetro automático Anton Paar, modelo DMA 4500.
Destilação	ABNT NBR 9619	Destilador automático Takana, modelo AD-6.
Enxofre Total	ABNT NBR 14533 (S500)	Espectrômetro de Fluorescência de raios-X por energia dispersiva da marca Shimadzu, modelo EDX-800 (S500)
	ASTM D5453 (S10)	Analizador de enxofre e nitrogênio por fluorescência de ultravioleta da marca Thermo Scientific, modelo IPro 500 NS Analyzer.
Ponto de Fulgor	ASTM D93	Analizador automático de ponto de fulgor com vaso fechado Pensky-Maters, modelo PMA-4 da marca Petrotest.

**Fonte:** Elaborado pela autora.

## 4.3 Desenvolvimento de método de separação cromatográfica para o óleo diesel

O equipamento utilizado foi o cromatógrafo a gás *Thermo Trace GC Ultra* equipado com amostrador automático AS 3000, detector FID com alta frequência de aquisição de dados e módulo de aquecimento ultrarrápido (Figura 29). O sistema é interfaceado com o software *ChromQuest 4.1* (*Thermo Fisher Scientific*, EUA).

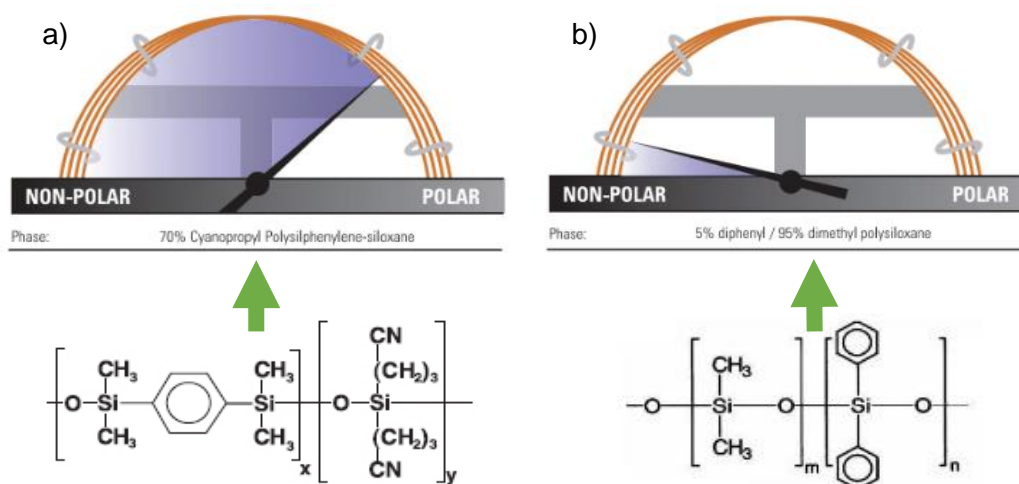




**Figura 29** - Módulo de aquecimento ultrarrápido instalado no forno de um cromatógrafo convencional.

**Fonte:** *Thermo Scientific* (2010).

As colunas utilizadas no desenvolvimento do método foram a UFM VF23MS (10m × 0,15mm × 0,15µm), de caráter polar, composta por 70% cianopropilpolissilfenilenossiloxano e a UFM PH SE54 (5m × 0,05mm × 0,3µm), de caráter apolar, composta por 5% difenil e 95% dimetilpolissiloxano, ambas da marca *Thermo Scientific* (Figura 30).



**Figura 30** - Polaridade e composição das fases estacionárias das colunas a) UFM VF23MS; b) UFM PH SE54.

**Fonte:** Adaptado de *Thermo Scientific*.

O gás de arraste utilizado foi o hidrogênio, e os gases utilizados pelo detector foram o gás hidrogênio, 45 mL min<sup>-1</sup>, o gás nitrogênio, 30 mL min<sup>-1</sup>, como *make-up* e o ar sintético, 350 mL min<sup>-1</sup>. Todos os gases utilizados possuem pureza mínima de

99,999% e são da marca *White Martins*. O injetor empregado foi o *Split/Splitless* no modo *split* e as amostras foram injetadas com uma seringa com capacidade de 5  $\mu\text{L}$ .

Para o desenvolvimento do método cromatográfico, utilizou-se uma amostra de óleo diesel comercial B8 (acrescida de 8% (v/v) de biodiesel) selecionada, classificada como conforme, e cedida pelo laboratório de monitoramento do CEMPEQC. Esta amostra foi analisada sem pré-tratamento ou adição de padrão interno. As condições cromatográficas iniciais foram baseadas no método desenvolvido por Munhoz (2013) e se encontram na Tabela 8.

**Tabela 8** - Condições da análise cromatográfica inicial.

<b>Parâmetros</b>	<b>Condição</b>
Temperatura do injetor	270°C
Temperatura do bloco	280°C
Temperatura Inicial	40°C
Taxa de aquecimento	100°C/min
Temperatura Final	250°C
Temperatura do detector	270°C
Coluna	UFM VF23MS (10m x 0,15mm x 0,15 $\mu\text{m}$ )
Gás de arraste	Hidrogênio
Modo de injeção	Split 1:100
Vazão do gás de arraste	0,5 mL/min
Volume de amostra injetado	0,1 $\mu\text{L}$
Tempo total de análise	2,85 min

**Fonte:** Elaborado pela autora.

A partir das condições iniciais, o método foi aperfeiçoado através da determinação das melhores condições de análise utilizando planejamento experimental e MSR, utilizando, para isso, os programas *Minitab® Statistical Software 17* da Minitab, Inc e *Chemoface 1.6.1*.

Primeiramente, foi elaborado um planejamento fatorial completo  $2^5$ , com 32 experimentos, como triagem para avaliar as variáveis mais importantes para a obtenção do maior número possível de picos cromatográficos da amostra complexa óleo diesel. Os fatores avaliados foram temperatura inicial (TI), temperatura final (TF), taxa de aquecimento, vazão e tipo de fase estacionária (com maior caráter polar – VF23MS ou maior caráter apolar – PH SE54), todos testados nos níveis +1 e -1 (Tabela 9).

**Tabela 9** - Níveis do planejamento fatorial completo 2<sup>5</sup>.

Variável	Nível inferior (-)	Nível superior (+)
Temperatura Inicial (°C)	30	80
Temperatura Final (°C)	200	250
Taxa de Aquecimento (°C/min)	80	150
Vazão (mL/min)	0,5	1,0
Coluna	PH SE54	VF23MS

Fonte: Elaborado pela autora.

A partir desses resultados de triagem, um planejamento composto central (CCD) foi realizado com os 3 fatores de grande influência remanescentes para uma verificação em maior número de níveis (5, ao invés de 2 realizados na etapa anterior) destes, provocando resultados mais aprofundados, a fim de realizar um refino do método. Os níveis +1 e -1 foram novamente avaliados e o ponto central, tido como a média aritmética entre os níveis superior e inferior, foi feito com quatro réplicas (20% dos experimentos). Os outros níveis -1,68 e +1,68 correspondem a região axial do CCD.

Juntamente com os resultados deste planejamento, o *software* utilizado gerou gráficos de contorno e superfícies de respostas para cada combinação entre variáveis, a equação do modelo e a tabela ANOVA, permitindo otimizar a resposta e verificar se o modelo proposto para esta otimização está bem ajustado.

#### 4.4 Validação analítica do método cromatográfico otimizado

Após desenvolvimento e otimização do método cromatográfico, foi realizada a etapa de validação analítica. Para isso, foi utilizado como solução estoque um *mix* contendo 1 mL de série homóloga de hidrocarbonetos parafínicos de cadeia linear (material de referência certificado – MRC), obtido da Supelco® (código 46855-U), que compreende a faixa de número de carbono de C<sub>8</sub> a C<sub>40</sub> solubilizados em diclorometano/dissulfeto de carbono na proporção 50:50. A concentração de cada componente desse *mix* é de 1000 µg.mL<sup>-1</sup>.

A partir dessa solução estoque, os pontos utilizados na curva analítica foram preparados em concentrações (m/m): 500  $\mu\text{g.g}^{-1}$ , 300  $\mu\text{g.g}^{-1}$ , 200  $\mu\text{g.g}^{-1}$ , 100  $\mu\text{g.g}^{-1}$  e 50  $\mu\text{g.g}^{-1}$  de cada hidrocarboneto (Tabela 10), com o auxílio da balança analítica *Ohaus*, série *Adventure* com capacidade de carga de 210 g e precisão de 0,0001 g (Certificado: BA 084\_08\_18 INMETRO). O solvente utilizado foi o hexano (reagente P.A., pureza 99% da marca *Vetec Química Fina*, obtido em frasco âmbar de 1L), escolhido por não interferir na análise cromatográfica dos demais alcanos analisados.

**Tabela 10** - Pontos utilizados na construção da curva analítica da etapa de validação, valores das concentrações nominais e reais dos hidrocarbonetos e massas de solução estoque (*mix*) e solvente utilizadas.

Ponto	Concentração nominal ( $\mu\text{g/g}$ )	Concentração real ( $\mu\text{g/g}$ )	Massa da solução <i>mix</i> (g)	Massa de hexano (g)
1	50	52,62	0,0354	0,4688
2	100	102,35	0,0694	0,4388
3	200	199,56	0,1343	0,3701
4	300	301,61	0,2033	0,3019
5	500	494,91	0,3342	0,1720
6	-	749,51	-	-

**Fonte:** Elaborado pela autora.

As amostras foram preparadas pesando as massas do *mix* referentes a cada concentração desejada e completando com a massa do solvente até atingir a massa total de 0,5 g. As soluções preparadas foram armazenadas em *vials* de 1,5 mL e transferidas no momento da análise cromatográfica para *inserts* com capacidade de 250  $\mu\text{L}$  localizados dentro de *vials* de 1,5 mL, de modo a garantir que a profundidade até onde a seringa consegue sorver a amostra inserida fosse alcançada.

A solução contida no *mix* também foi utilizada como ponto da curva analítica com unidade de concentração convertida de  $\mu\text{g.mL}^{-1}$  para  $\mu\text{g.g}^{-1}$ , resultando em 749,51  $\mu\text{g.g}^{-1}$ . Este valor foi obtido pesando o conteúdo do frasco do *mix* em um *vial* previamente tarado (1,3342g) e efetuando a divisão: (1000  $\mu\text{g}/1,3342$  g). Através do valor da massa da solução *mix*, foram feitos os cálculos para chegar aos valores das concentrações das demais amostras. Ela foi designada como o ponto de maior concentração ou limite superior de quantificação e armazenada em *insert* nas mesmas condições dos demais pontos.

As figuras de mérito avaliadas na validação analítica do método cromatográfico foram: seletividade, linearidade, precisão e exatidão (intradia e interdia), limite de detecção, limite de quantificação, estabilidade em autoamostrador e robustez. Os

cálculos referentes à etapa de validação foram efetivados pelo programa *Action Stat Pro* e pelo programa *Excel*.

#### **4.5 Seleção de amostras de diesel por HCA**

Com a finalidade de obter um conjunto representativo de amostras de óleo diesel e economizar tempo e custos que seriam gastos analisando todas as amostras coletadas, foi feita uma seleção de amostras através do método HCA.

Os valores dos parâmetros físico-químicos determinados pelo laboratório de monitoramento foram utilizados como variáveis neste método para a construção dos dendogramas. Como os parâmetros não se apresentam na mesma escala, o pré-processamento autoescalamamento foi empregado para a garantia de que todas as variáveis possuíssem mesma influência nos resultados da HCA. O critério de distância entre as amostras empregado foi a euclidiana, no qual as duas amostras mais próximas se conectam e formam um agrupamento e foi utilizado repetidamente até que um único grupo, que pudesse conectar todas as amostras, fosse formado. Então, foi utilizada a técnica de agrupamento incremental para definir a distância de uma amostra a um grupo ou entre os grupos de amostras, através de agrupamento obtido por conexão pela distância média. O programa *Pirouette 3.11*, *Infometrix Co.* foi utilizado em todos esses processos.

Após a definição de todos esses parâmetros, foram feitas as seleções de 50% do conjunto de amostras total por mês, durante o período de outubro de 2017 a fevereiro de 2018, selecionando as amostras com maiores distâncias entre si, ou seja, com a maior diferença possível para manter a representatividade.

#### **4.6 Análise cromatográfica das amostras selecionadas**

As 617 amostras de óleo diesel comercial selecionadas pelo método HCA foram analisadas pelo método otimizado através de CG-UR e os dados contidos nos cromatogramas (área dos picos) foram armazenados juntamente com os dados dos ensaios físico-químicos cedidos pelo monitoramento do CEMPEQC para utilização na construção dos modelos de calibração multivariada.

#### **4.7 Desenvolvimento de modelo de calibração multivariada pelo método PLS**

Esta etapa foi realizada em parceria com o grupo do Prof. Dr. Leonardo Pezza (IQ-Unesp, *Campus Araraquara*, onde foi utilizado o programa *Pirouette* na versão 4.5, que apresenta o pacote mais atual e completo, contemplando maior número de funções e mais recursos para auxiliar no desenvolvimento dos modelos de regressão multivariada, tais como o algoritmo Kennard-Stone para separação das amostras em conjuntos de calibração e de validação.

Os cromatogramas das amostras de diesel apresentam lacunas/*drifts* sobre a linha de base, o que é muito comum em separações cromatográficas de óleo diesel devido à complexidade da amostra em número de componentes, provocando uma saturação na coluna. Esse *drift* dificulta a integração das áreas e alinhamento dos picos cromatográficos, por isso, foi necessário realizar uma correção antes de desenvolver o modelo de PLS e o programa *Origin*<sup>®</sup> 8.6 foi empregado para essa finalidade. Nesta etapa, o conjunto amostral foi reduzido para um total de 100 devido às correções serem feitas manualmente através de escolhas de pontos da linha de base original e apenas os picos principais foram integrados, também pelo programa *Origin*<sup>®</sup> 8.6 de forma manual.

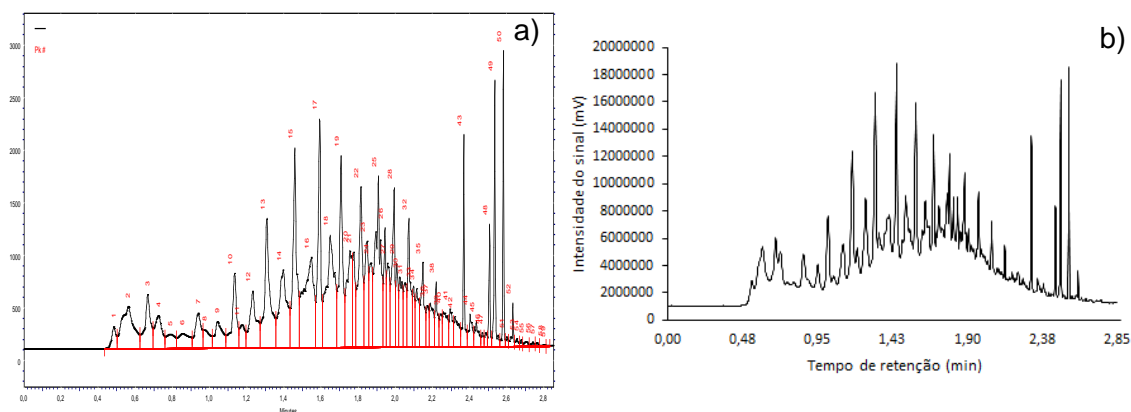
Com as linhas de base corrigidas, a PLS foi conduzida nos conjuntos para avaliar a possibilidade de elaboração de um modelo de regressão que correlacione as variáveis independentes (dados cromatográficos) com as variáveis dependentes (parâmetros físico-químicos), então, foram formados através do algoritmo *Kennard-Stone* os conjuntos de calibração (matriz 70 x 30 x 8 – amostras vs variáveis independentes vs variáveis dependentes) e o de previsão (matriz 30 x 30 x 8 – amostras vs variáveis independentes vs variáveis dependentes). As análises foram realizadas utilizando o *software* estatístico *Pirouette* 4.5.

Para a construção dos modelos de previsão por PLS foram avaliados os pré-processamentos: nenhum, autoescalado, centrado na média, escalado pela faixa e pela variância; as transformações matemáticas: nenhuma, primeira e segunda derivada, logarítmica e normalizada. Foi realizada a validação cruzada após a definição dos melhores critérios para prever cada parâmetro físico-químico. O intervalo de confiança foi de 95%.

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Desenvolvimento e otimização do método cromatográfico

Os resultados experimentais têm início com a análise realizada a partir das condições iniciais, onde foi obtido um cromatograma com 59 picos cromatográficos e resolução global igual a 1,08 (Figura 31). O número de picos cromatográficos foi obtido pelo *ChromQuest*, utilizando a opção ferramentas de integração numa configuração com largura (*Width*) igual a 0,05; limite (*Threshold*) igual a 10000; sensibilidade para detectar picos coeluídos em “ombros” (*Shoulder Sensitivity*) igual a 50 e área mínima de integração igual a 1000. As resoluções entre os pares de picos foram obtidas também pelo *ChromQuest* segundo o método da *United States Pharmacopeia* – USP, que utiliza a relação da diferença entre o tempo de retenção de dois analitos eluídos seguidamente com a metade da soma da largura de seus picos. A resolução global, referente ao cromatograma inteiro, foi calculada pela somatória dos valores das resoluções, dividida pelo número de picos obtidos menos um (FLUMIGNAN, 2005). Um exemplo de integração de picos pode ser observado na Figura 31a, enquanto a Figura 31b mostra o mesmo cromatograma com escalas legíveis nos eixos.



**Figura 31** - Cromatograma obtido sob as condições iniciais de análise, onde a temperatura inicial foi igual a 40 °C, taxa de aquecimento de 100 °C/min até a temperatura final de 250 °C; volume injetado de 0,1 µL e *split* de 1:100. Vazão de 0,5 mL.min<sup>-1</sup>. A coluna VF23MS foi utilizada como fase estacionária a) mostrando a integração dos picos; b) mostrando as escalas nos eixos.

Assim, a fim de melhorar a metodologia de separação cromatográfica do diesel, foi desenvolvido um planejamento experimental com o objetivo de gerar um perfil

cromatográfico com o maior número possível de picos, resultando em boa separação cromatográfica entre os diversos componentes do óleo diesel.

Na primeira etapa, de triagem, foi realizado um planejamento fatorial completo  $2^5$  e os resultados se encontram na Tabela 11.

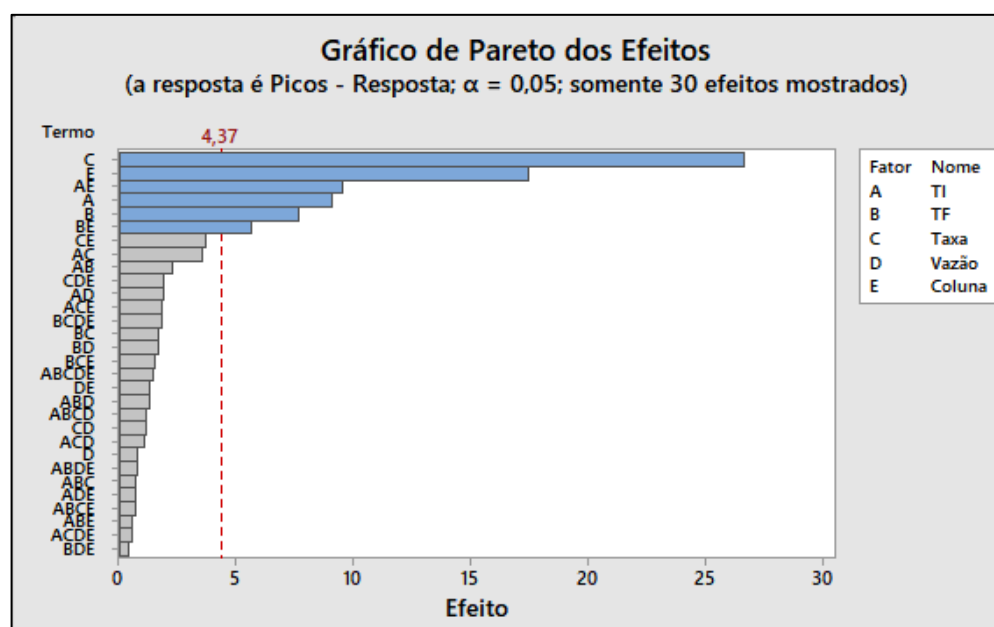
**Tabela 11** - Resultados obtidos no planejamento fatorial completo  $2^5$ .

Experimento	TI (°C)	TF (°C)	Taxa de aquecimento (°C/min)	Vazão (mL/min)	Coluna	Picos	Resolução global
1	-	-	-	-	+	82	0,62
2	-	-	-	+	+	77	0,96
3	-	-	+	-	+	59	1,53
4	-	+	-	-	+	85	1,13
5	+	-	-	-	+	81	1,24
6	+	+	-	-	+	82	1,09
7	-	+	+	-	+	59	1,77
8	-	-	+	+	+	62	1,07
9	+	-	+	-	+	60	1,45
10	-	+	-	+	+	88	1,25
11	+	-	-	+	+	79	1,00
12	+	+	+	-	+	58	2,20
13	+	+	-	+	+	85	1,68
14	+	-	+	+	+	64	0,59
15	-	+	+	+	+	58	0,84
16	+	+	+	+	+	60	1,46
17	-	-	-	-	-	108	1,08
18	-	-	-	+	-	108	1,22
19	-	-	+	-	-	80	1,23
20	-	+	-	-	-	123	1,75
21	+	-	-	-	-	85	0,81
22	+	+	-	-	-	96	0,87
23	-	+	+	-	-	89	0,89
24	-	-	+	+	-	79	1,39
25	+	-	+	-	-	62	0,56
26	-	+	-	+	-	123	1,11
27	+	-	-	+	-	86	1,15
28	+	+	+	-	-	72	1,51



29	+	+	+	-	-	99	1,22
30	+	-	+	+	-	62	0,76
31	-	+	+	+	-	88	1,53
32	+	+	+	+	-	70	1,10

Como pode ser observado, houve variação significativa nas respostas obtidas para os 32 experimentos. As condições experimentais que forneceram o maior número de picos foram aquelas referentes à utilização da coluna de caráter apolar (PH SE54) e apresentaram nível mínimo (-) para a temperatura inicial e taxa de aquecimento e nível máximo (+) para a temperatura final. Estes resultados foram representados em um gráfico de Pareto a fim de verificar quais parâmetros e combinações entre eles são mais significativos na otimização desta metodologia cromatográfica com a finalidade de elaboração de um perfil com o maior número de picos cromatográficos (Figura 32):



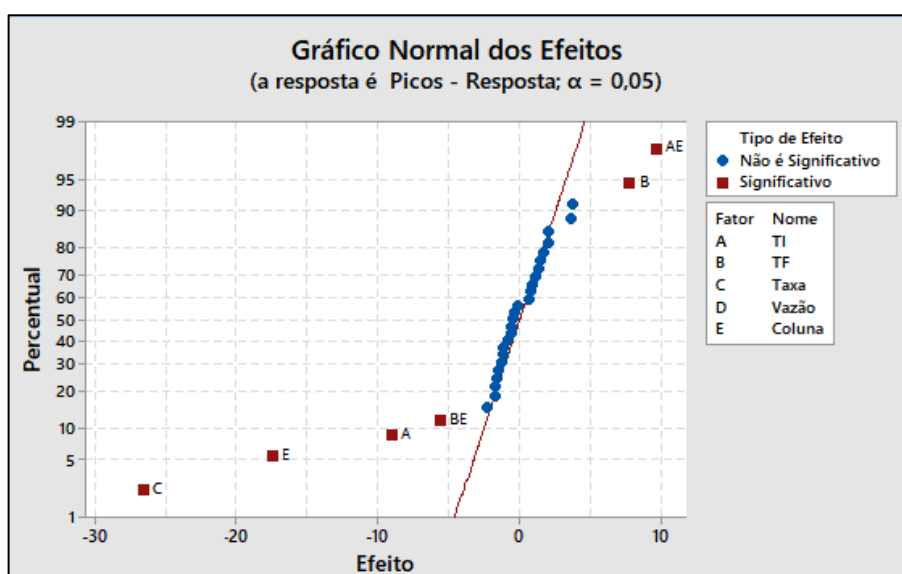
**Figura 32** - Gráfico de Pareto das variáveis avaliadas no planejamento fatorial completo  $2^5$ . Os efeitos lineares ou quadráticos em azul são considerados significativos ( $\alpha = 0,05$ ) para a contribuição com o aumento de picos cromatográficos. Os efeitos em cinza podem ser desprezados.

Analisando o gráfico de Pareto podemos confirmar os efeitos que influenciam na obtenção da maior quantidade de picos. O valor mínimo em módulo dos efeitos para serem considerados significativos depende do nível de significância (denotado

por  $\alpha$ ). Quanto mais à direita da reta pontilhada vermelha ( $\alpha = 0,05$ ), mais significativo é esse efeito principal ou de interação entre as variáveis.

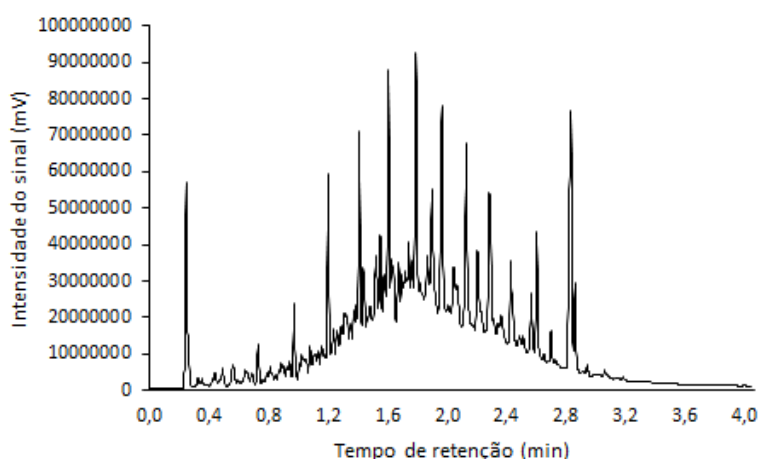
A observação do gráfico normal dos efeitos (Figura 33) além de confirmar quais fatores contribuem com os maiores efeitos, demonstrados no gráfico de Pareto, indica também se o efeito é positivo ou negativo, ou seja, se o fator utilizado no nível (+) ou no nível (-) irá contribuir mais para o aumento no número de picos cromatográficos.

Quando um efeito principal é positivo, admite-se que a utilização do maior nível proporciona uma melhor resposta, neste caso, maior TF. O inverso ocorre quando o efeito é negativo, como menor taxa, coluna PH SE54 e menor TI, corroborando com as informações extraídas da Tabela 11. Quando se trata de um termo de interação linear positivo, pode-se inferir que as variáveis podem estar ambas no maior ou menor nível, devido a multiplicação dos sinais codificados, então provavelmente, em análise da Figura 33 em conjunto com a Tabela 11, extrai-se a informação de que coluna PH SE54 combinada com menor TI (-,-) apresenta grande efeito positivo e, com a mesma linha de raciocínio, que a combinação da coluna PH SE54 com maior TF (-,+) resulta em efeito significativo negativo. Efeitos mais afastados de zero que estão no eixo x têm maior magnitude e são estatisticamente mais significativos. A distância em que os pontos devem estar da linha de referência para serem estatisticamente significativos depende do nível de significância ( $\alpha$ ).

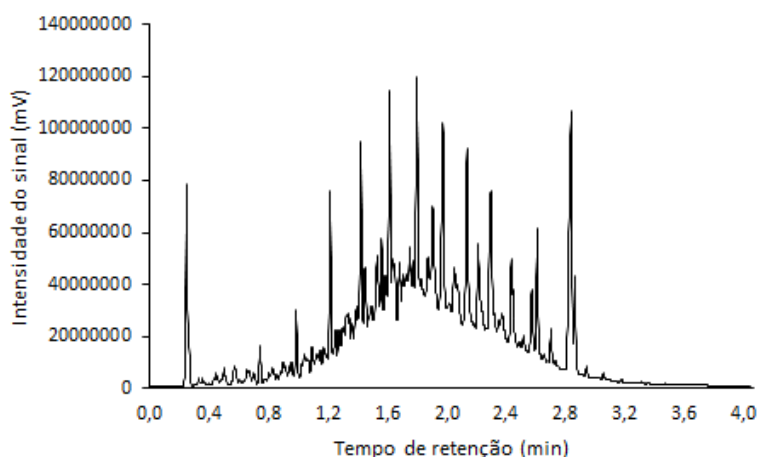


**Figura 33** - Gráfico normal dos efeitos calculados para o planejamento fatorial completo  $2^5$ . Apenas os efeitos representados em vermelho são significativos para o modelo de separação cromatográfica. Os efeitos negativos mostram que estes fatores ou combinações entre fatores influenciam positivamente o modelo quando estão no menor nível e os efeitos positivos indicam que estes fatores e combinações apresentam maior influência quando estão no maior nível.

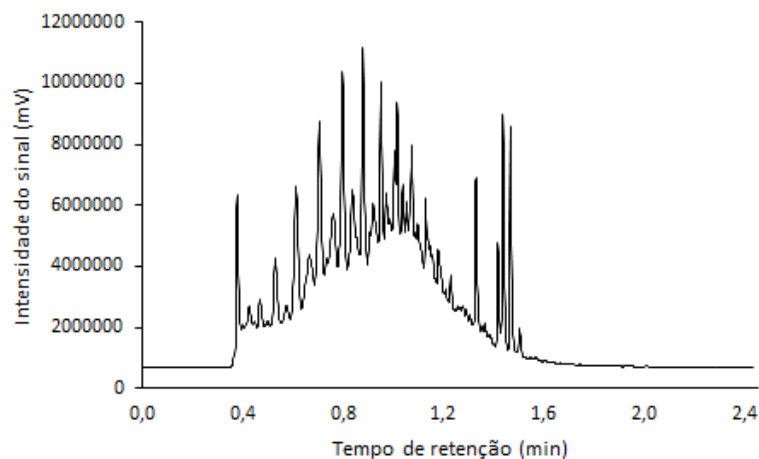
Entre as variáveis avaliadas, a vazão não se mostrou importante na contribuição para o aumento de número de picos. As demais variáveis se mostraram importantes e devem ser estudadas em maior número de níveis, exceto a coluna que é um fator qualitativo. Neste caso, a coluna PH SE54, verificada no modelo como nível qualitativo inferior, apresentou melhores resultados e foi mantida fixa para a próxima etapa. Para efeito de comparação, as Figuras 34, 35, 36 e 37 representam os dois cromatogramas com maior e menor quantidade de picos, respectivamente, a saber: maior quantidade de picos experimentos 20 (Fig. 34) e 26 (Fig. 35) enquanto os experimentos 12 (Fig. 36) e 15 (Fig. 37) correspondem os cromatogramas obtidos com menor quantidade de picos.



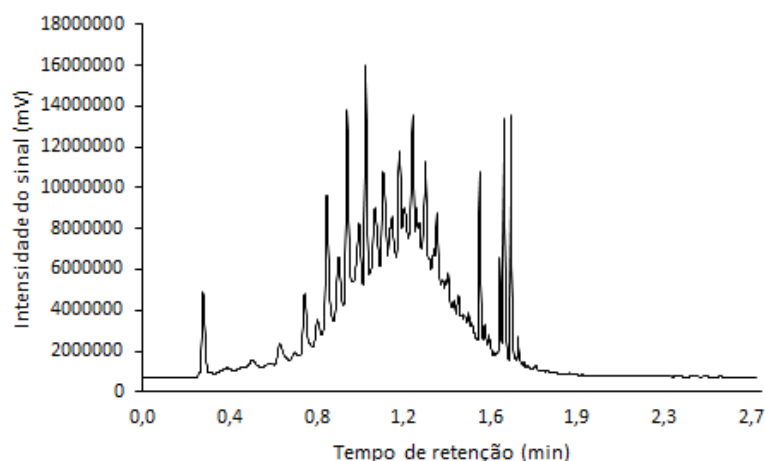
**Figura 34** - Cromatograma obtido do experimento número 20, onde temperatura final foi mantida no nível máximo e temperatura inicial, taxa de aquecimento e vazão foram mantidos no nível mínimo. A coluna empregada foi a de maior caráter apolar PH SE54.



**Figura 35** - Cromatograma obtido do experimento número 26, onde temperatura final e a vazão foram mantidos no nível máximo e temperatura inicial e taxa de aquecimento foram mantidos no nível mínimo. A coluna empregada foi a de maior caráter apolar PH SE54.



**Figura 36** - Cromatograma obtido do experimento número 12, onde temperatura inicial, temperatura final e taxa de aquecimento foram mantidos no nível máximo e a vazão foi mantido no nível mínimo. A coluna empregada foi a de caráter polar VF23MS.



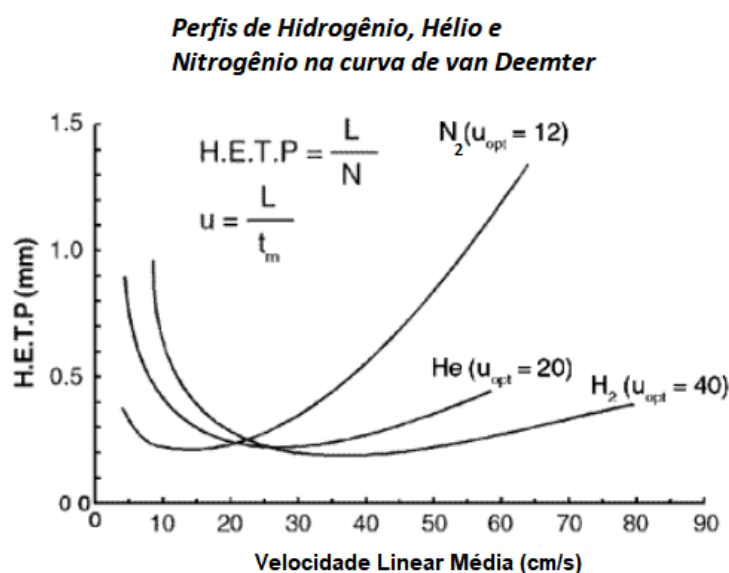
**Figura 37** - Cromatograma obtido do experimento número 15, onde a temperatura final, taxa de aquecimento e vazão foram mantidos no nível máximo e a temperatura inicial foi mantida no nível mínimo. A coluna empregada foi a de caráter polar VF23MS.

Vale observar que, com o mesmo número de picos e apresentando como diferença apenas o nível de vazão, o Experimento 20 (-) apresentou resolução global de 1,75, enquanto que o Experimento 26 (+) teve o valor de 1,11. Isto mostra que embora este fator não seja significativo para o aumento do número de picos segundo o planejamento experimental, ele colabora com a separação cromatográfica, mesmo porque, a vazão constante é diretamente relacionada com a velocidade linear da equação de *van Deemter*, pela equação 40:

$$u = \frac{V_c}{A} = \frac{V_c}{\pi r^2} \quad (\text{equação 40})$$

onde  $u$  é a velocidade linear do gás de arraste,  $V_c$  é a vazão constante,  $A$  corresponde à área da secção transversal da tubulação da coluna e  $r$  ao raio da tubulação da coluna (GROB; BARRY, 2004).

A partir os dados de diâmetro interno da coluna, e, portanto, de raio (0,0025 cm), substituídos na equação para uma vazão igual a 0,5 mL.min<sup>-1</sup>, o valor de velocidade linear encontrado é de aproximadamente 42 cm.s<sup>-1</sup>, sendo o hidrogênio o gás de arraste, é um valor bem próximo ao ideal, conforme a curva de *van Deemter* (Figura 38). Assim, o valor do fator vazão no nível inferior foi mantido fixo para as demais etapas.



**Figura 38** - Curva de *van Deemter* para gases hidrogênio, hélio e nitrogênio, comparando as velocidades lineares ideais.

**Fonte:** Adaptado de <https://www.restek.com/Restek-Capillary-Column-Installation-Guide-Section-II-In-depth-Installation-Information> (2018) - Adaptado pela autora.

Com os resultados obtidos no planejamento fatorial completo 2<sup>5</sup> foi possível determinar quais fatores exercem influência sob o número de picos e elaborar um detalhamento, objetivando melhorias no método cromatográfico ao estudar estas variáveis em mais níveis, não apenas no superior (+) e inferior (-), sendo possível obter uma otimização da resposta.

Assim, foi realizado um estudo de CCD em cinco níveis para as três variáveis, no caso, TI, TF e taxa de aquecimento. O ponto central é a média aritmética entre os níveis superior e inferior e foi feito com quatro réplicas (20% dos experimentos) e os

níveis -1,68 e +1,68 correspondem a região axial de um planejamento composto central. Neste planejamento foi utilizada a coluna PH SE54 e a vazão do gás de arraste mantida constante igual a 0,5 mL.min<sup>-1</sup>. Outras configurações do equipamento como volume de injeção, temperatura do injetor, dentre outros não foram alterados em relação ao método inicial. A Tabela 12 descreve os níveis de cada fator utilizado e a Tabela 13 apresenta o roteiro deste planejamento e a resposta (número de picos cromatográficos).

**Tabela 12 - Níveis do CDD com 3 fatores.**

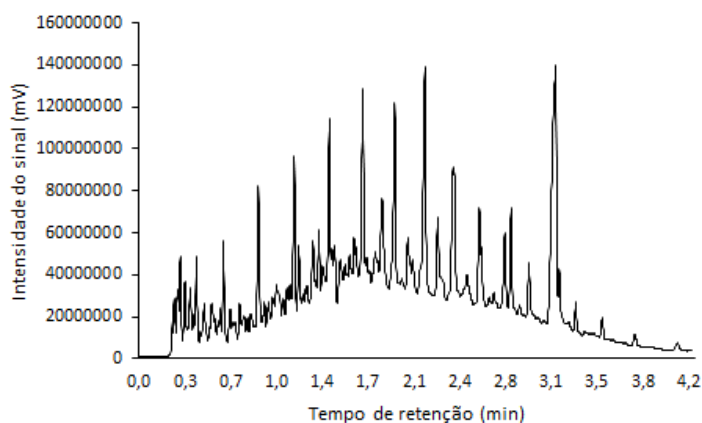
<b>Variável</b>	<b>-1,68</b>	<b>+1,68</b>	<b>-1</b>	<b>+1</b>	<b>0</b>
Temperatura Inicial (°C)	30	93	43	80	61
Temperatura Final (°C)	183	267	200	250	225
Taxa de Aquecimento (°C/min)	56	174	80	150	115

Devido à limitação instrumental, o mínimo valor possível para TI é 30 °C, então este valor real passou para o menor nível codificado (-1,68) e o novo valor real para o nível -1 foi calculado e resultou em 43 °C.

**Tabela 13 - Resultados do planejamento CCD. A resposta está representada por picos cromatográficos.**

<b>Experimento</b>	<b>TI (°C)</b>	<b>TF (°C)</b>	<b>Taxa de Aquecimento (°C/min)</b>	<b>Picos</b>	<b>Resolução global</b>
1	-	-	-	106	1,23
2	+	-	-	86	1,04
3	-	+	-	123	1,56
4	-	-	+	80	1,09
5	-	+	+	90	1,21
6	+	-	+	65	1,37
7	+	+	-	99	1,37
8	+	+	+	73	1,24
9	0	0	0	92	0,83
10	0	0	0	92	0,83
11	0	0	0	92	0,81
12	0	0	0	92	0,83
13	-1,68	0	0	106	0,77
14	0	-1,68	0	81	1,04
15	0	0	-1,68	134	0,94
16	1,68	0	0	72	0,71
17	0	1,68	0	93	1,06
18	0	0	1,68	76	0,69

A Figura 39 representa o cromatograma com maior quantidade de picos (Experimento 15), onde a melhor condição experimental foi obtida com a temperatura inicial e a temperatura final no ponto central e a taxa de aquecimento mantida no nível axial -1,68.



**Figura 39** - Cromatograma obtido do experimento número 15, onde a temperatura inicial e a temperatura final foram mantidas no ponto central e a taxa de aquecimento foi mantida no nível mínimo axial (-1,68). Coloque aqui a coluna e a vazão do gás de arraste.

A partir das respostas obtidas pelo CDD, foi construída uma tabela ANOVA para verificar se o modelo proposto está ajustado (Tabela 14)

**Tabela 14** - Análise de variância para o método cromatográfico.

Parâmetro	Soma quadrática	Graus de liberdade	Média quadrática	Valor de F	Valor de p
TI	1298,77	1	1298,77	114,23	0,000
TF	340,39	1	340,39	29,94	0,000
Taxa	3033,65	1	3033,65	266,81	0,000
TI x TI	46,85	1	46,85	4,12	0,070
TF x TF	90,80	1	90,80	7,99	0,018
Taxa x Taxa	214,04	1	214,04	18,82	0,001
TI x TF	4,50	1	4,50	0,40	0,543
TI x Taxa	18,00	1	18,00	1,58	0,237
TF x Taxa	18,00	1	18,00	1,58	0,237
Regressão	5103,50	9	567,06	49,87	0,000

<b>Resíduos</b>	113,70	10	11,37		
<b>Falta de ajuste</b>	113,70	5	22,74		
<b>Erro puro</b>	0,00	5	0,00		
<b>Total</b>	5217,20	19			
<b>Varição explicada (R<sup>2</sup>)</b>	0,9782				
<b>Máxima variação explicada (R<sup>2</sup>)</b>	1,000				

A análise do Teste *F* demonstra que o modelo de regressão para o método cromatográfico desenvolvido foi significativo, visto que o valor calculado de *F* (49,89) é superior ao valor de *F* tabelado (3,02), a um intervalo de confiança de 95%.

Na Tabela 15 constam todos os coeficientes da equação de regressão para o modelo proposto de otimização da quantidade de picos cromatográficos, que são TI, TF, taxa de aquecimento e as interações entre eles. Pode-se notar pela Tabela 15 que todos os parâmetros estudados nesta parte de refino do método são estatisticamente significativos a um intervalo de confiança de 95%, pois apresentam valores de *p* menores que 0,05.

**Tabela 15 - Coeficientes codificados e efeitos.**

<b>Parâmetro</b>	<b>Efeitos</b>	<b>Coeficiente</b>	<b>IC de 95%</b>	<b>Valor-<i>p</i></b>
<b>Constante (b<sub>0</sub>)</b>	-	92,11	(89,05; 95,18)	0,000
<b>TI</b>	-19,504	-9,752	(-11,785; -7,719)	0,000
<b>TF</b>	9,985	4,992	(2,959; 7,026)	0,000
<b>Taxa</b>	-29,808	-14,904	(-16,937; -12,871)	0,000
<b>TI x TI</b>	-3,606	-1,803	(-3,782; 0,176)	0,070
<b>TF x TF</b>	-5,020	-2,510	(-4,489; -0,531)	0,018
<b>Taxa x Taxa</b>	7,708	3,854	(1,875; 5,833)	0,001
<b>TI x TF</b>	-1,50	-0,75	(-3,41; 1,91)	0,543
<b>TI x Taxa</b>	3,00	1,50	(-1,16; 4,16)	0,237
<b>TF x Taxa</b>	-3,00	-1,50	(-4,16; 1,16)	0,237

Desta forma, foi obtida a equação de regressão para a resposta (número de picos cromatográficos) (equação 41).



$$\text{Resposta} = 92,11 - 9,752 \text{ TI} + 4,992 \text{ TF} - 14,904 \text{ Taxa} - 1,803 \text{ TI}^2 - 2,510 \text{ TF}^2 + 3,854 \text{ Taxa}^2 - 0,75 \text{ TI} \cdot \text{TF} + 1,50 \text{ TI} \cdot \text{Taxa} - 1,50 \text{ TF} \cdot \text{Taxa} \text{ (equação 41)}$$

Como os valores de  $p$  devem ser menores que 0,05 para que o valor do parâmetro seja significativo num intervalo de confiança de 95%, de maneira que haja contribuição com a equação da resposta, foram retirados os valores de  $\text{TI}^2$ ,  $\text{TI} \cdot \text{TF}$ ,  $\text{TI} \cdot \text{Taxa}$  e  $\text{TF} \cdot \text{Taxa}$ , e, com isso, foi criado um novo modelo. A Tabela ANOVA deste modelo recalculado está presente na Tabela 16.

**Tabela 16** - Análise de variância para o novo modelo calculado para o método cromatográfico.

Parâmetro	Soma quadrática	Graus de liberdade	Média quadrática	Valor de F	Valor de $p$
TI	1298,77	1	1298,77	90,44	0,000
TF	340,39	1	340,39	23,70	0,000
Taxa	3033,65	1	3033,65	211,25	0,000
TF x TF	106,61	1	79,09	5,51	0,034
Taxa x Taxa	236,72	1	236,72	16,48	0,001
Regressão	5016,15	5	1003,23	69,86	0,000
Resíduos	201,05	14	14,36		
Falta de ajuste	201,05	9	22,34		
Erro puro	0,00	5	0,00		
Total	5217,20	19			
Varição explicada ( $R^2$ )	0,9615				
Máxima variação explicada ( $R^2$ )	1,000				

A análise do Teste F demonstra que o modelo de regressão para a obtenção de maior número de picos cromatográficos foi significativa, pois o valor calculado de F (69,86) é maior que o valor de F tabelado (2,96), com 95% de confiança, apresentando um coeficiente de determinação  $R^2$  igual a 0,9615, o que indica que 96% das variâncias da obtenção de número de picos pode ser explicada pelo modelo quadrático construído.

É importante observar que o valor do coeficiente de determinação diminuiu neste novo modelo, mas o valor de F calculado e sua relação com o F tabelado para a regressão aumentou, o que mostra que ele está melhor ajustado, além de continuar consistente e confiável, e, portanto, não há necessidade de novos ajustes.

Na Tabela 17 constam todos os coeficientes da equação de regressão para o novo modelo calculado.

**Tabela 17** - Coeficientes codificados e efeitos para o novo modelo calculado do método cromatográfico.

Parâmetro	Efeitos	Coef	IC de 95%	Valor-p
<b>Constante (b<sub>0</sub>)</b>	-	90,64	(87,82; 93,45)	0,000
<b>TI</b>	-19,504	-9,75	(-11,95; -7,55)	0,000
<b>TF</b>	9,985	4,99	(2,79; 7,19)	0,000
<b>Taxa</b>	-29,808	-14,90	(-17,10; -12,70)	0,000
<b>TF x TF</b>	-4,662	-2,331	(-4,462; -0,201)	0,034
<b>Taxa x Taxa</b>	8,066	4,033	(1,902; 6,163)	0,001

A equação obtida para resposta (número de picos cromatográficos) do novo modelo, fica sendo (equação 42)

$$\text{Resposta} = 90,64 - 9,75 \text{ TI} + 4,99 \text{ TF} - 14,90 \text{ Taxa} - 2,331 \text{ TF}^2 + 4,033 \text{ Taxa}^2$$

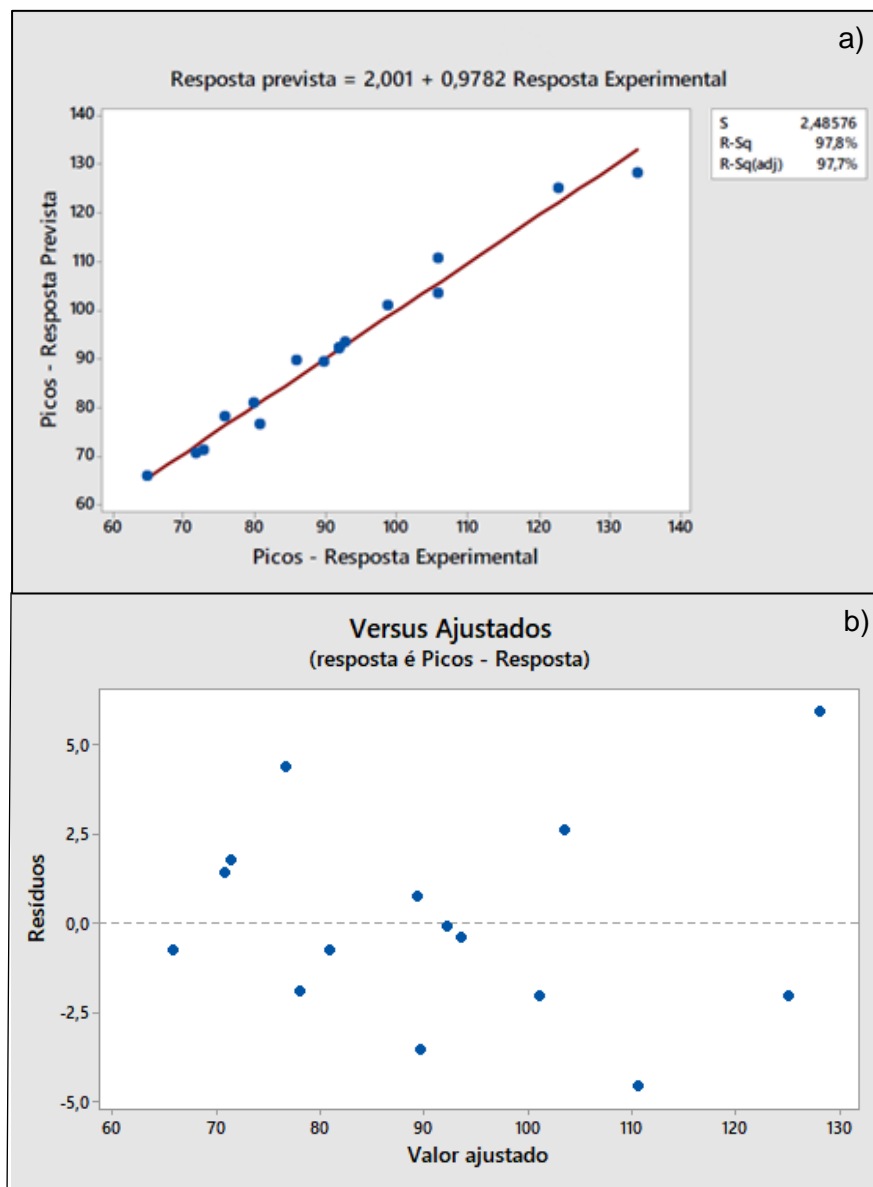
(equação 42)

A qualidade do ajuste do modelo também pode ser verificada através da resposta prevista *versus* a resposta experimental (Tabela 18 e Figura 40). Observa-se boa concordância entre as respostas experimentais e as previstas pelo modelo, sendo a relação entre os resultados aleatória, o que confirma o ajuste do modelo proposto ser adequado.

**Tabela 18** - Valores da resposta experimental e da resposta prevista pelo modelo de regressão.

<b>Resposta Experimental (Y)</b>	<b>Resposta Prevista (<math>\hat{Y}</math>)</b>
72	70
106	110
92	92
92	92
99	101
106	103
65	65
93	93
92	92
80	80
90	89
92	92
134	128
73	71
86	89
123	125
81	77
76	77

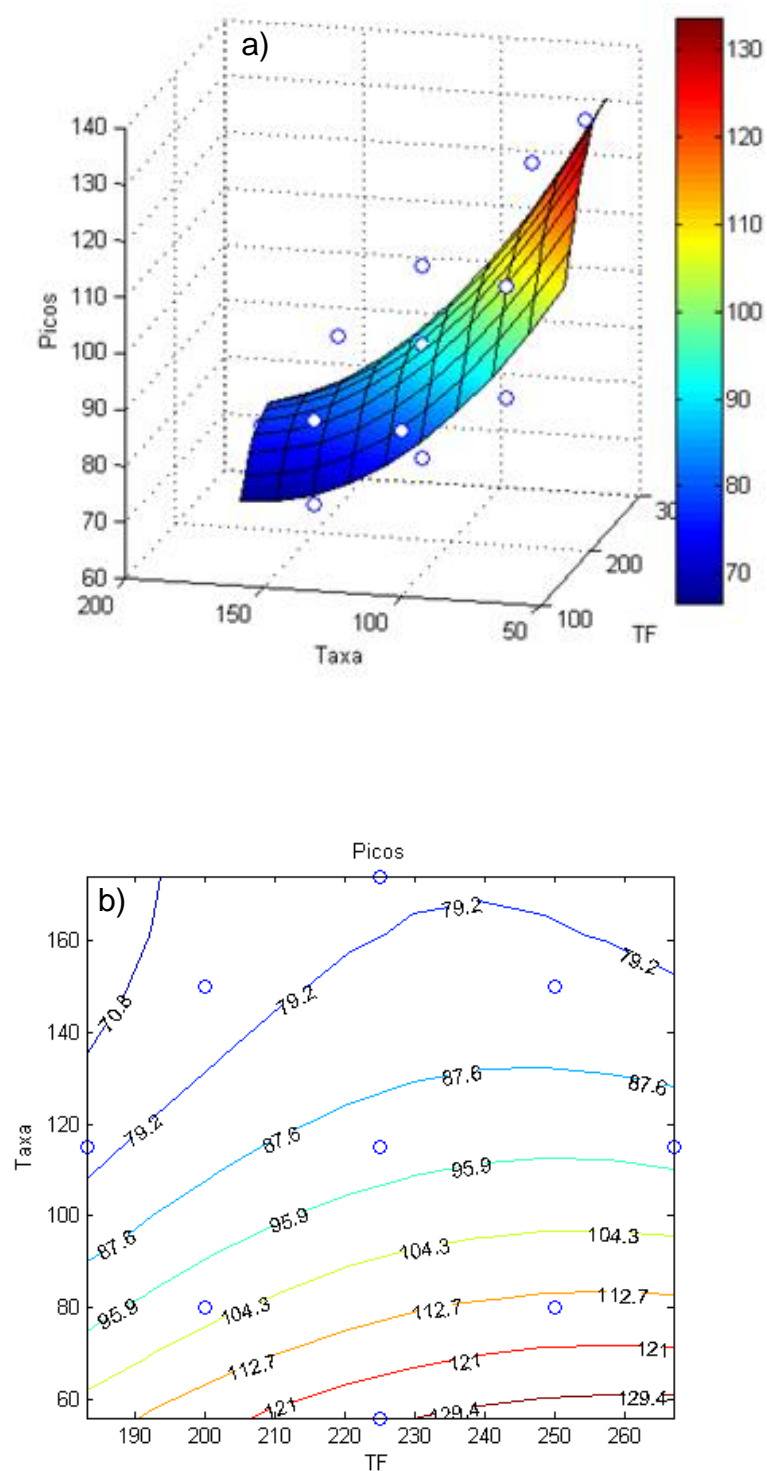
O gráfico com comparação das respostas experimental e prevista para o modelo de regressão visando a obtenção de maior número de picos cromatográficos (Figura 40a) resultou na equação Resposta Prevista = 2,001 + 0,9782 Resposta Experimental. O  $R^2$  para esta regressão linear é igual a 0,978 e percebe-se a boa correlação entre estas respostas. Outro parâmetro observado foi o gráfico de resíduos dos valores preditos (Figura 40b), uma vez que, se os resíduos se apresentam de forma aleatória, ou seja, são homocedásticos, as hipóteses dos resíduos são satisfatórias, podendo-se concluir que os resíduos do modelo ajustado seguem uma distribuição normal sem tendência.



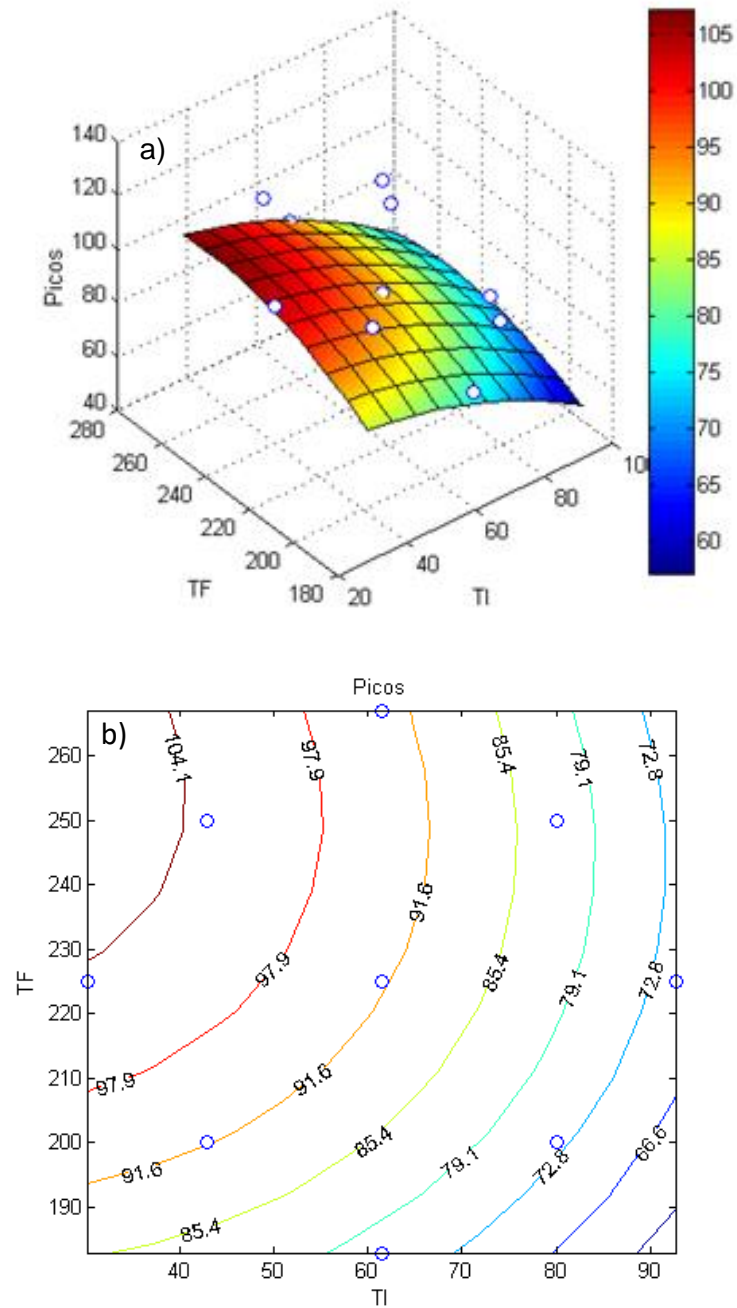
**Figura 40** - Gráfico com comparação das respostas experimental e prevista (a) e gráfico de análise de resíduos(b).

Como o objetivo da aplicação de planejamento experimental e modelo de otimização da resposta é gerar um cromatograma com o maior número possível de picos originados da separação cromatográfica do óleo diesel, foram observados os gráficos de superfície de contorno e de superfície de resposta gerados pelo modelo proposto, empregando dois parâmetros por vez para visualizar a interação entre eles e o efeito resultante no número de picos. A Figura 41 mostra a interação entre temperatura final e taxa de aquecimento, a Figura 42 mostra a interação entre a temperatura final e a temperatura inicial, enquanto a Figura 43 mostra a interação entre a temperatura inicial e a taxa de aquecimento.

Observando os gráficos da Fig. 41, pode-se notar que para um valor fixo da TI, quanto menor for a taxa de aquecimento e maior for a temperatura final, maior será o número de picos obtidos.

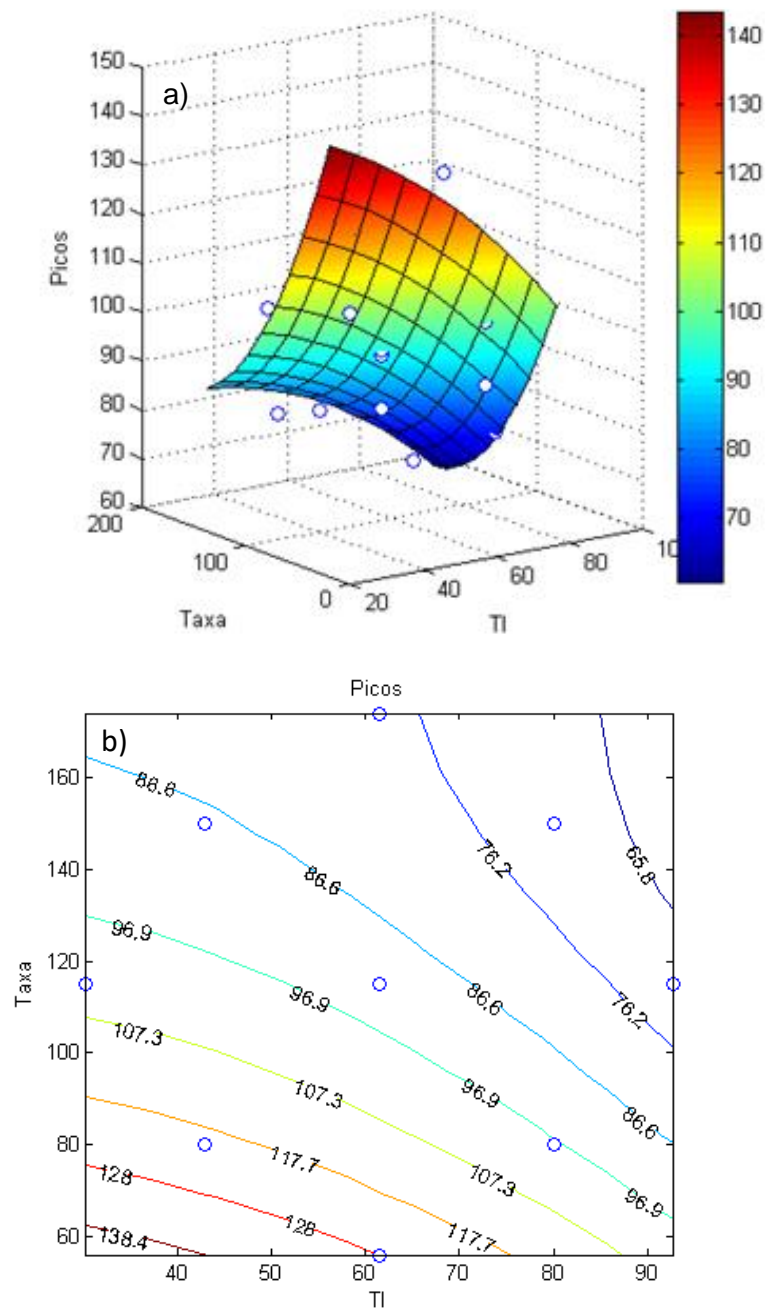


**Figura 41** - Superfície de resposta (a) e Gráfico de contorno (b) de Resposta vs TF; Taxa de aquecimento.



**Figura 42** - Superfície de resposta (a) e Gráfico de contorno (b) de Resposta vs TF; TI.

Avaliando estes gráficos (Fig. 42), é possível notar que, mantendo a taxa de aquecimento em um valor fixo, a medida em que a TF aumenta e a TI diminui, melhor será a resposta desejada, ou seja, um aumento do número de picos cromatográficos.

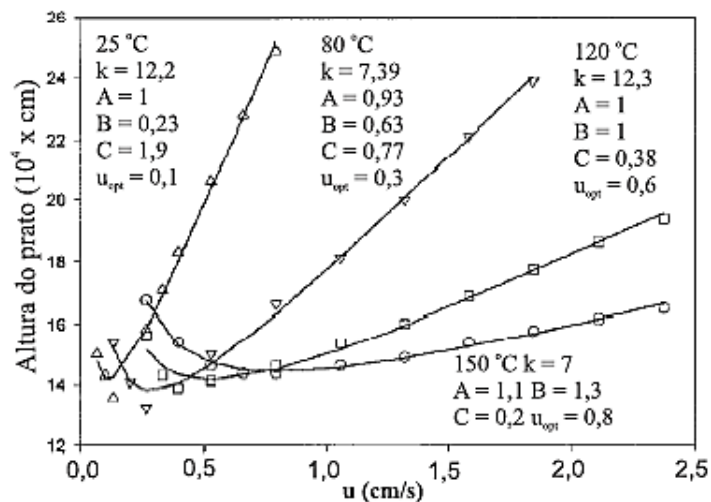


**Figura 43** - Superfície de resposta (a) e Gráfico de contorno (b) de Resposta vs TI; Taxa de aquecimento.

Nestes gráficos (Fig. 43), é possível observar que, para um valor fixo de temperatura final, quanto menor for a taxa de aquecimento e a temperatura inicial, maior será a resposta com o número de picos obtidos.

A temperatura final ser um fator importante e quanto maior for seu valor, melhor ser a resposta está de acordo com o esperado em CG-UR, visto que as vantagens obtidas com o aumento da temperatura frente à redução do tempo de análise são descritas pelas curvas de van Deemter (Figura 44). O aumento da temperatura facilita

a transferência de massa entre a coluna e o gás de arraste, reduzindo a termo C da equação de van Deemter (equação 43), permitindo o uso de velocidades lineares maiores que a velocidade linear ótima, em detrimento de apenas um pequeno aumento na altura de prato teórico. A difusão longitudinal (termo B) aumenta com o aumento da temperatura, fazendo com que a velocidade linear ótima se desloque para valores mais elevados, já que em baixas velocidades lineares as curvas de van Deemter são dominadas pela difusão longitudinal (BORGES, BOTOLLI, COLLINS, 2010). A Figura 44 exemplifica a relação da curva de van Deemter com o aumento da temperatura final do método.



**Figura 44** - Curvas de van Deemter (altura do prato vs velocidade linear) ilustrando como o aumento da temperatura auxilia na velocidade linear ótima, permitindo separações mais rápidas.

**Fonte:** Borges, Botolli, Collins (2010).

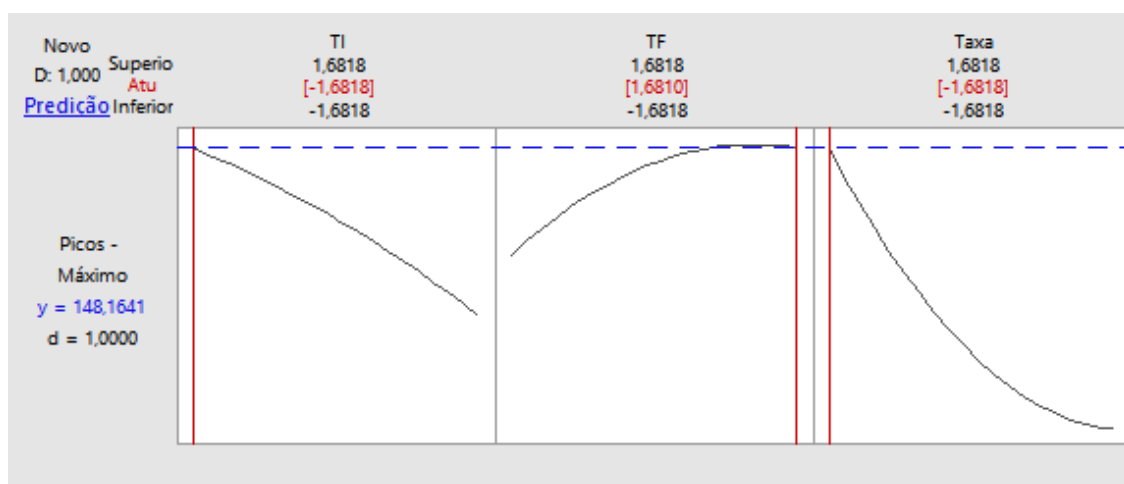
$$H = A + \frac{B}{\bar{u}} + C_s \bar{u} + C_m \bar{u} \quad (\text{equação 43})$$

A temperatura inicial e a taxa de aquecimento serem eficientes no menor valor possível condiz com a complexidade da mistura que consiste o óleo diesel, onde lentamente os componentes vão atingindo seus pontos de ebulição na pressão interna do cromatógrafo e são mais facilmente separados, resultando em um maior número de picos cromatográficos identificados.

Com o modelo obtido e de acordo com os gráficos de contorno e de superfície de resposta, foi possível conhecer a desejabilidade do planejamento experimental realizado (Figura 45). Segundo esta desejabilidade, a melhor condição do método pode ser obtida mantendo o parâmetro temperatura inicial no nível axial -1,68, o



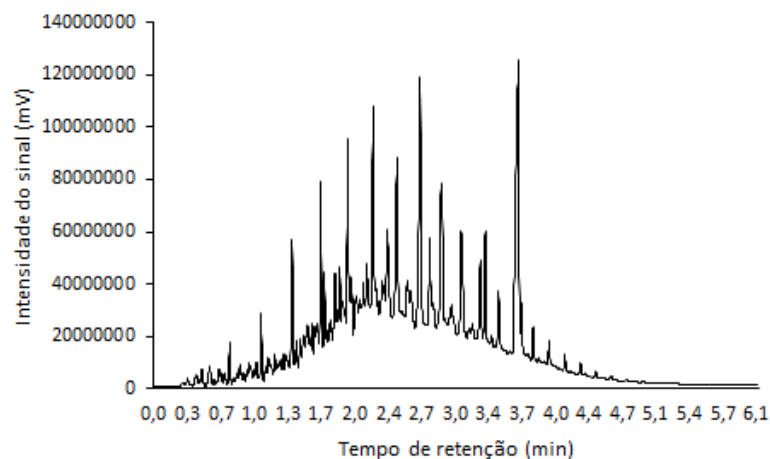
parâmetro temperatura final no nível axial 1,68 e o parâmetro taxa de aquecimento no nível axial -1,68. Em valores numéricos, a desejabilidade seria dada por temperatura inicial igual a 30 °C, temperatura final igual a 297 °C e taxa de aquecimento igual a 56 °C.min<sup>-1</sup>. Segundo o modelo, esse experimento produziria um cromatograma com a resposta maximizada de 148 picos, aproximadamente.



**Figura 45** - Desejabilidade da melhor condição do método para gerar o maior número de picos. Em vermelho, estão as condições consideradas desejáveis.

Foi realizado um experimento nas condições indicadas de desejabilidade do método. O cromatograma resultante está representado na Figura 46 e obteve como resposta, em triplicata, 163 picos. Este resultado não foi concordante com o resultado sugerido pelo modelo, o que pode indicar que, para uma previsão de resposta nessa faixa de valores, ele perde a linearidade e não funciona corretamente, tendo sua aplicabilidade apenas para os valores mostrados na Tabela 18 e Figura 44.

De acordo com os gráficos de contorno demonstrados nas figuras 41, 42 e 43, não há uma região específica que aponte o máximo da resposta do modelo, mas há regiões de ascendência (TI vs TF; TF vs Taxa) e região de descendência (TI vs Taxa) que indicam o caminho para um maior valor de número de picos, assim, quanto maior for a temperatura final e menor forem a taxa de aquecimento e temperatura inicial, maior então será o número de picos cromatográficos obtidos, corroborando com a equação 42, onde o coeficiente TF apresenta um valor positivo e os coeficientes Taxa e TI possuem valores negativos.



**Figura 46** - Cromatograma obtido do experimento reproduzindo as condições de desejabilidade, onde a temperatura inicial e a taxa de aquecimento foram mantidas no menor nível (-1,68) e a temperatura final foi mantida no maior nível (1,68). A resolução global é de 1,08.

A desejabilidade (Figura 45) também corrobora com estes dados, pois indicou a maximização do modelo utilizando os menores valores experimentais para temperatura inicial e taxa de aquecimento (codificados em -1,68) e o maior valor experimental para a temperatura final (codificado em 1,68).

Com isto, foram feitos experimentos de ascendência e descendência, em triplicatas, com a finalidade de explorar estes parâmetros e suas interações, no menor tempo possível, visto que este é um dos objetivos do trabalho. Analisando a Figura 46, percebe-se que nenhum pico elui após 5,2 minutos, o que é condizente com a metodologia da cromatografia gasosa ultrarrápida, e então, os novos experimentos foram delineados para terem um tempo total em torno de 5 minutos.

A Tabela 19 mostra os experimentos de exploração da superfície de resposta realizados com os respectivos valores de temperatura inicial, temperatura final e taxa de aquecimento.

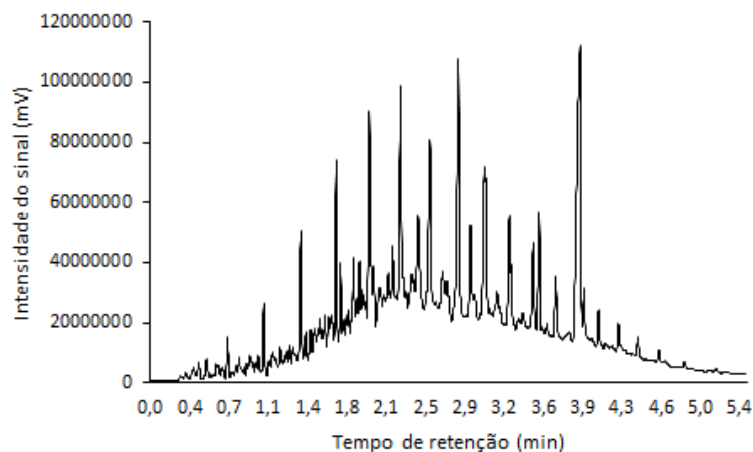
**Tabela 19** - Experimentos de exploração da superfície de resposta (ascendência e descendência).

Experimento	TI (°C)	TF (°C)	Taxa (°C/min)	Tempo (min)	Picos $\pm$ s	Resolução global
Asc/Desc 1	35	240	56	4,96	162 $\pm$ 1	0,91
Asc/Desc 2	40	220	50	4,90	156 $\pm$ 0	0,86
Asc/Desc 3	40	210	45	5,08	157 $\pm$ 0	0,51
Asc/Desc 4	35	210	50	4,80	154 $\pm$ 1	0,97
Asc/Desc 5	30	220	60	4,47	144 $\pm$ 2	1,02

Asc/Desc 6	30	240	60	4,80	150 ± 1	1,03
Asc/Desc 7	30	220	50	5,10	158 ± 1	1,03
Asc/Desc 8	30	240	54	5,19	155 ± 1	1,08
Asc/Desc 9	35	240	50	5,40	164 ± 0	1,09
Asc/Desc10	35	280	60	5,38	145 ± 1	0,87

O termo “Picos ± s”, presente nesta tabela é referente às medidas feitas em triplicata, onde s é o desvio padrão entre elas. Observando a Tabela 19 nota-se que o maior número de picos é observado, respeitando a minimização do tempo de análise, nas condições de menor taxa de aquecimento possível, não necessariamente mantendo a temperatura final e a temperatura inicial em suas otimizações, mas apenas não deixando a final muito baixa e a inicial muito alta, o que corrobora com a equação 42, onde nota-se que o valor do coeficiente que multiplica “Taxa” é maior em módulo.

Todas estas análises foram realizadas mantendo os demais parâmetros constantes: Split 1:100; 0,1 µL de volume de injeção de amostra, vazão igual a 0,5 mL/min; *hold time* inicial de 0,3 minutos, *hold time* final de 1,0 minuto, temperatura do bloco igual a 280 °C, temperatura do injetor igual a 270 °C e temperatura do detector a 280 °C. O experimento escolhido nesta etapa foi o Asc/Desc 9 por apresentar a maior quantidade de picos e resolução global em relação aos demais. A Figura 47 representa o cromatograma obtido com este experimento. Após estabelecidas as condições de temperatura final, temperatura inicial e taxa de aquecimento, foram feitos experimentos variando o volume de injeção de amostra e a taxa de split (Tabela 20).



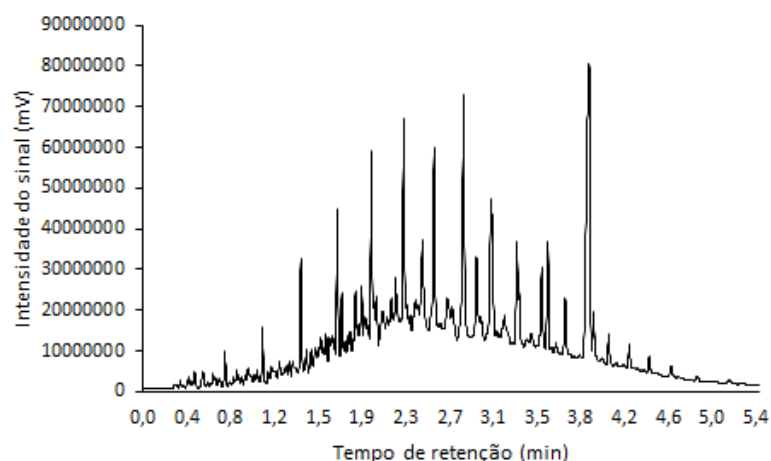
**Figura 47** - Cromatograma obtido do experimento “Asc/Desc 9”, onde a temperatura inicial foi igual a 30 °C, a taxa de aquecimento igual a 50 °C/min e temperatura final igual a 240 °C, com tempo total de análise igual a 5,40 minutos.

**Tabela 20** - Variação de volume de injeção e razão de split.

Experimento	Split	Volume de injeção (µL)	Picos ± s	Resolução Global
Refino 1	1:100	0,05	163 ± 1	0,99
Refino 2	1:100	0,1	164 ± 1	0,94
Refino 3	1:100	0,15	155 ± 1	0,52
Refino 4	1:100	0,2	155 ± 1	0,53
Refino 5	1:100	0,3	152 ± 0	1,23
Refino 6	1:100	0,5	150 ± 0	1,22
Refino 7	1:200	0,1	169 ± 0	1,08
Refino 8	1:150	0,1	167 ± 1	0,94
Refino 9	1:130	0,1	167 ± 1	0,90
Refino 10	1:70	0,1	159 ± 1	1,00

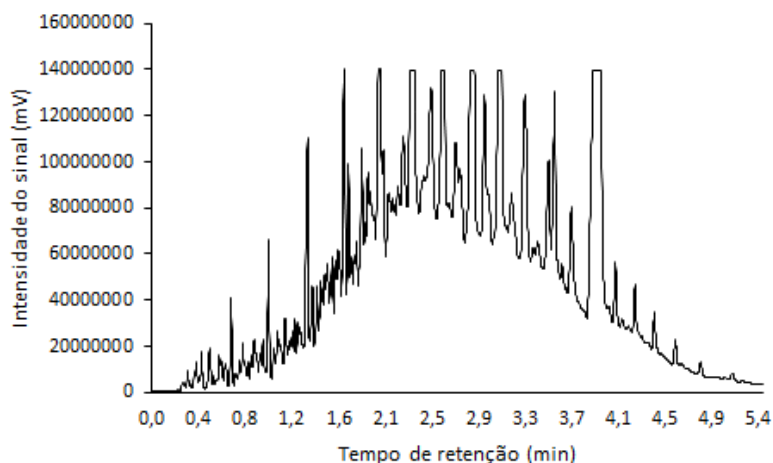
Esta etapa foi dividida em duas partes. Primeiramente, foram realizados experimentos mantendo o split 1:100 e variando o volume de injeção. O melhor resultado de volume de injeção foi 0,1 µL. Então, na segunda parte, manteve-se fixo o volume de injeção em 0,1 µL e variou-se a taxa de split. Assim, de acordo com os resultados obtidos na Tabela 20, para as posteriores análises, o volume de injeção

continuou sendo igual a 0,1  $\mu\text{L}$  e a razão de split passou a ser 1:200 (Figura 48), escolha realizada pela quantidade de picos conjunta a resolução global.



**Figura 48** - Cromatograma obtido do experimento “Refino 7”, onde a razão de split é igual a 1:200 e o volume de injeção é igual a 0,1  $\mu\text{L}$

Os experimentos Refino 3, Refino 4, Refino 5 e Refino 6 estouraram a escala do cromatograma (Figura 49, exemplo Refino 5) devido ao volume excedente injetado. A partir deste ponto, foram feitos experimentos para ajustar o melhor valor de *hold time* da temperatura inicial, ou seja, o tempo em que a temperatura inicial permanece constante antes de começar o aquecimento através da taxa definida. Experimentos realizados anteriormente ao planejamento experimental mostraram que este tempo auxilia o aumento do número de picos e, no início da otimização dos parâmetros ficou fixado em 0,3 minutos. Nos experimentos realizados, o *hold time* inicial variou de zero a um minuto. A descrição desta variação, cada tempo final da análise e seus respectivos picos cromatográficos encontram-se na Tabela 21.

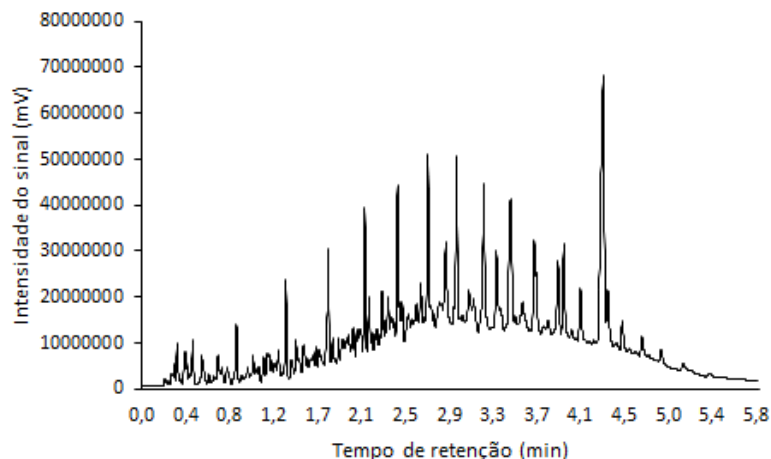


**Figura 49** - Cromatograma obtido do experimento "Refino 5", onde a razão de split é igual a 1:100 e o volume de injeção é igual a 0,3  $\mu$ L

**Tabela 21** - Variação do *hold time* inicial

<b>Experimento</b>	<b>Hold time (min)</b>	<b>Tempo final (min)</b>	<b>Picos <math>\pm</math> s</b>	<b>Resolução global</b>
<i>Hold Time 1</i>	0,0	5,10	138 $\pm$ 1	0,98
<i>Hold Time 2</i>	0,1	5,20	158 $\pm$ 2	0,98
<i>Hold Time 3</i>	0,3	5,40	169 $\pm$ 1	0,95
<i>Hold Time 4</i>	0,5	5,60	174 $\pm$ 1	0,96
<i>Hold Time 5</i>	0,7	5,80	183 $\pm$ 0	0,97
<i>Hold Time 6</i>	1,0	6,10	186 $\pm$ 0	0,97

Neste caso, levando em consideração a otimização do tempo e a pouca diferença de número de picos entre os experimentos *Hold Time 5* e *Hold Time 6*, de 3 unidades, o experimento *Hold Time 5* foi o que apresentou uma resposta mais interessante ao propósito. A Figura 50 mostra o cromatograma obtido nas condições deste experimento.



**Figura 50** - Cromatograma obtido do experimento *Hold Time 5*, onde a temperatura inicial, igual a 35 °C, foi mantida por 0,7 minutos. O tempo total desta análise foi igual a 5,80 minutos.

No primeiro planejamento experimental, a temperatura final teve seu valor no maior nível (+1) igual a 250 °C devido à temperatura máxima suportada pela coluna VF23MS (caráter polar) ser 260 °C. Nas etapas posteriores, mesmo estando definido que a coluna utilizada seria a PH SE54, manteve-se a temperatura de 250 °C, previamente definida, como referência para as próximas etapas. Entretanto, a coluna PH SE54 suporta até 350 °C e o diesel apresenta compostos de alta massa molecular com pontos de ebulição elevados. Ficou então definido aumentar a temperatura final para 340 °C, assim como a temperatura do detector. A temperatura do injetor não foi alterada, pois seu valor mantido a 270 °C, juntamente com a presença de um *hold time* inicial antes do aquecimento da fase estacionária, contribuiriam para a separação dos compostos com menor peso molecular e pontos de ebulição menores presentes no óleo diesel, contribuindo para um maior número de picos, além de o valor máximo de temperatura suportado pelo septo do injetor ser de 280 °C.

No intuito de aumentar bastante a temperatura final do método sem aumentar muito o tempo total de análise, foram testadas novas taxas de aquecimento, buscando explorar regiões do cromatograma onde há possibilidade de resolver mais picos.

Nestes experimentos, a temperatura inicial foi mantida em 35 °C, de acordo com otimizações anteriores, a taxa de aquecimento foi mantida com o valor mais baixo possível que não prejudicasse o tempo até o valor de temperatura onde eluem mais picos (280 °C) e a segunda taxa de aquecimento foi avaliada nos mais altos valores possíveis, para otimizar tempo de análise e número de picos, simultaneamente, ou seja, sem perder eficiência de separação. Foi observada a importância de estudar o

*hold time* na temperatura final antes de encerrar a análise, pois alguns picos não estavam eluindo e permaneciam na coluna, saindo na análise posterior, gerando efeito memória e alterando o resultado entre repetições. Visto a necessidade de minimizar o tempo total, o *hold time* inicial foi mantido a 0,1 minuto para esta otimização. Deixar a primeira taxa de aquecimento menor em detrimento a um *hold time* inicial também menor foi mais eficiente para o objetivo proposto, visto que a taxa de aquecimento é o parâmetro que mais influencia a maximização da resposta analítica. Desta forma, a Tabela 22 descreve os experimentos desta etapa e os resultados.

**Tabela 22** - Experimentos adicionando mais taxas de aquecimento.

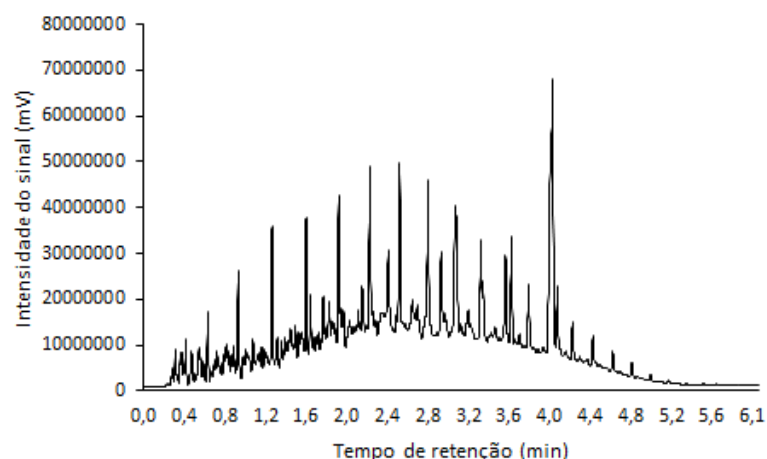
Exp.	TI (°C)	H.T.I (min)	Taxa (°C/min)	T (°C)	H.T (min)	Taxa (°C/min)	TF (°C)	H.T.F. (min)	t (min)	Picos ± s	R.G.
1	35	0,1	45	280	0,0	800	340	0,5	6,12	187 ± 1	1,09
2	35	0,1	45	280	0,0	600	340	0,5	6,14	183 ± 2	0,90
3	35	0,1	45	280	0,0	300	340	0,5	6,24	182 ± 0	0,81
5	35	0,1	45	280	0,0	800	340	1,0	6,62	187 ± 0	1,09
6	35	0,1	45	280	0,0	800	340	0,7	6,32	187 ± 1	1,08
7	35	0,1	45	280	0,0	800	340	1,5	7,12	187 ± 0	1,09

Exp.: Experimento; TI: Temperatura inicial; H.T.I.: Hold time inicial; T: Temperatura; H.T. Hold time; TF: Temperatura final; H.T.F.: Hold time final; t: tempo e R.G.: Resolução global.

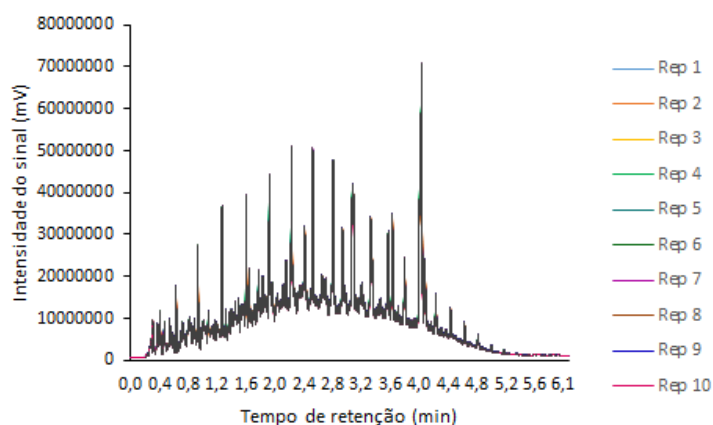
Como pode ser observado na Tabela 22, o experimento 1 foi o que melhor atingiu o objetivo proposto, de menor tempo em relação aos outros com número de picos e resolução global semelhantes. Um tempo de análise maior que 5 minutos não é muito interessante para cromatografia gasosa ultrarrápida, entretanto, para uma amostra complexa como o óleo diesel, com a finalidade de garantir uma análise sem picos espúrios e variações brutas entre as replicatas, foi necessário um tempo de 6,12 minutos, o qual não compromete a aplicabilidade de técnica. A Figura 51 apresenta o cromatograma correspondente ao experimento 1 e a Figura 52 apresenta a sobreposição de 10 replicatas desta condição representada no experimento 1. Apesar do *hold time* na temperatura final do método desenvolvido, as replicatas apresentam diferença na quantidade de picos (187, 185, 186, 187, 187, 187, 188, 188, 185, 186; média igual a 187 e desvio padrão igual a 1,02). Esta diferença é pequena e se deve possivelmente ao fato de a separação desta amostra complexa não ser total, havendo



muita coeluição, de forma que o *software* reconheça numa dessas coeluições, por exemplo, um único pico e em outra análise, reconheça dois ou mais, devido a variações normais da exatidão do tempo de retenção em cromatografia.



**Figura 51** - Cromatograma obtido do experimento “1”, onde a temperatura inicial foi igual a 35 °C, mantida por 0,1 minuto e numa taxa de 45 °C/min até 280 °C e então numa taxa de aquecimento igual a 800°C/min até temperatura final igual a 340°C que se mantém por 0,5 minuto, com tempo total de análise igual a 6,12 minutos.



**Figura 52** - Replicatas ( $n = 10$ ) sobrepostas de cromatogramas obtidos do experimento “1”, onde a temperatura inicial foi igual a 35 °C, mantida por 0,1 minuto e numa taxa de 45 °C/min até 280 °C e então numa taxa de aquecimento igual a 800 °C/min até temperatura final igual a 340 °C que se mantém por 0,5 minuto, com tempo total de análise igual a 6,12 minutos.

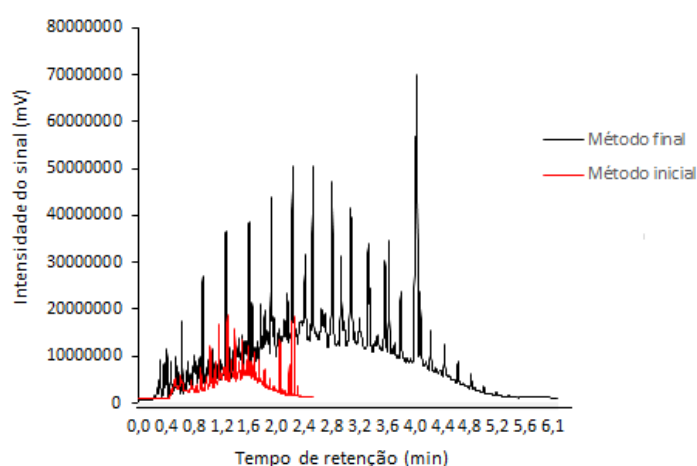
O método desenvolvido se mostrou promissor no objetivo proposto de apresentar o maior número de picos cromatográficos possível em um tempo reduzido de análise. O valor da resolução global (1,09) é razoável, estando dentro do intervalo considerado adequado  $1 \leq R_s \leq 1,5$  (SNYDER, 1997; FLUMIGNAN, 2005). Devido à complexidade da composição do óleo diesel, uma separação cromatográfica com

resolução global não muito alta e muitos picos coeluídos e assimétricos já era esperada.

Por fim, comparando o método inicial com o aperfeiçoado (Tabela 23 e Figura 53), conclui-se que a otimização do método permitiu que com um aumento de aproximadamente 3,3 minutos no tempo de análise fossem identificados um número 3,17 vezes maior de picos cromatográficos. Não houve variação significativa na resolução global, entretanto, a área total dos cromatogramas apresentou um aumento de 15,38 vezes.

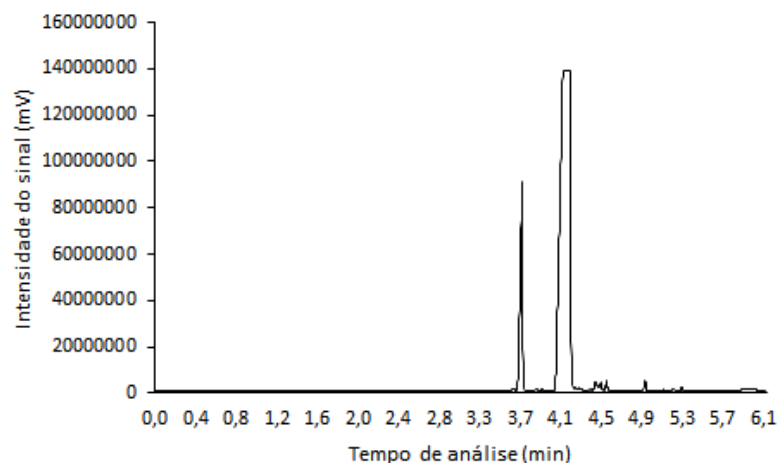
**Tabela 23** - Comparação dos métodos cromatográficos antes e depois da otimização.

Metodologia	Tempo (min)	Picos	R.G.	Área total
Inicial	2,85	59	1,08	237114828
Final	6,12	187	1,09	3647657643



**Figura 53** - Comparação dos cromatogramas obtidos nas condições de análise inicial (em vermelho) e final (em preto).

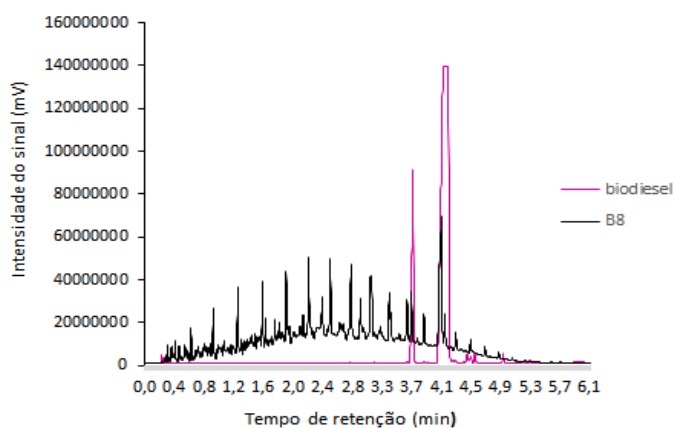
Uma amostra de biodiesel de soja puro foi cedida pelo CEMPEQC e avaliada pelo método cromatográfico aperfeiçoado para elucidar em quais tempos de retenção seus ésteres metílicos de ácidos graxos (*fatty acid methyl ester* – FAME) componentes estavam sendo eluídos (Figura 54).



**Figura 54** - Cromatograma obtido de amostra de biodiesel de soja puro utilizando as condições do método cromatográfico otimizado.

Conforme pode ser observado nesta figura, há dois picos no cromatograma correspondentes à presença da mistura de biodiesel no diesel, que são aproximadamente 3,7 minutos e 4,1 minutos, sendo que este último provavelmente corresponde a uma mistura de ésteres não separadas pelo método cromatográfico, visto a intensidade do sinal. A composição do biodiesel de óleo de soja se dá majoritariamente por ácido linoleico (C18:2), correspondente a mais da metade do total de FAMES. Há também a presença de ácido palmítico (C16:2), ácido esteárico (C18:0), ácido oleico (C18:1) e ácido linolênico (C18:3) (FERRARI, OLIVEIRA e SCABIO, 2005; ROCHA *et al*, 2008). Possivelmente, o pico cromatográfico mais intenso é composto por todos os FAMES de 18 carbonos.

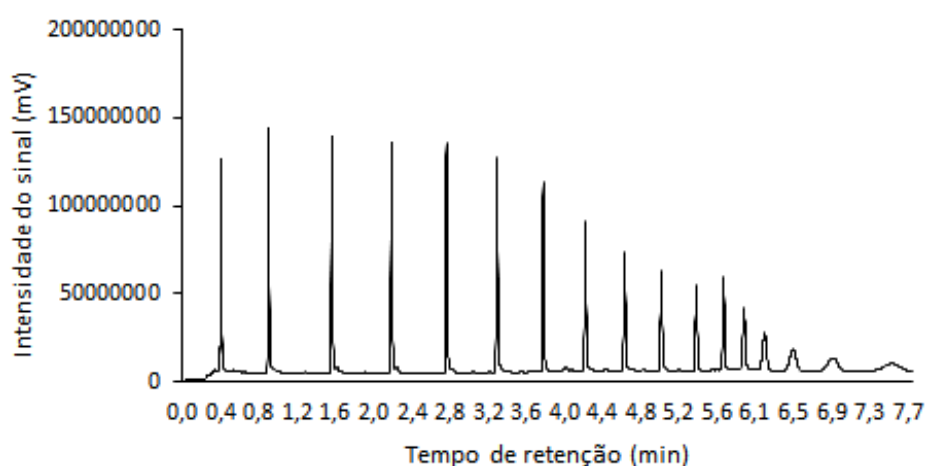
A Figura 55 contém a sobreposição dos picos de biodiesel puro ao cromatograma da amostra de mistura B8.



**Figura 55** - Comparação dos cromatogramas obtidos entre amostra de biodiesel de soja e a amostra de diesel comercial B8 utilizada no desenvolvimento do método de separação.

A separação cromatográfica dos 17 hidrocarbonetos parafínicos da série homóloga (C<sub>8</sub>-C<sub>40</sub>) foi realizada nas condições descritas do método otimizado com um aumento no valor do hold time final para 2,9; resultando em uma análise cromatográfica de 8 minutos, com o objetivo da observação de todos os analitos (Figura 56 e Tabela 24).

É possível concluir por esta figura, que o método de separação por CG-UR otimizado abrange toda a faixa de hidrocarbonetos correspondentes ao óleo diesel (C<sub>8</sub>-C<sub>28</sub>), sendo possível visualizar por este método até o C<sub>32</sub> com o tempo de análise. Essa mesma série homóloga foi utilizada para validação do método.



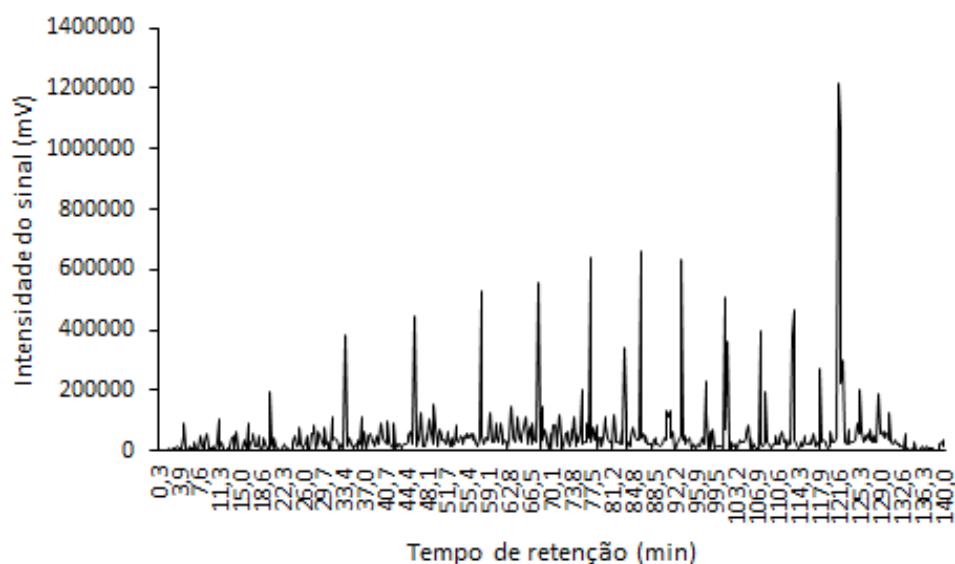
**Figura 56** - Cromatograma obtido ao analisar a série homóloga parafínica (C<sub>8</sub>-C<sub>40</sub>) pelo método cromatográfico otimizado.

**Tabela 24** - Tempo retenção (min) de cada hidrocarboneto da série homóloga parafínica correspondente à faixa do óleo diesel.

Hidrocarbonetos	Tempo de retenção (min)
C <sub>8</sub> H <sub>18</sub>	0,42
C <sub>10</sub> H <sub>22</sub>	0,88
C <sub>12</sub> H <sub>26</sub>	1,67
C <sub>14</sub> H <sub>30</sub>	2,19
C <sub>16</sub> H <sub>34</sub>	2,75
C <sub>18</sub> H <sub>38</sub>	3,39
C <sub>20</sub> H <sub>42</sub>	3,79
C <sub>22</sub> H <sub>46</sub>	4,32
C <sub>24</sub> H <sub>50</sub>	4,71
C <sub>26</sub> H <sub>54</sub>	5,03
C <sub>28</sub> H <sub>58</sub>	5,49

### 5.1.2 Análise da amostra de diesel comercial por cromatografia gasosa convencional

Paralelamente, foi realizada uma análise da mesma amostra de óleo diesel B8 em um cromatógrafo convencional, *Shimadzu*, modelo GC-17A, com auto mostrador AOC20i e software *GCSolution* (Shimadzu, Kyoto, Japão). A análise foi feita utilizando a coluna DB-Petro da *J&W Scientific* (50m x 0,15 mm x 0,50  $\mu$ m) com razão de split 1:100, injetor a 300 °C, 0,4  $\mu$ L de volume de injeção, gás de arraste hélio a uma vazão igual a 1,56 mL/min, temperatura do forno: 50 °C a 2 °C/min até 160 °C a 4 °C/min a 220 a 12 °C/min até 320 °C, mantendo por 30 minutos, totalizando um tempo de análise igual 140 minutos. O detector FID foi mantido a 300 °C. Este método foi desenvolvido por Dos Santos e colaboradores (2012) e resultou em 529 picos cromatográficos e resolução global 1,62. (Figura 57).



**Figura 57** - Cromatograma obtido em cromatógrafo a gás convencional, onde a análise foi feita utilizando a coluna DB-Petro da J&W Scientific com temperatura inicial igual a 50 °C com taxa de 2 °C/min até 160 °C e taxa de 4 °C/min até 220 °C com taxa de aquecimento de 12 °C/min até 320 °C, mantendo por 30 minutos, totalizando um tempo de análise igual 140 minutos. O detector FID foi mantido a 300 °C.

Apesar de apresentar maior número de picos e de resolução global, no método por CG convencional a razão de picos por minuto foi igual a 4,22; enquanto no método desenvolvido por CG-UR foi de 31,17. O método de CG-UR mostrou que apesar do tempo reduzido, não houve perda de informação, já que é possível identificar todos os hidrocarbonetos na faixa do óleo diesel neste método, bem como o biodiesel. A área total do cromatograma obtido por CG convencional foi de 220210133, um valor menor que o da área obtida pelo cromatograma gerado por CG-UR (3647657643),

entretanto, nos cromatogramas por CG-UR há uma lacuna entre a linha de base e o perfil cromatográfico que interfere no valor real de área. Esta lacuna/*drift* não foi encontrada no método convencional.

## 5.2 Validação analítica do método cromatográfico

Na validação do método cromatográfico, foram escolhidos os 11 hidrocarbonetos que correspondem à faixa de destilação do petróleo equivalente ao óleo diesel (C<sub>8</sub>-C<sub>28</sub>) para monitorar. A série homóloga foi analisada da forma descrita anteriormente, com tempo final de 8 minutos e último hidrocarboneto (C<sub>40</sub>) eluindo em aproximadamente 7,5 minutos. Isso foi feito para garantir a eluição de toda a sua composição e não permanecer traços na coluna, a fim de não interferir na repetibilidade do método.

Os resultados dos experimentos realizados na determinação de cada parâmetro de validação do método desenvolvido serão descritos a seguir.

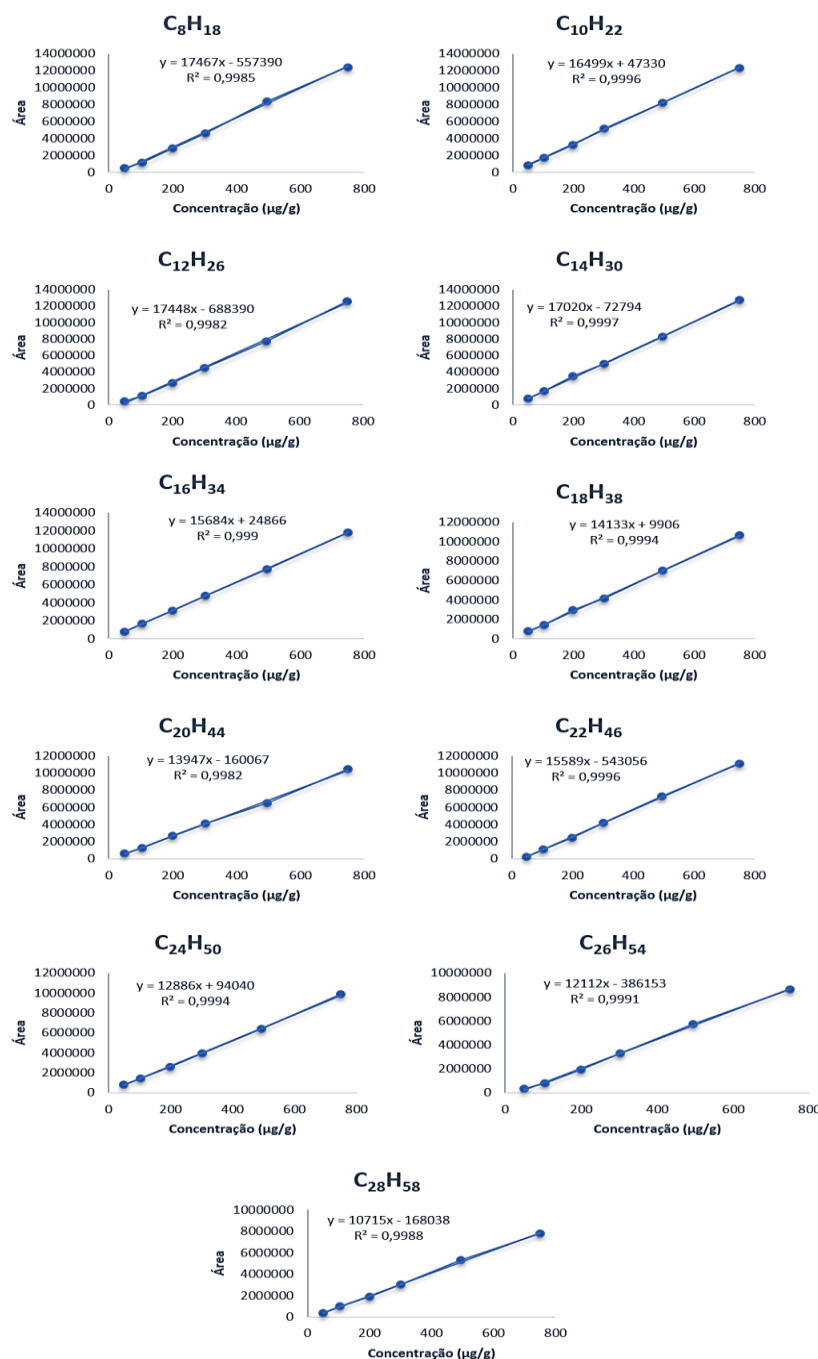
### 5.2.1 Linearidade

A linearidade do método foi determinada através de uma curva analítica preparada com 6 pontos de diferentes concentrações, avaliados em triplicata. O resultado para cada hidrocarboneto analisado está presente na Figura 58 e na Tabela 25.

Foram realizados estudos de verificação de homocedasticidade pelo teste de Cochran e de adequação do modelo linear e falta de ajuste pela ANOVA, para 95% de confiança. Os resultados mostram que todas as curvas são homocedásticas, o modelo linear é adequado e não sofrem falta de ajuste. Estes resultados se encontram no APÊNDICE A e no APÊNDICE B, respectivamente.

**Tabela 25** - Equação de regressão da curva analítica e seus parâmetros para cada hidrocarboneto.

$y = ax + b$				
<b>HIDROCARBONETOS</b>	<b>a</b>	<b>b</b>	<b>R<sup>2</sup></b>	<b>r</b>
$C_8H_{18}$	17467	-557390	0,9985	0,9992
$C_{10}H_{22}$	16499	47330	0,9996	0,9998
$C_{12}H_{26}$	17448	-688390	0,9982	0,9991
$C_{14}H_{30}$	17020	-72794	0,9997	0,9998
$C_{16}H_{34}$	15684	24866	0,9990	0,9995
$C_{18}H_{38}$	14133	9906	0,9994	0,9997
$C_{20}H_{42}$	13947	-160067	0,9982	0,9991
$C_{22}H_{46}$	15589	-543056	0,9996	0,9998
$C_{24}H_{50}$	12886	94040	0,9994	0,9997
$C_{26}H_{54}$	12112	-386153	0,9991	0,9995
$C_{28}H_{58}$	10715	-168038	0,9988	0,9994



**Figura 58** - Curvas analíticas de cada hidrocarboneto para elucidação do parâmetro linearidade.

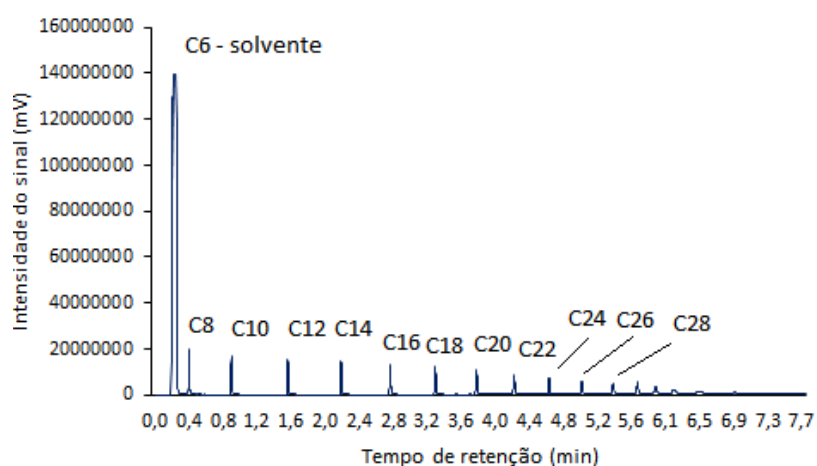
Verificou-se que as linearidades para todos os hidrocarbonetos são satisfatórias no intervalo de concentração estudado, uma vez que se adequam ao critério mínimo aceitável para o coeficiente de correlação ( $r \geq 0,99$ ) (INMETRO, 2018; ANVISA, 2017). A garantia dessa ampla faixa de linearidade é importante para a etapa de elaboração dos modelos de PLS, pois os componentes do óleo diesel se encontram em variadas concentrações.



### 5.2.2 Seletividade

Segundo a Agência Nacional de Vigilância Sanitária – ANVISA, em sua Resolução nº 166 de 24 de julho de 2017 sobre validação de métodos analíticos, a seletividade do método analítico deve ser demonstrada por meio da sua capacidade de identificar ou quantificar o analito de interesse, inequivocamente, na presença de componentes que podem estar presentes na amostra, como impurezas, diluentes e componentes da matriz. Caso a seletividade do método não seja assegurada, a linearidade e a precisão estarão seriamente comprometidas (INMETRO, 2018). A seletividade pode ser obtida de várias maneiras, uma delas é comparando a matriz isenta da substância de interesse e a matriz adicionada com esta substância (padrão), sendo que, nesse caso, nenhum interferente deve eluir no tempo de retenção da substância de interesse, que deve estar bem separada dos demais compostos presentes na amostra (RIBANI, 2004).

A Figura 59 mostra o cromatograma de uma amostra com a menor concentração da curva analítica ( $52,62 \mu\text{g}\cdot\text{g}^{-1}$ ) dos hidrocarbonetos estudados no solvente hexano), onde é visualmente perceptível que não ocorre interferência do solvente.



**Figura 59** - Cromatograma referente à amostra de hidrocarbonetos na concentração de  $52,62 \mu\text{g}\cdot\text{g}^{-1}$  e hexano como solvente, sendo possível visualizar que o hexano não interfere na seletividade dos hidrocarbonetos estudados.

Outra maneira de averiguar a seletividade é comparar amostras contendo apenas o hexano (branco) com amostras de hidrocarbonetos nele dissolvida. Assim, foram comparadas amostras nas concentrações de  $102,35 \mu\text{g}\cdot\text{g}^{-1}$ ;  $301,61 \mu\text{g}\cdot\text{g}^{-1}$  e

494,91  $\mu\text{g.g}^{-1}$  (chamadas concentrações baixa, média e alta), injetadas em triplicatas. A maior concentração da curva analítica (749,51  $\mu\text{g.g}^{-1}$ ) não foi utilizada como a alta neste caso por não ser solubilizada em hexano para a realização dos ensaios.

No tratamento estatístico da seletividade pode ser usado tanto o teste F (*Snedecor*) que avalia a homogeneidade da variância quanto o teste t (*Student*) que faz comparação entre as médias. Para o teste F, deve-se comparar o valor F calculado com o tabelado. Quando menor o valor calculado, a matriz não exerce efeito significativo. O mesmo ocorre para o teste t, se o teste t calculado for menor que o t tabelado a matriz não interfere na detecção do método (EURACHEM, 2014). Aqui, foi optado por utilizar o teste F.

A comparação foi feita através da média dos valores da área total dos cromatogramas, realizadas em  $n=7$  ensaios. Os resultados estão na Tabela 26.

**Tabela 26** - Teste F para avaliação da seletividade do método.

Amostra	hexano	Concentração baixa	Concentração média	Concentração alta
Média	95435002,67	171295733	250630911	293110086
(Desvio padrão) <sup>2</sup>	4,39.10 <sup>14</sup>	1,22.10 <sup>14</sup>	4,41.10 <sup>14</sup>	1,95.10 <sup>13</sup>
F calculado		0,277774956	0,99633906	0,0444781
F tabelado (95% de confiança)		19	19	19

O valor de  $F$  calculado é obtido através da divisão entre o quadrado do desvio padrão das áreas de amostra sem hexano pelo quadrado do desvio padrão das amostras de hexano puro, para cada concentração avaliada.

Devido ao  $F$  calculado ser menor que o  $F$  tabelado, conclui-se que o hexano e possíveis interferentes presentes no solvente não prejudicam a determinação dos analitos estudados, provando a seletividade do método.

### 5.2.3 Precisão e exatidão

A precisão foi avaliada em repetibilidade ou interdia e precisão intermediária ou interdia (realizada em 5 dias diferentes) e foi obtida através da estimativa do desvio padrão relativo percentual (*relative standard deviation* – RSD%), também conhecido como coeficiente de variação percentual (CV%). A exatidão foi obtida através do erro

relativo percentual (E%). Todas as avaliações foram realizadas a partir das áreas dos analitos após análises em triplicata de três diferentes concentrações (alta, média e baixa). Os cálculos para avaliar os ensaios interdia foram baseados em curvas analíticas processadas diariamente. Os resultados se encontram na Tabela 27.

**Tabela 27** - Resultados dos ensaios de repetibilidade, precisão intermediária e exatidão intradia e interdia.

Hidrocarbonetos		Repetibilidade			Intermediária		
		Concentração (µg/g)	CV %	E %	Concentração (µg/g)	CV %	E %
C <sub>8</sub> H <sub>18</sub>	Alta	749,51	0,4	-0,4	749,51	0,7	-1,0
	Média	301,61	3,5	0,4	301,61	4,3	0,8
	Baixa	102,35	1,8	1,9	102,35	2,1	2,3
C <sub>10</sub> H <sub>22</sub>	Alta	749,51	0,5	0,3	749,51	0,7	-0,2
	Média	301,61	3,5	2,1	301,61	3,4	1,9
	Baixa	102,35	1,8	-1,6	102,35	2,3	-2,0
C <sub>12</sub> H <sub>26</sub>	Alta	749,51	1,8	-0,1	749,51	0,8	0,4
	Média	301,61	3,4	2,2	301,61	3,7	-2,4
	Baixa	102,35	2,2	1,5	102,35	3,5	-3,1
C <sub>14</sub> H <sub>30</sub>	Alta	749,51	2,3	0,6	749,51	1,0	1,2
	Média	301,61	4,1	3,3	301,61	4,2	3,9
	Baixa	102,35	2,5	4,8	102,35	3,7	4,9
C <sub>16</sub> H <sub>34</sub>	Alta	749,51	2,5	-1,6	749,51	1,2	1,3
	Média	301,61	4,4	-3,5	301,61	4,6	2,8
	Baixa	102,35	4,8	-1,3	102,35	3,1	-2,3
C <sub>18</sub> H <sub>38</sub>	Alta	749,51	2,3	1,0	749,51	1,1	-0,7
	Média	301,61	4,8	1,4	301,61	4,9	-0,9
	Baixa	102,35	4,2	3,0	102,35	3,8	2,7
C <sub>20</sub> H <sub>42</sub>	Alta	749,51	2,4	-0,4	749,51	1,5	0,6
	Média	301,61	2,7	1,1	301,61	4,8	2,2
	Baixa	102,35	4,0	-4,6	102,35	2,1	-4,9
C <sub>22</sub> H <sub>46</sub>	Alta	749,51	3,6	-0,6	749,51	4,1	1,0
	Média	301,61	4,6	0,8	301,61	4,4	1,1
	Baixa	102,35	3,2	-3,9	102,35	5,0	-4,3
C <sub>24</sub> H <sub>50</sub>	Alta	749,51	4,4	-1,7	749,51	3,4	-3,0
	Média	301,61	2,8	-3,1	301,61	4,9	-4,8
	Baixa	102,35	2,5	4,4	102,35	4,2	4,8
C <sub>26</sub> H <sub>54</sub>	Alta	749,51	4,7	-1,0	749,51	3,6	-3,9
	Média	301,61	2,5	-4,4	301,61	3,2	-3,7
	Baixa	102,35	3,4	4,1	102,35	4,8	-2,4
C <sub>28</sub> H <sub>58</sub>	Alta	749,51	4,7	0,2	749,51	3,9	2,1
	Média	301,61	2,9	0,5	301,61	3,6	-0,8
	Baixa	102,35	1,9	-2,6	102,35	3,6	4,6

De acordo com o guia de validação e controle de qualidade analítica do Ministério da Agricultura, Pecuária e Abastecimento (MAPA), o critério de aceitação de CV% para repetibilidade é de 13,3% para amostras com concentração de 10 a 100 mg/kg e 10% para amostras com concentração de 100 a 1000 mg/kg. Então, os resultados obtidos estão dentro do limite aceito. (BRASIL, 2011). Não há nesse guia de validação utilizado critérios para aceitação de precisão intermediária, apenas para repetibilidade e reprodutibilidade. Entretanto, os resultados obtidos de precisão intermediária também estão abaixo dos limites permitidos para repetibilidade nas concentrações que foram avaliadas.

#### 5.2.4 Limite de quantificação

O limite de quantificação (LQ) é dado pela menor concentração de hidrocarbonetos que pode ser determinada com exatidão, representada pelo erro relativo percentual, e precisão, representada pelo desvio padrão relativo. Assim, o LQ foi considerado como a menor concentração da curva analítica (INMETRO,2018), ou seja, 52,62  $\mu\text{g}\cdot\text{g}^{-1}$ . As análises para o LQ foram feitas em sextuplicata e os resultados se encontram na Tabela 28.

Os valores de CV% obtidos no ensaio de limite de quantificação estão em conformidade com o indicado pelo guia de validação e controle de qualidade analítica do MAPA, sendo o critério de aceitação de CV% para repetibilidade de 13,3% para amostras com concentração de 10 a 100 mg/kg.

**Tabela 28** - Resultados do ensaio para verificar LQ.

Hidrocarboneto	LQ	
	CV %	E %
C <sub>8</sub> H <sub>18</sub>	3,9	1,2
C <sub>10</sub> H <sub>22</sub>	4,3	-1,2
C <sub>12</sub> H <sub>26</sub>	5,2	1,2
C <sub>14</sub> H <sub>30</sub>	6,1	-4,8
C <sub>16</sub> H <sub>34</sub>	4,9	-5,0
C <sub>18</sub> H <sub>38</sub>	2,4	0,3

C <sub>20</sub> H <sub>42</sub>	5,0	-4,1
C <sub>22</sub> H <sub>46</sub>	4,5	-2,1
C <sub>24</sub> H <sub>50</sub>	8,0	4,8
C <sub>26</sub> H <sub>54</sub>	1,4	-2,5
C <sub>28</sub> H <sub>58</sub>	4,4	-3,9

### 5.2.5 Limite de detecção

O limite de detecção (LD) representa a menor concentração que pode ser detectada, mas não necessariamente quantificada, utilizando um determinado procedimento experimental. Segundo orientações do INMETRO (2018) sobre validações de métodos, ele deve ser calculado através do valor do LQ dividido por 3,3.

Como o valor de LQ para todos os hidrocarbonetos foi igual a concentração da menor amostra da curva analítica (52,62 µg.g<sup>-1</sup>), então, o valor de LD é igual a 15,95 µg.g<sup>-1</sup> para todos os analitos estudados.

### 5.2.6 Estabilidade no autoamostrador

Foram realizados experimentos para avaliar a estabilidade das amostras armazenadas no autoamostrador durante 6h e durante 24h, avaliando a precisão, pelo desvio padrão relativo percentual e a exatidão, pelo erro relativo percentual. Estes experimentos foram feitos com o intuito de verificar se o tempo que as amostras podem ficar no autoamostrador, numa sala com temperatura média de 22 °C durante uma rotina de análise, influenciam nos resultados. A Tabela 29 contém os resultados verificados em concentração alta e baixa para cada analito, realizadas em triplicata.

**Tabela 29** - Resultados dos experimentos de estabilidade em autoamostrador por 6h e por 24 h para os hidrocarbonetos avaliados.

Hidrocarbonetos		Estabilidade 6h			Estabilidade 24h		
		Concentração (µg/g)	CV %	E %	Concentração (µg/g)	CV %	E %
C <sub>8</sub> H <sub>18</sub>	Alta	749,51	2,2	-3,1	749,51	1,4	-2,0
	Baixa	102,35	3,4	4,9	102,35	2,1	0,6
C <sub>10</sub> H <sub>22</sub>	Alta	749,51	3,1	4,2	749,51	3,6	-5,2
	Baixa	102,35	4,6	6,7	102,35	7,3	0,7
C <sub>12</sub> H <sub>26</sub>	Alta	749,51	1,8	2,6	749,51	4,4	5,0
	Baixa	102,35	0,5	0,7	102,35	2,5	2,2

<b>C<sub>14</sub>H<sub>30</sub></b>	<b>Alta</b>	749,51	0,3	-0,4	749,51	1,9	-2,7
	<b>Baixa</b>	102,35	3,8	-5,3	102,35	2,9	4,5
<b>C<sub>16</sub>H<sub>34</sub></b>	<b>Alta</b>	749,51	0,8	1,1	749,51	2,6	-3,7
	<b>Baixa</b>	102,35	3,9	5,7	102,35	2,8	4,8
<b>C<sub>18</sub>H<sub>38</sub></b>	<b>Alta</b>	749,51	3,2	-4,6	749,51	3,4	-4,9
	<b>Baixa</b>	102,35	3,6	5,3	102,35	4,0	1,5
<b>C<sub>20</sub>H<sub>42</sub></b>	<b>Alta</b>	749,51	1,6	-2,2	749,51	3,2	-4,6
	<b>Baixa</b>	102,35	3,3	-1,0	102,35	4,2	2,9
<b>C<sub>22</sub>H<sub>46</sub></b>	<b>Alta</b>	749,51	0,1	-0,2	749,51	4,0	-5,9
	<b>Baixa</b>	102,35	0,8	4,8	102,35	1,8	4,3
<b>C<sub>24</sub>H<sub>50</sub></b>	<b>Alta</b>	749,51	1,5	-2,2	749,51	3,1	-4,4
	<b>Baixa</b>	102,35	3,6	-1,1	102,35	8,2	4,3
<b>C<sub>26</sub>H<sub>54</sub></b>	<b>Alta</b>	749,51	0,4	0,6	749,51	2,9	-4,2
	<b>Baixa</b>	102,35	2,4	5,2	102,35	4,4	4,2
<b>C<sub>28</sub>H<sub>58</sub></b>	<b>Alta</b>	749,51	1,7	-2,4	749,51	1,1	-1,5
	<b>Baixa</b>	102,35	1,0	1,5	102,35	2,6	2,9

Pode-se concluir, pelos valores de CV% abaixo do limite aceito para a precisão, que as amostras não sofrem alterações significativas nas condições avaliadas.

### 5.2.7 Robustez

O ensaio de robustez foi verificado através de um planejamento fatorial completo 2<sup>4</sup>, com 16 experimentos sem réplicas, avaliando pequenas modificações do método cromatográfico nos fatores temperatura inicial, taxa de aquecimento 1, taxa de aquecimento 2 e vazão. A descrição destes experimentos está presente na Tabela 30.

**Tabela 30** - Descrição dos experimentos do planejamento fatorial completo 2<sup>4</sup> para o ensaio de robustez.

<b>Variável</b>	<b>Nível inferior (-)</b>	<b>Nível superior (+)</b>
Temperatura Inicial (°C)	33	37
Taxa de Aquecimento1 (°C/min)	35	55
Taxa de Aquecimento 2 (°C/min)	790	810
Vazão (mL/min)	0,4	0,6

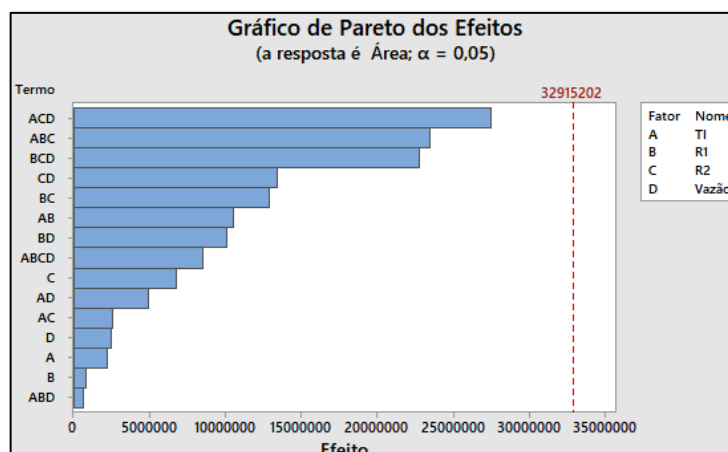
A variável analisada como resposta foi a área total, ou seja, a soma das áreas de todos os picos correspondentes aos hidrocarbonetos estudados na etapa de

validação. Estes estudos foram realizados utilizando a amostra de concentração igual a  $301,61 \mu\text{g}\cdot\text{g}^{-1}$ , que também foi utilizada como concentração média nos ensaios de precisão e exatidão. Os resultados obtidos neste planejamento se encontram na Tabela 31.

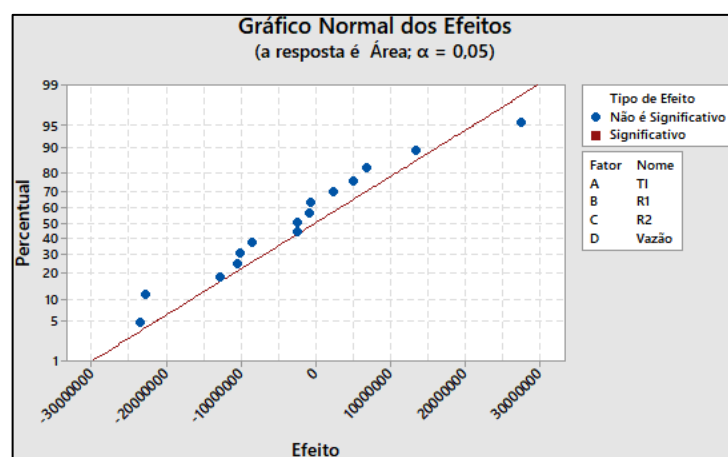
**Tabela 31** - Resultados do planejamento fatorial completo  $2^4$  para o ensaio de robustez do método desenvolvido.

Experimento	TI (°C)	Taxa de Aquecimento 1 (°C/min)	Taxa de Aquecimento 2 (°C/min)	Vazão (mL/min)	Área total
1	-	-	-	-	152903916
2	-	-	-	+	154708608
3	-	-	+	-	151481482
4	-	+	-	-	147209443
5	+	-	-	-	175156822
6	-	-	+	+	153693774
7	-	+	+	-	195428120
8	-	+	-	+	158509117
9	+	+	-	-	179636640
10	+	-	-	+	116122250
11	+	-	+	-	143527942
12	+	+	+	-	137822053
13	-	+	+	+	150234872
14	+	-	+	+	228945871
15	+	+	-	+	161691214
16	+	+	+	+	139137552

Observando os gráficos de Pareto (Figura 60) e de distribuição normal (Figura 61), observa-se que nenhuma variável nos níveis estudados é significativa para contribuir com a resposta (área), ou seja, as pequenas modificações feitas não interferem no método desenvolvido, provando que ele pode ser considerado robusto nestas condições.



**Figura 60** - Gráfico de Pareto com os fatores estudados, onde TI é temperatura inicial, R1 é taxa de aquecimento 1 e R2 é taxa de aquecimento 2. Nota-se que nenhum efeito ou combinação entre efeitos mostrou ter influência sobre a resposta estudada, ou seja, a área total dos hidrocarbonetos no nível de significância ( $\alpha=0,05$ ).



**Figura 61** - Gráfico normal dos efeitos para os fatores estudados, onde TI é temperatura inicial, R1 é taxa de aquecimento 1 e R2 é taxa de aquecimento 2. Nota-se que nenhum efeito ou combinação entre efeitos mostrou ter influência sobre a resposta estudada, ou seja, a área total dos hidrocarbonetos no nível de significância ( $\alpha=0,05$ ).

### 5.3 Seleção de amostras de óleo diesel por HCA

Com o método de separação cromatográfica desenvolvido e validado, a próxima etapa consistiu em selecionar quais amostras de óleo diesel seriam analisadas para que a representatividade de todo o conjunto fosse garantida. Aproximadamente metade das amostras totais foram mantidas. A Tabela 32



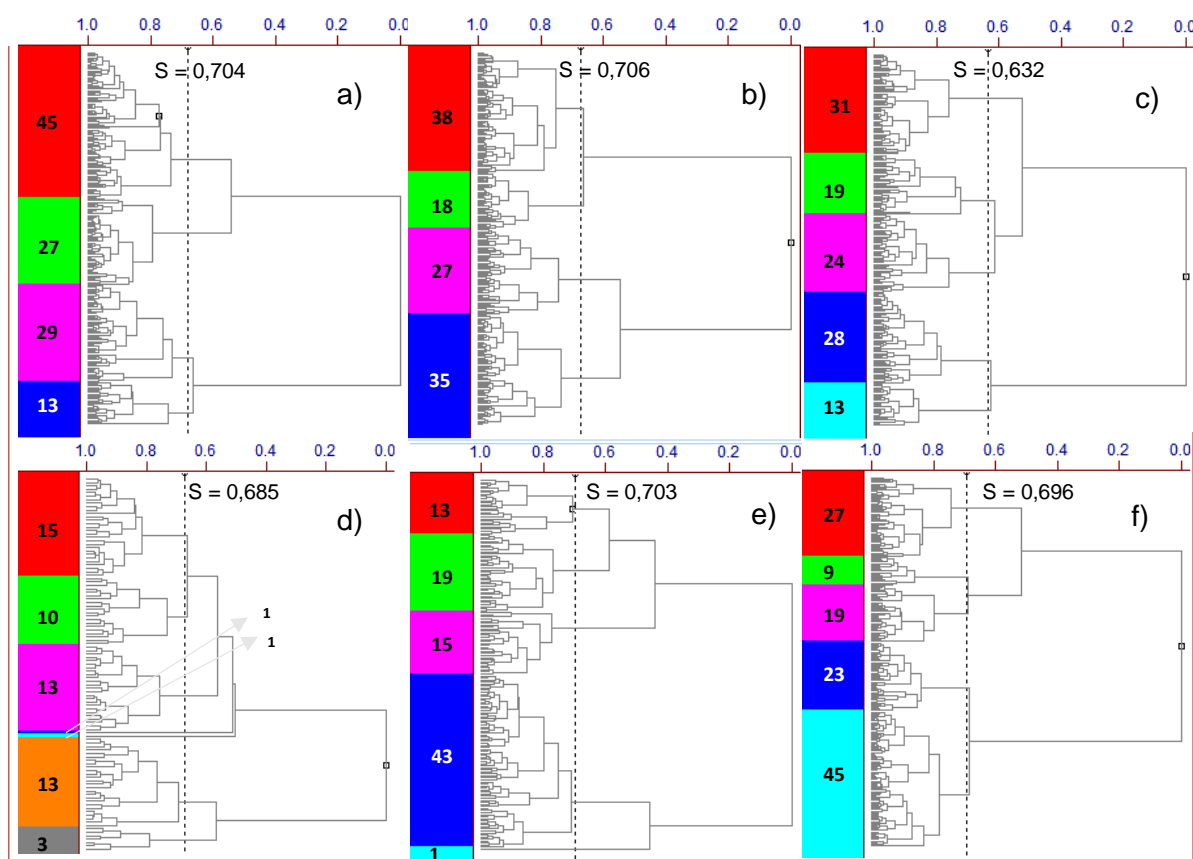
apresenta o total de amostras disponíveis em cada mês de coleta e a quantidade entre elas que foi selecionada.

**Tabela 32** - Quantidade de amostras de óleo diesel selecionadas a cada mês.

<b>Mês</b>	<b>Total de amostras</b>	<b>Amostras selecionadas para cromatografia</b>
Setembro/2017	227	114
Outubro/2017	236	118
Novembro/2017	231	115
Dezembro/2017	109	56
Janeiro/2018	182	91
Fevereiro/2018	246	123
<b>TOTAL</b>	<b>1231</b>	<b>617</b>

A Figura 62 consiste nos dendogramas obtidos por HCA na verificação de cada mês, mostrando os *clusters* formados e as similaridades entre eles. As amostras de óleo diesel foram selecionadas através dos dendogramas obtidos a cada mês em número proporcional ao tamanho dos *clusters*. Em setembro de 2017, foram selecionadas 45 amostras do vermelho; 27 do verde; 29 do magenta e 13 do azul, totalizando as 114 amostras. A similaridade escolhida para discriminação dos clusters foi 0,704 (Figura 62a). Em outubro de 2017, foram selecionadas ao todo 118 amostras, sendo estas, 38 do vermelho; 18 do verde; 27 do magenta; 35 do azul e a similaridade escolhida foi 0,706 (Figura 62b). No mês de novembro de 2017, foram selecionadas 31 amostras do vermelho; 19 do verde; 24 do magenta, 28 do azul, e 13 do azul claro, totalizando as 115 amostras e a similaridade escolhida foi 0,632 (Figura 62c). Em dezembro de 2017, foram selecionadas 15 amostras do vermelho; 10 do verde; 13 do magenta, 1 do azul, 1 do azul claro, 13 do laranja e 3 do cinza, totalizando as 56 amostras e a similaridade escolhida para discriminação dos clusters foi 0,685 (Figura 62d). As amostras selecionadas em janeiro de 2018 foram 13 do vermelho; 19 do verde; 15 do magenta; 43 do azul e 1 do azul claro, somando as 91 e a similaridade escolhida foi 0,703 (Figura 62e). Por fim, no mês de fevereiro de 2018 foram selecionadas 123 amostras, sendo elas obtidas: 27 do vermelho; 9 do verde; 19 do magenta; 23 do azul e 45 do azul claro, com similaridade escolhida de 0,696 (Figura

62f). Dentro dos *clusters* foram escolhidas as amostras pertencentes aos subgrupos mais distantes entre si (com menores valores de similaridade) a fim de manter a amostragem homogênea e fidedigna em relação ao conjunto original. O conjunto selecionado possui 18 amostras não conformes (2,9% do total), enquanto o original possui 38 amostras não conformes (3,1% do total).



**Figura 62** - Dendrogramas com os resultados de HCA para seleção das amostras de diesel a cada mês: a) setembro/2017; b) outubro/2017; c) novembro/2017; d) dezembro/2017; e) janeiro/2018; f) fevereiro/2018.

## 5.4 Desenvolvimento de modelos de calibração multivariada

### 5.4.1 Ajustes nos picos cromatográficos

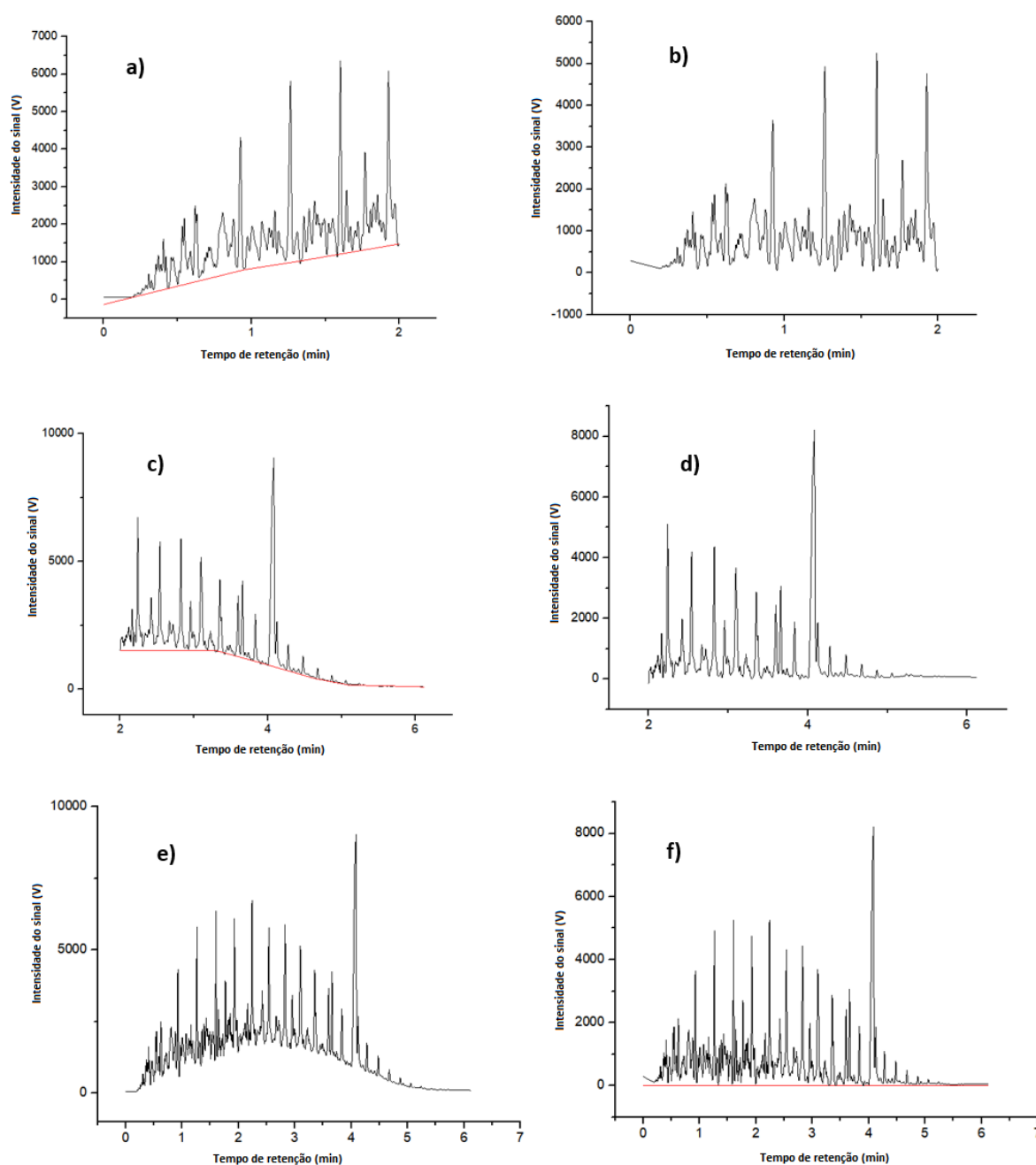
As amostras selecionadas por HCA foram analisadas por CG-UR pelo método desenvolvido e validado e, então, os dados referentes às áreas de cada pico cromatográfico presente em cada amostra foram exportados ao *Excel* e salvos de arquivos *.ASC* para *.xls*, a fim de poder criar um banco de dados juntamente com os

dados dos parâmetros físico-químicos, onde os cromatogramas formaram a matriz X e os parâmetros a matriz Y.

A princípio, foi necessário corrigir o deslocamento existente entre os picos e a linha de base nos cromatogramas. Para isso, o cromatograma da amostra mais complexa, isto é, com maior quantidade de picos foi utilizado como referência (194 picos) e um tratamento dessa linha de base foi realizado dividindo o perfil em duas partes (até 2,0 min, correspondendo à parte ascendente da linha de base e de 2,0 a 6,12 min, correspondendo à parte descendente) (Figura 63).

Foi utilizado uma função de correção de linha de base do programa *Origin*, onde a seleção de pontos é feita manualmente. Então, realizou uma subtração entre o cromatograma original e a linha de base, resultando em um novo cromatograma, para então ser reinserido ao *software ChromQuest* e refazer a integração dos picos, sem a interferência da lacuna/*drift* existente sob a linha de base. Para isso, foi necessário converter novamente o arquivo .xls em .ASC e, a única forma possível de realizar essa conversão é exportar novamente os cromatogramas originais do *ChromQuest* para o *Excel* ainda em formato .ASC e, então, substituir os dados originais pelos modificados com a correção da linha de base na planilha. Em seguida, o arquivo foi salvo na pasta do programa, substituindo o original, para então abri-lo e visualizar o cromatograma atualizado e poder integrar os picos e, novamente, exportar os dados das áreas em planilha do *Excel* salvas como .xls para posterior utilização.

Por fim, a área total do cromatograma sem o *drift* passou a ser de 1947469012, correspondente a 53,4% da área do cromatograma com o *drift*.



**Figura 63** - Tratamento da linha de base de amostra de diesel a) primeira parte do cromatograma (até 2,0 min) antes da correção; b) primeira parte após correção; c) segunda parte do cromatograma (2,0 a 6,12 min) antes da correção; d) segunda parte após correção; e) cromatograma inteiro antes da correção; f) cromatograma inteiro após correção.

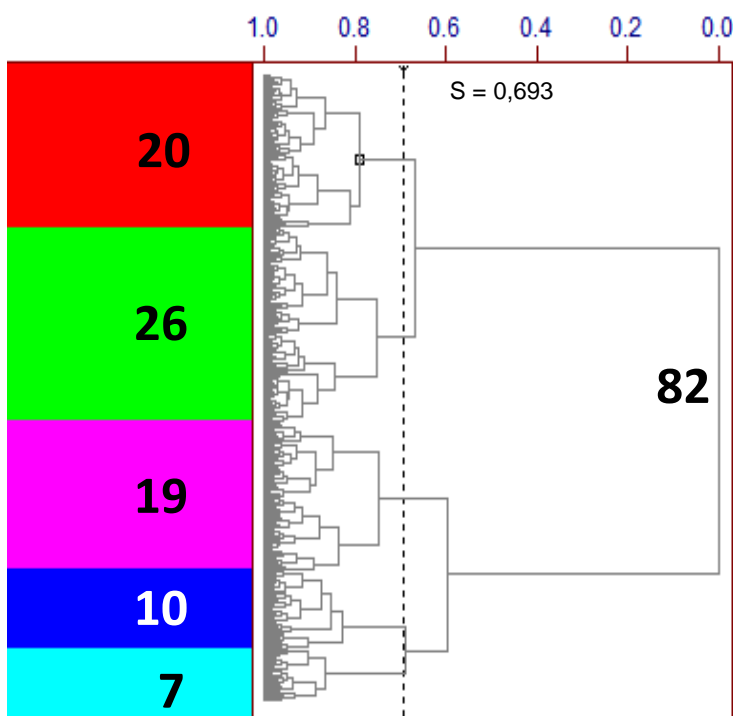
#### 5.4.1.1 Nova seleção de amostras de óleo diesel por HCA

Devido à morosidade desse procedimento, optou-se por reduzir o número do conjunto amostral de 617 para 100 amostras. Para isto, foi realizado um novo procedimento de HCA (Figura 64) com o conjunto formado pelas 599 amostras

conformes já analisadas por CG-UR anteriormente, no intuito de selecionar 82. As 18 amostras não conformes presentes no conjunto original foram mantidas a fim de aumentar o percentual de não conformidade do conjunto (18%) e resultar em uma faixa de valores mais amplas aos modelos.

Os critérios utilizados nesta HCA foram os mesmos da anterior (parâmetros físico-químicos como variáveis, autoescalamento dos dados das variáveis, distância euclidiana e grupamento incremental).

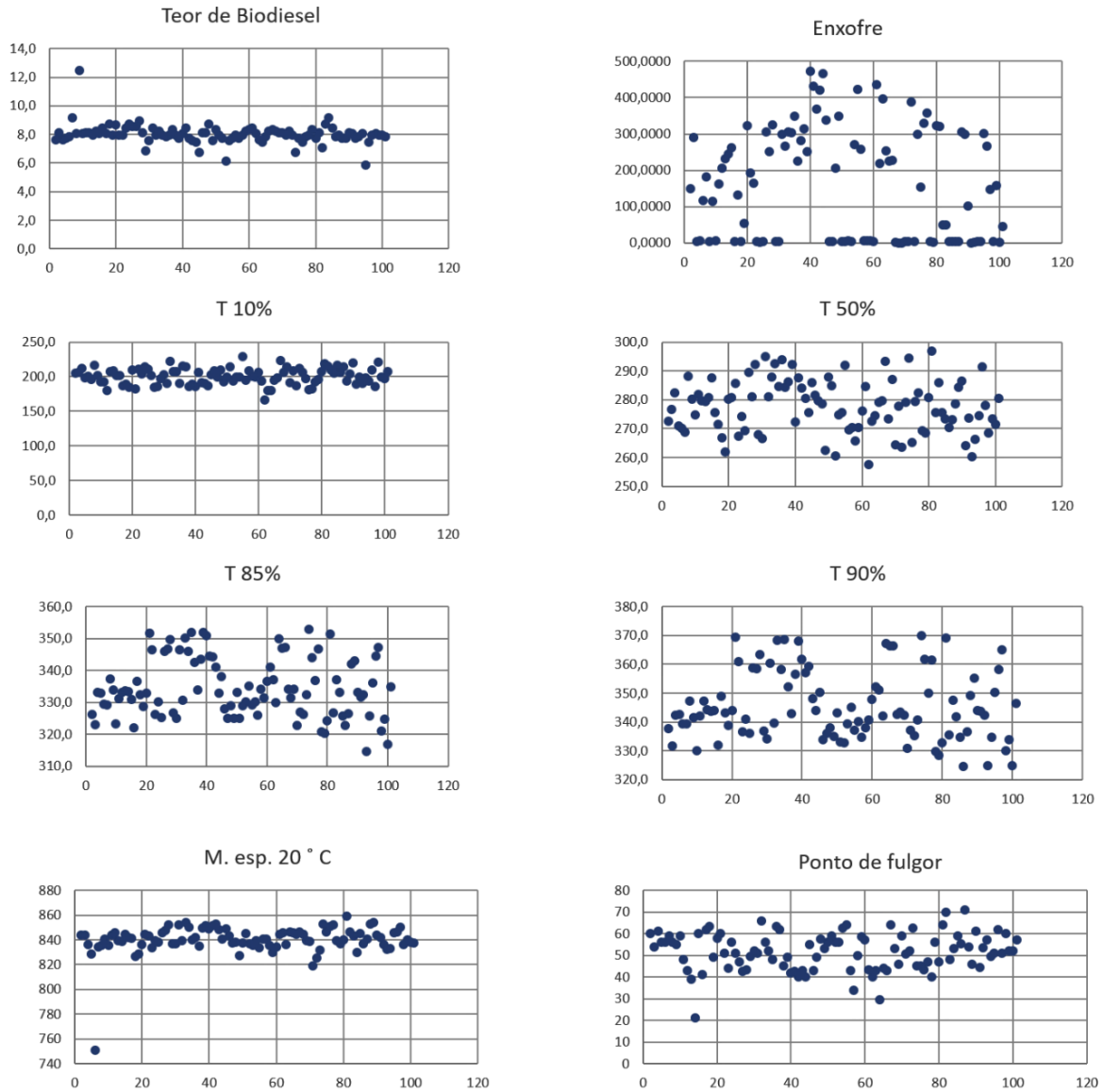
A partir do dendograma apresentado na Figura 64, foram selecionadas as 82 amostras com menor similaridade entre si, extraídas dos grupos mais com maior distância possível. A similaridade escolhida para discriminação dos *clusters* foi 0,693, o que indica que amostras presentes em um mesmo agrupamento devem possuir no mínimo 69,3% de similaridade. Este valor foi escolhido para que um número razoável de *clusters* fosse formado, facilitando a escolha. O número de amostras selecionadas foi proporcional ao tamanho de cada *cluster*, sendo 20 amostras obtidas do vermelho; 26 obtidas do verde; 19 do magenta; 10 do azul e 7 do azul claro.



**Figura 64** - Dendograma resultante de HCA para seleção de novo conjunto representativo de amostras de óleo diesel.

A Figura 65 mostra um gráfico de valores vs quantidade de amostras, mostrando a faixa de variação de cada parâmetro físico-químico, também presente na

Tabela 33. Das 18 amostras não conformes do conjunto, 6 apresentaram valores de biodiesel abaixo do especificado; 8 apresentaram valores de biodiesel acima do especificado; 3 apresentaram ponto de fulgor abaixo do especificado e 1 apresentou massa específica abaixo do valor especificado (Tabela 34).



**Figura 65** - Variação dos resultados de todos os parâmetros físico-químicos das amostras selecionadas.

**Tabela 33** - Intervalo de valores de cada parâmetro físico-químico presente nas amostras do conjunto representativo de óleo diesel.

Parâmetro físico-químico	Unidade	Conjunto representativo de óleo diesel				
		Mínimo	Máximo	Varição	Média	Não conformidades
Massa específica	kg/m <sup>3</sup>	751,0	859,1	108,1	840,3	1
Temperatura de destilação						
10% Evaporada, máx	°C	167,0	229,4	62,4	200,2	0
50% Evaporada, máx	°C	260,4	296,8	36,4	277,5	0
85% Evaporada, máx	°C	314,7	353,0	38,3	334,3	0
90% Evaporada, máx	°C	324,7	369,2	44,5	345,4	0
Teor de biodiesel	% vol.	5,9	12,5	6,6	8,1	14
Teor de enxofre		1,2	475,0	463,0	165,8	0
S10	mg/kg	1,2	8,9	7,7	6,2	0
S500	mg/kg	48,0	475,0	427,0	263,7	0
Ponto de fulgor	°C	29,5	71,0	41,5	51,6	3

Todos os parâmetros físico-químicos foram obtidos através de normas oficiais estabelecidas pela resolução ANP n° 50, de 2013.

**Tabela 34** - Valores dos parâmetros físico-químicos fora da especificação nas amostras selecionadas de óleo diesel.

Amostra	Não conformidade	Valor especificado *	Valor determinado
1	Biodiesel	8,0 ± 0,5 (%volume)	7,1
2	Biodiesel	8,0 ± 0,5 (%volume)	5,9
3	Biodiesel	8,0 ± 0,5 (%volume)	6,8
4	Biodiesel	8,0 ± 0,5 (%volume)	6,8
5	Biodiesel	8,0 ± 0,5 (%volume)	9,2
6	Biodiesel	8,0 ± 0,5 (%volume)	6,2
7	Biodiesel	8,0 ± 0,5 (%volume)	12,5
8	Biodiesel	8,0 ± 0,5 (%volume)	8,7
9	Biodiesel	8,0 ± 0,5 (%volume)	8,8
10	Biodiesel	8,0 ± 0,5 (%volume)	8,8
11	Biodiesel	8,0 ± 0,5 (%volume)	8,6
12	Biodiesel	8,0 ± 0,5 (%volume)	9,0
13	Biodiesel	8,0 ± 0,5 (%volume)	6,9
14	Biodiesel	8,0 ± 0,5 (%volume)	8,6
15	Ponto de fulgor	38 °C (valor mínimo)	29,5
16	Ponto de fulgor	38 °C (valor mínimo)	34,0
17	Ponto de fulgor	38 °C (valor mínimo)	21,0
18	Massa específica	815 – 850 (kg/m <sup>3</sup> ) – S10 815 – 865 (kg/m <sup>3</sup> ) – S500	751 – S500

\* Valores especificados de acordo com a Resolução n° 50 da ANP. O valor do biodiesel está de acordo com o estabelecido na Lei 13.263/2016, vigente quando as amostras foram analisadas.

#### 5.4.2. Desenvolvimento dos modelos PLS

À priori, todas as regiões dos cromatogramas foram utilizadas para a construção dos modelos, entretanto, devido à presença de muitos picos pouco resolvidos, o uso das áreas e tempos de retenção de todos eles para a construção da matriz X prejudicou o alinhamento dos cromatogramas, dificultando a correlação com os parâmetros analisados. Então, apesar de todo o desenvolvimento do método cromatográfico objetivar a obtenção do maior número possível de picos cromatográficos, optou-se nesta fase por escolher apenas os principais, com maior área e intensidade e com boa separação em todas as amostras, incluindo os 11 picos referentes aos hidrocarbonetos parafínicos identificados na série homóloga e os 2 picos identificados como FAMES do biodiesel. Assim, foram construídos novos modelos com os conjuntos de calibração (matriz 70 x 30 x 8 – amostras vs variáveis independentes vs variáveis dependentes) e o de previsão (matriz 30 x 30 x 8 – amostras vs variáveis independentes vs variáveis dependentes).

A capacidade preditiva do modelo é estabelecida nos conjuntos de calibração e de previsão através de seus respectivos erros padrão (RMSEC e RMSEV) e também pelos seus coeficientes de correlação ( $r_{cal}$  e  $r_{val}$ ). Tanto o RMSEV quanto o  $r_{val}$  são



valores obtidos através da validação cruzada *leave-one-out*, onde uma amostra do conjunto de calibração é excluída e o modelo construído é utilizado para prever suas propriedades. Por fim, o conjunto de previsão foi utilizado para avaliar o modelo desenvolvido. Os parâmetros RMSEP (erro padrão de previsão) e erro relativo médio são utilizados para medir a capacidade de previsão para o modelo criado.

Para a construção dos modelos de previsão por PLS foram avaliados os pré-processamentos: nenhum, autoescalado, centrado na média, escalado pela faixa e pela variância e as transformações matemáticas: nenhuma, primeira e segunda derivada, logarítmica e normalizada. Foi realizada a validação cruzada após a definição dos melhores critérios para prever cada parâmetro físico-químico. Os pré-processamentos e transformações matemáticas que resultaram nos melhores modelos de regressão multivariada (isto é, modelos com menores erros de calibração, validação e previsão e maiores coeficientes de correlação para a calibração e a validação) para os parâmetros físico-químicos constam na Tabela 35.

**Tabela 35** - Pré-processamento e transformações dos dados para os modelos de calibração.

Parâmetros físico-químicos	Pré-processamento	Transformações
Massa específica	Centrado na média	Normalizada
Ponto de fulgor	Centrado na média	Logarítmica
Teor de enxofre	Autoescalado	Normalizada
Teor de biodiesel		
Temperaturas de destilação		
T 10%		
T 50%		
T 85%		
T 90%		

Diferentes pré-processamentos foram utilizados para as propriedades em questão porque os algoritmos multivariados criados em cada caso fornecem resultados ótimos de acordo com os padrões de variância nas variáveis independentes (áreas dos picos). Escalas e/ou magnitudes de medidas arbitrárias podem produzir resultados errôneos se pré-processamentos inapropriados forem efetuados (DOS SANTOS, 2009; FERREIRA, 2015).

A normalização é usada principalmente para remover variações sistemática, em geral associadas com o tamanho da amostra. É bastante utilizada com a finalidade de corrigir o efeito da variação no volume de injeção em cromatografia, normalizando cada cromatograma para área unitária. (FERREIRA, 2015). Com exceção do

parâmetro ponto de fulgor, este tipo de transformação foi o que produziu melhores resultados nos modelos. A transformação utilizada para o ponto de fulgor foi a logarítmica ( $\log 10$ ), aplicada com o objetivo de linearizar os dados, não afetando a interpretação dos resultados. Esta transformação também pode ser usada para enfatizar intensidades baixas (FERREIRA, 2015), o que está de acordo com a baixa intensidade dos primeiros picos cromatográficos, que são correspondentes às substâncias mais voláteis do diesel que caracterizam o ponto de fulgor.

Quando a faixa de variação dos dados é grande, recomenda-se o autoescalamento de cada variável por um valor. Assim, minimizamos o efeito (influência) de uma variável dominante em cálculos posteriores. A maioria dos parâmetros estudados foi autoescalada, apenas em ponto de fulgor e massa específica os dados foram centrados na média, que é apenas fazer uma translação de eixos para o valor médio de cada um deles, onde a estrutura dos dados é totalmente preservada (FERREIRA, 2015). Este é o pré-processamento mais utilizado quando se trabalha com espectros, mas também é aplicado a cromatogramas com frequência, como por exemplo, Flumignan e colaboradores (2008) utilizaram os dados centrados na média para modelo de previsão por PLS para a massa específica e composição de hidrocarbonetos da gasolina. Após definição dos melhores pré-processamentos e transformações para cada parâmetro, foram determinados os números de variáveis latentes ótimos na construção de cada modelo pela observação do valor correspondente ao menor valor de RMSEV.

Por fim, os cálculos de repetibilidade e reprodutibilidade dos guias normativos foram feitos com aplicação das equações presentes nas normas para cada parâmetro, onde foram utilizados os valores da média do conjunto amostral de óleo diesel (ANEXO A). O erro médio relativo foi calculado para todos os parâmetros pela média dos módulos das diferenças entre o valor medido e o valor previsto, em porcentagem. Através deste valor percentual, foram obtidos os erros padronizados pela multiplicação dessa porcentagem de erro pelos valores médios dos parâmetros de acordo com as normas. Assim, pode-se realizar a comparação entre os erros dos modelos com os valores de repetibilidade e reprodutibilidade definidos pelas referências normativas.

A Tabela 36 mostra o número de amostras anômalas (*outliers*) retirados do modelo, o número de variáveis latentes (VL) utilizadas, erros padrões de previsão

(RMSEP), calibração (RMSEC) e validação (RMSEV), coeficientes de correlação de calibração ( $r_{cal}$ ) e validação ( $r_{val}$ ) e os valores de erro padronizado.

A Figura 66 ilustra as regressões lineares obtidas a partir dos dados cromatográficos das amostras contidas na etapa de calibração e validação para cada parâmetro físico-químico com a aplicação dos pré-processamentos e transformações definidos.

**Tabela 36** - Características gerais dos modelos de PLS obtidos nas melhores condições de pré-processamento, transformações e seleção de variáveis para cada parâmetro físico-químico do óleo diesel e comparação dos modelos com valores obtidos através dos guias de referência.

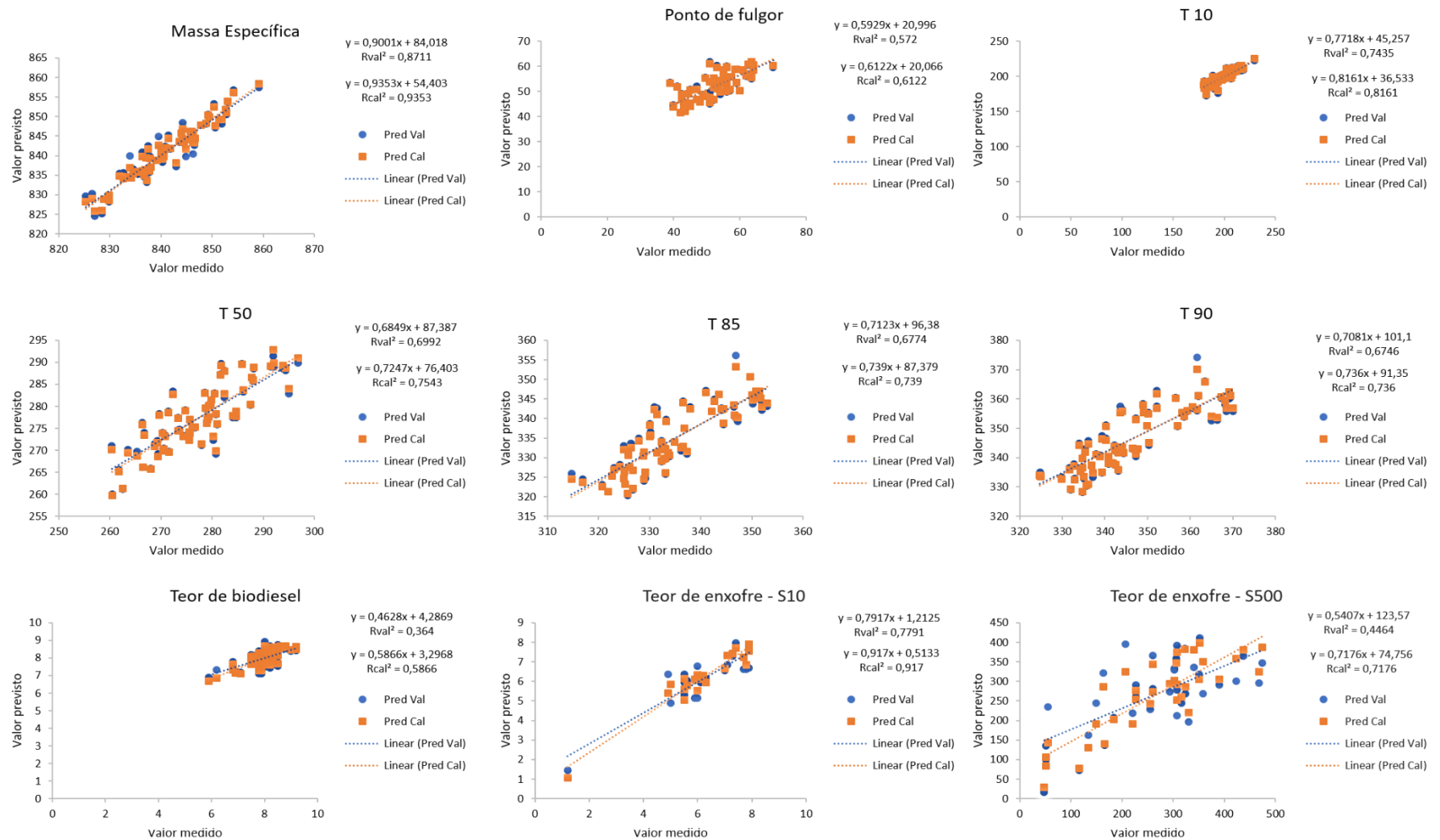
Parâmetros físico-químicos	Guias de referência*					Modelos PLS						E.P.***
	Norma	Unidade	Repetibilidade	Reprodutibilidade	Outliers**	Previsão		Calibração		Validação		
						VL	RMSEP	RMSEC	r <sub>cal</sub>	RMSEV	r <sub>val</sub>	
Massa específica	NBR 14065	kg.m <sup>-3</sup>	0,0001	0,0005	5 <sup>a</sup> -0 <sup>b</sup>	7	2,5601	2,0328	0,9671	2,6970	0,9333	2,00
Ponto de fulgor	ASTM D93	° C	1,50	3,66	6 <sup>a</sup> -2 <sup>b</sup>	2	5,4299	4,6338	0,7824	4,7570	0,7563	2,80
Temperaturas de destilação												
10% Evaporada (T10)		° C	1,88	4,40	3 <sup>a</sup> -1 <sup>b</sup>	3	6,9588	4,7842	0,9034	5,4911	0,8623	4,52
50% Evaporada (T50)	NBR 9619	° C	0,94	2,97	3 <sup>a</sup> -1 <sup>b</sup>	2	4,5237	4,4821	0,8685	4,8521	0,8362	3,16
85% Evaporada (T85)		° C	-	-	2 <sup>a</sup> -0 <sup>b</sup>	2	5,0130	4,9648	0,8597	5,4121	0,8230	3,87
90% Evaporada (T90)		° C	1,42	5,18	2 <sup>a</sup> -0 <sup>b</sup>	2	6,3238	6,4270	0,8579	6,9945	0,8214	4,68
Teor de biodiesel	EN 14078	% vol.	0,11	0,23	1 <sup>a</sup> -1 <sup>b</sup>	5	0,3712	0,3810	0,7659	0,4610	0,6033	0,27
Teor de enxofre	NBR 14533 (S500)	mg/kg	15,81	69,74	0 <sup>a</sup> -0 <sup>b</sup>	5	38,8577	60,5343	0,8471	86,2984	0,6681	28,00
	ASTM D5453 (S10)		1,0	3,3	1 <sup>a</sup> -0 <sup>b</sup>	7	0,7151	0,4210	0,9575	1,7084	0,8827	0,35

VL: número de variáveis latentes; r<sub>cal</sub> e r<sub>val</sub>: coeficientes de correlação dos conjuntos de calibração e validação, respectivamente.

\* Valores provenientes do uso das equações de cálculo de repetibilidade e reprodutibilidade das normas com a aplicação no valor médio do conjunto amostral.

\*\*Número de amostras anômalas retiradas dos conjuntos a) calibração; b) previsão.

\*\*\* Erro padronizado com base no erro médio relativo, aplicado aos valores das médias de cada parâmetro.



**Figura 66** - Regressões lineares obtidas com os dados de valores medidos vs previstos na calibração e validação cruzada nos modelos PLS de cada parâmetro.

Para o parâmetro teor de enxofre, primeiramente foi construído um modelo para todas as amostras, entretanto, o erro de previsão resultante foi bastante elevado e praticamente todas as amostras S10 foram detectadas como anômalas, então, foi melhor separá-las em dois conjuntos (38 amostras S10, onde 23 amostras foram para o conjunto de calibração e 15 foram para o conjunto de previsão; e 62 amostras S500, onde 44 amostras foram para o conjunto de calibração e 18 para o conjunto de previsão). Essa divisão auxiliou também na interpretação dos resultados, uma vez que os métodos oficiais de análise são diferentes e os valores de repetibilidade e reprodutibilidade também.

As características massa específica, teor de enxofre e temperaturas de destilação são dependentes dos componentes que estão presentes ao longo de todo o cromatograma, enquanto o biodiesel é correlacionado a dois picos cromatográficos (Figura 54) e o ponto de fulgor com os primeiros picos cromatográficos, os quais correspondem aos componentes mais voláteis da amostra. Por esta razão, a melhor condição de modelo de previsão para o ponto de fulgor foi obtida utilizando apenas os 3 primeiros picos cromatográficos, enquanto que os modelos para os demais parâmetros foram construídos utilizando todos os 30 picos cromatográficos e, mesmo o biodiesel, que não possui correlação com o cromatograma todo, apresentou melhor resultado quando foi modelado com todos os picos. O átomo de enxofre, apesar de não ser detectado pelo FID, se apresenta no óleo diesel ligado a átomos de carbono e hidrogênio, como mercaptanos, sulfetos, dissulfetos e compostos heterocíclicos (BREITKREITZ, 2003) e, por isso, foi possível obter modelos bem ajustados para o parâmetro teor de enxofre.

Pela análise dos modelos (Tabela 36 e Figura 66), observa-se que, com exceção do teor de enxofre (tanto para o conjunto S10 quanto para o conjunto S500), os valores de RMSEP de todos os parâmetros encontram-se acima do intervalo de valores entre repetibilidade e reprodutibilidade obtido pelas referências normativas. Entre estes, o parâmetro teor de biodiesel é o que mais se aproxima do valor da norma. Além do RMSEP, o erro padronizado também foi utilizado para avaliar a capacidade de previsão dos modelos de PLS, e, novamente, para o teor de enxofre está dentro da faixa de variação da norma e para o biodiesel, o valor modelado está bem próximo ao valor normativo. Os erros padronizados da temperatura da amostra 90% destilada e do ponto de fulgor estão contidos na faixa normativa também e os erros dos demais parâmetros, apesar de acima da faixa, não estão tão acima. Vale

ressaltar que o parâmetro temperatura da amostra 85% destilada não apresenta valores de repetibilidade e reprodutibilidade indicados pela norma referente às temperaturas de destilação.

Com base nestas observações, os modelos de PLS construídos neste trabalho não estão inadequados para a previsão dos 8 parâmetros físico-químicos do óleo diesel aqui estudados e poucas amostras anômalas foram retiradas em suas construções. Vale ressaltar a rapidez com que os dados são adquiridos ao acoplar a CG-UR com a PLS, comparados com o tempo total de análise pelos métodos oficiais e a vantagem de obtenção dos 8 parâmetros simultaneamente gastando apenas 0,1  $\mu\text{L}$  (volume de injeção cromatográfica) de amostra de óleo diesel, tornando estes modelos ambientalmente atrativos.

Comparando os resultados aqui obtidos com um trabalho semelhante do grupo CEMPEQC utilizando CG convencional (ZANÃO *et al.*, 2018) para os parâmetros avaliados por ambos através de metodologias oficiais, temos que (Tabela 37):

**Tabela 37** - Comparação dos modelos PLS obtidos por CG-UR e por CG.

PARÂMETROS	VL		RMSEP		RMSEC		$r_{cal}$		RMSEV		$r_{val}$		E <sub>relativo</sub> (%)	
	CG-UR	CG	CG-UR	CG	CG-UR	CG	CG-UR	CG	CG-UR	CG	CG-UR	CG	CG-UR	CG
<b>Massa específica</b>	7	4	2,5601	1,3241	2,3373	1,1906	0,9554	0,9104	2,8437	1,4462	0,9264	0,8574	0,24	0,11
<b>Ponto de fulgor</b>	2	4	5,4299	2,9285	4,6338	4,4462	0,7824	0,7399	4,7570	4,9730	0,7563	0,6376	9,39	4,79
<b>Temperaturas de destilação</b>														
<b>T10</b>	3	3	6,9588	6,6085	4,7842	4,8090	0,9034	0,7688	5,4911	5,7857	0,8623	0,6271	2,26	2,71
<b>T50</b>	2	3	4,5237	3,6503	4,4821	2,8534	0,8685	0,8050	4,8521	3,5686	0,8362	0,6601	1,14	0,94
<b>T85</b>	2	4	5,0130	3,3602	4,9648	2,9268	0,8597	0,7663	5,4121	3,8351	0,8230	0,5128	1,16	0,67

O teor de biodiesel na época vigente do trabalho por CG era facultativo de 2% e não havia metodologia oficial especificada pela ANP, enquanto que na época vigente ao trabalho por CG-UR era obrigatório de 8% e já normatizado pela EN 14078.

O teor de enxofre na época vigente do trabalho era de 500 mg/kg ou 2000 mg/kg e foi avaliado num mesmo conjunto, sendo os valores obtidos por uma só norma referenciada, a ABNT NBR 14533.

Outros parâmetros não foram avaliados em ambos os trabalhos, por isso não há comparações.



A análise da Tabela 37 mostra que os modelos de PLS construídos a partir de dados de CG-UR apresentam em geral maiores coeficientes de correlação e erros de calibração, validação cruzada, previsão e relativo médio pouco acima dos modelos criados a partir de dados de CG. Apenas o parâmetro ponto de fulgor apresenta erros de previsão e relativo médio consideravelmente maiores, mas que quando comparados com a faixa de valores da norma (Tabela 36) não estão muito acima do especificado. Isto confirma mais uma vez que a CG-UR não perdeu informações amostrais devido à sua rapidez, bem como a escolha de apenas 30 picos cromatográficos com maior área e intensidade não impossibilitou a criação de modelos de previsão. Com o apelo de apenas 6,12 minutos de análise cromatográfica, os modelos desenvolvidos neste trabalho apresentam uma alternativa atrativa para a análise dos parâmetros avaliados em rotinas laboratoriais.

## 6. CONCLUSÕES

Com a utilização de planejamento experimental e análise por superfície de respostas, foi possível aperfeiçoar um método de cromatografia gasosa ultrarrápida, no qual em apenas 6,12 minutos foi possível identificar em média 187 picos com resolução global de 1,09, onde todos os hidrocarbonetos parafínicos presentes na faixa de destilação do petróleo correspondente ao óleo diesel foram eluídos e bem resolvidos, tal qual o biodiesel, mostrando que apesar da rapidez do método, não houve perda de informações.

A validação analítica deste método apresentou seletividade, linearidade adequada ( $r > 0,99$ ) sem falta de ajuste e heterocedasticidade. Apresentou também valores, segundo o guia de validação e controle de qualidade analítica elaborado pelo MAPA, adequados de precisão intermediária e repetibilidade (menores que 5,0 e 4,8, respectivamente) e de erro avaliados intradia e interdias (menores que 4,8 e 4,9, respectivamente); limites de quantificação (52,62  $\mu\text{g/g}$  com CV% menores que 8%) e detecção (15,95  $\mu\text{g/g}$ ) condizentes com concentrações que podem ser encontradas em amostras de óleo diesel, e robustez. A análise dessas figuras de mérito confirma a adequação do método cromatográfico para o fim proposto, sugerindo que ele é apropriado para a identificação e quantificação dos componentes do óleo diesel para composição da matriz de dados e construção dos modelos de calibração multivariada.

O uso da ferramenta HCA permitiu selecionar adequadamente amostras de óleo diesel a partir dos dados dos parâmetros físico-químicos obtidos através dos métodos oficiais, reduzindo o número de amostras a serem analisadas pelo método analítico desenvolvido, otimizando assim o tempo total gasto em análises cromatográficas e calibração multivariada, porém mantendo a representatividade do conjunto total disponível.

O *drift* na linha de base é comumente encontrado em análises cromatográficas do óleo diesel devido à complexidade da amostra e saturação dos componentes na coluna, e sua correção nas linhas de base dos cromatogramas se mostrou adequada e possibilitou integração fidedigna das áreas dos picos. A seleção de 30 picos cromatográficos, entre eles os 11 correspondentes aos alcanos verificados pela série homóloga e 2 referentes aos

FAMEs do biodiesel, foi adequada e possibilitou a criação de modelos bem ajustados de previsão para os parâmetros avaliados.

A utilização de pré-processamentos e transformações resultaram em modelos com maiores coeficientes de correlação e menores erros. O autoescalamento e a normalização dos dados foram os mais utilizados entre os parâmetros. O modelo para ponto de fulgor apresentou melhor resultado quando foi utilizado apenas os 3 primeiros picos cromatográficos, confirmando a relação entre este parâmetro e os componentes mais voláteis do óleo diesel.

Em geral, os modelos de regressão construídos por PLS utilizando área dos cromatogramas e valores adquiridos através de ensaios por métodos de referência ASTM, NBR e EN dos parâmetros físico-químicos como banco de dados apresentaram resultados com boa capacidade de previsão com erros padronizados não muito acima do intervalo entre repetibilidade e reprodutibilidade obtido pelas normas oficiais dos parâmetros, com a vantagem da rapidez, possibilidade de previsão simultânea de 8 parâmetros e do pequeno volume de amostra utilizado (0,1 µL).

Em contrapartida aos não tão menores valores de erros, com o uso dos métodos oficiais NBR, ASTM e EN para monitorar estes parâmetros, há um gasto de aproximadamente 280 mL de cada amostra e tempo total de cerca de 90 minutos, onde um analista consegue analisar 8 amostras de óleo diesel por dia. Isto torna o método cromatográfico e os modelos de PLS desenvolvidos neste trabalho interessantes para uma possível aplicação em rotinas laboratoriais, uma vez que, com o método analítico desenvolvido nesse trabalho, uma pessoa consegue analisar 5 amostras por hora, em cromatografia e, no período em que um grupo de 5 amostras estiverem eluindo, é possível efetuar as correções da linha de base, calcular as áreas dos picos e inserir os dados no *software*, realizando assim a previsão simultânea dos 8 parâmetros avaliados nas 5 que eluíram anteriormente pelo método cromatográfico.

Essa facilidade possibilita que os modelos de PLS sejam utilizados em rotinas laboratoriais como etapa de *screening*, onde apenas as amostras previstas como não conformes necessitariam da etapa de análise pelos métodos normativos oficiais, gerando uma inovação em controle de qualidade de óleo diesel, o combustível mais utilizado no setor de transporte de maior importância no país, o rodoviário.

## REFERÊNCIAS

- ABADIE, E. **Apostila sobre processos de refinação**. Rio de Janeiro: Universidade Corporativa Petrobras/RH/UC/DTA, 2002. 76 p.
- ABRAHAM, M. H. 100 years of chromatography or is it 171? **Journal of Chromatography A**, v. 1061, p. 113-114, 2006.
- ALMEIDA, J. A. S. et al. Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering. **Chemometrics and Intelligent Laboratory Systems**, v. 87, n. 2, p. 208–217, 2007.
- ANTUNES, A.M.S. **A Indústria Química Orgânica: Classificações e Características**. *IN: Setores da Indústria Química Orgânica*, Antunes, A.M.S. (Org.), Rio de Janeiro: E-papers, 2007. 243 p.
- AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA - ANVISA. **Guia para validação de métodos analíticos e bioanalíticos**. Brasília: ANVISA, 2017. Disponível em: [http://portal.anvisa.gov.br/documents/10181/2721567/RDC\\_166\\_2017\\_COMP.pdf/d5fb92b3-6c6b-4130-8670-4e3263763401](http://portal.anvisa.gov.br/documents/10181/2721567/RDC_166_2017_COMP.pdf/d5fb92b3-6c6b-4130-8670-4e3263763401). Acesso em: 28 abr. 2018.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS – ANP. **Anuário estatístico brasileiro do petróleo, gás natural e biocombustíveis 2018**. Brasília: 2018a. Disponível em: <http://www.anp.gov.br/publicacoes/anuario-estatistico/anuario-estatistico-2018>. Acesso em: 9 jan. 2019.
- AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA - ANVISA. **Resolução RDC nº 166, 24 de julho de 2017**. Dispõe sobre a validação de métodos analíticos e dá outras providências. Diário Oficial da União, Brasília, DF, 25 de julho de 2017. Disponível em: [http://portal.anvisa.gov.br/documents/10181/2721567/RDC\\_166\\_2017\\_COMP.pdf/d5fb92b3-6c6b-4130-8670-4e3263763401](http://portal.anvisa.gov.br/documents/10181/2721567/RDC_166_2017_COMP.pdf/d5fb92b3-6c6b-4130-8670-4e3263763401). Acesso em: 23 ago. 2018.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP. **Resolução nº 63**. Determina que os revendedores varejistas de combustíveis automotivos que comercializarem óleo diesel deverão confeccionar adesivos plásticos coloridos, afixando-os em local de destaque, a partir de 1º de janeiro de 2012, nas bombas abastecedoras de óleo diesel. Diário Oficial da União, Brasília, DF, 07 dez. 2011a.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP. **Resolução nº 65**. Regulamenta as especificações dos óleos diesel de uso rodoviário, consoante as disposições contidas no Regulamento Técnico ANP nº 08/2011 e as obrigações quanto ao controle da qualidade a serem atendidas pelos diversos agentes econômicos que comercializam o produto em todo o território nacional. Diário Oficial da União, Brasília, DF, 09 dez. 2011b.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP, **Regulação do biodiesel: especificação e controle de qualidade**. Rio de Janeiro, 2005. Disponível em: [http://www.inmetro.gov.br/noticias/conteudo/painelbio\\_rosangela.pdf](http://www.inmetro.gov.br/noticias/conteudo/painelbio_rosangela.pdf). Acesso em: 19 dez. 2018.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP). **Programa de Monitoramento da Qualidade de combustíveis (PMQC)**. Brasília: 2018b. Disponível em: <http://www.anp.gov.br/qualidade-produtos/158-programas-de-monitoramento/1864-pmqc>. Data de acesso: 10 jan. 2019.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP. **Biodiesel**. Disponível em: <http://www.anp.gov.br/biocombustiveis/biodiesel>. 2018c. Acesso em: 11 jan. 2019.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP.  
**Portaria nº 180, de 3/12/1998.** Estabelece a regulamentação para utilização de combustíveis líquidos ou gasosos não especificados. Diário Oficial da União, Brasília, DF. 4 dez. 1998. Disponível em: <http://legislacao.anp.gov.br/?path=legislacao-anp/portarias-anp/tecnicas/1998/dezembro&item=panp-180--1998>. Acesso em: 10 jan. 2019.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS -ANP.  
**Regulamento Técnico ANP nº 2/2016.** Diário Oficial da União, Brasília, DF. 11 abr. 2016. Disponível em: <http://legislacao.anp.gov.br/?path=legislacao-anp/resol-anp/2016/abril&item=ranp-17--2016>. Acesso em: 11 jan. 2019.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP.  
**Resolução nº 50 de 23/12/2013.** Regulamenta as especificações do óleo diesel de uso rodoviário, contidas no Regulamento Técnico ANP nº 4/2013, e as obrigações quanto ao controle da qualidade a serem atendidas pelos diversos agentes econômicos que comercializam o produto em todo o território nacional. Diário Oficial da União, Brasília, DF. 24 dez. 2013. Disponível em: <https://www.legisweb.com.br/legislacao/?id=263587/>. Acesso em: 10 jan. 2019.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP.  
**Resolução nº 69 de 23/12/2014.** Altera as Resoluções ANP nº 45 de 2012, nº 50 de 2013 e nº 50 de 2014. Diário Oficial da União, Brasília, DF. 24 dez. 2014. Disponível em: <https://www.legisweb.com.br/legislacao/?id=279318>. Acesso em: 9 jan. 2019.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP.  
**Resolução ANP nº 30/2016.** Estabelece a especificação de óleo diesel BX a B30, em caráter autorizativo, nos termos dos incisos I, II e III do art. 1º da Resolução CNPE nº 03, de 21 de setembro de 2015. Diário Oficial da União, Brasília, DF. 24 jun. 2016. Disponível em: [http://www.lex.com.br/legis\\_27160107\\_RESOLUCAO\\_N\\_30\\_DE\\_23\\_DE\\_JUNHO\\_DE\\_2016.aspx](http://www.lex.com.br/legis_27160107_RESOLUCAO_N_30_DE_23_DE_JUNHO_DE_2016.aspx). Acesso em: 11 jan. 2019.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS -ANP.  
**Resolução ANP nº 45/2014.** Estabelece a especificação do biodiesel contida no Regulamento Técnico ANP nº 3/2014 e as obrigações quanto ao controle da qualidade a serem atendidas pelos diversos agentes econômicos que comercializam o produto em todo o território nacional. Diário Oficial da União, Brasília, DF. 24 ago. 2014c. Disponível em: [http://www.lex.com.br/legis\\_25883261\\_Resolucao\\_N\\_45\\_DE\\_25\\_de\\_agosto\\_DE\\_2014.aspx](http://www.lex.com.br/legis_25883261_Resolucao_N_45_DE_25_de_agosto_DE_2014.aspx). Acesso em: 11 jan. 2019.

AGÊNCIA NACIONAL DE VIGILÂNCIA SANITARIA - ANVISA. **Guia para validação de métodos analíticos e bioanalíticos.** Brasília: ANVISA, 2003. Disponível em: <http://www.latinigroup.com/legislacao/RE899-2011.pdf>. Acesso em: 28 abr. 2018.

AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D5453:** Standard test method for determination of total sulphur in light hydrocarbons, spark ignition engine fuel, diesel engine fuel, and engine oil by ultraviolet fluorescence. West Conshohocken, PA ; 2016. 11p

AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D86:** Standard Test Method for Distillation of Petroleum Products at Atmospheric Pressure. West Conshohocken, PA; 2007. 28 p

AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D93:** Standard test method for flash point by pensky-martens closed cup tester. West Conshohocken, PA; 2011. 18 p.

AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D1500:** Standard Test Method for ASTM Color of Petroleum Products (ASTM Color Scale) West Conshohocken, PA; 2017. 5 p.

- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D6045**: Standard Test Method for Color of Petroleum Products by the Automatic Tristimulus Method. West Conshohocken, PA; 2017. 8 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D4052**: Standard test method for density, relative density, and api gravity of liquids by digital density meter. West Conshohocken, 2011. 8 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D1298**: Standard Test Method for Density, Relative Density, or API Gravity of Crude Petroleum and Liquid Petroleum Products by Hydrometer Method. West Conshohocken, 2017. 8 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D2622**: Standard Test Method for Sulfur in Petroleum Products by Wavelength Dispersive X-ray Fluorescence Spectrometry. West Conshohocken, 2016. 12 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D7039**: Standard Test Method for Sulfur in Gasoline, Diesel Fuel, Jet Fuel, Kerosine, Biodiesel, Biodiesel Blends, and Gasoline-Ethanol Blends by Monochromatic Wavelength Dispersive X-ray Fluorescence Spectrometry. West Conshohocken, 2015. 12 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D7212**: Standard Test Method for Low Sulfur in Automotive Fuels by Energy-Dispersive X-ray Fluorescence Spectrometry Using a Low-Background Proportional Counter. West Conshohocken, 2018. 6 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D7220**: Standard Test Method for Sulfur in Automotive, Heating, and Jet Fuels by Monochromatic Energy Dispersive X-ray Fluorescence Spectrometry. West Conshohocken, 2017. 7 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D3828**: Standard Test Methods for Flash Point by Small Scale Closed Cup Tester. West Conshohocken, 2016. 10 p.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. **D7094**: Standard Test Method for Flash Point by Modified Continuously Closed Cup (MCCCFP) Tester. West Conshohocken, 2017. 10 p.
- ARMENTA, S.; GARRIGUES, S.; GUARDIA, M. Green analytical chemistry, **Trends Analytical Chemistry**, v. 27, n. 6, p. 497-511, 2008.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR9619**: Produtos de petróleo - Destilação à pressão atmosférica. Rio de Janeiro, 2009. 42 p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR14065**: Destilados de petróleo e óleos viscosos - Determinação da massa específica e da densidade relativa pelo densímetro digital. Rio de Janeiro, 2013. 8 p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR 14483**: Produtos de petróleo - Determinação da cor - Método do colorímetro ASTM. Rio de Janeiro, 2015. 7 p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR14533**: Petróleo e produtos de petróleo — Determinação de enxofre por espectrometria de fluorescência de raios X (energia dispersiva). Rio de Janeiro, 2011. 10 p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR ISSO/IEC 17025:2005**. Disponível em: <http://www.smarnet.com.br/qualidade/metrologia/17025.pdf>. Acesso em: 30 out. 2018. 37 p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR 7148**: Petróleo e derivados de petróleo — Determinação da massa específica, densidade relativa e °API - Método do densímetro. Rio de Janeiro, 2013. 11 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR 15568**: Biodiesel - determinação do teor de biodiesel em óleo diesel por espectroscopia na região do infravermelho médio. Rio de Janeiro, 2008. 13 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR 14598**: Produtos de petróleo - determinação do ponto de fulgor pelo aparelho de vaso fechado Pensky- Martens. Rio de Janeiro, 2007. 15 p.

ASSOCIAÇÃO NACIONAL DOS FABRICANTES DE VEÍCULOS AUTOMOTORES - ANFAVEA. **Anuário da indústria automobilística brasileira**. 2018. Disponível em: <http://www.virapagina.com.br/anfavea2018/>. Acesso em: 14 dezembro 2018.

BANDEIRA, K. F. **Otimização e comparação de metodologia de extração de Casearinas em Casearia sylvestris**. Orientador: Alberto José Cavalheiro. 2004. 117 f. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Estadual Paulista, Araraquara.

BARMAN B. B. *et al.* Petroleum and coal. **Analytical Chemistry**, v. 73, n. 12, p. 2791-804, 2001.

BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. 25 Anos de quimiometria no Brasil. **Química Nova**, v. 29, n. 6, p. 1401-1406, 2006.

BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. **Como fazer experimentos: pesquisa e desenvolvimento na ciência e na indústria**. 2 ed. Campinas: Editora Unicamp, 2001. 401 p.

BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E., **Planejamento e otimização de experimentos**. Campinas: Editora da UNICAMP, 1995. 300 p.

BARTLE, K. D.; MYERS, P. History of gas chromatography. **Trends in Analytical Chemistry**, v. 21, n. 9, p. 547-556, 2002.

BAUMANN, K.; WAETZIG, H. Regression and calibration for analytical separation techniques. Part I: Design considerations. **Process Quality Control and Quality**, v. 10, n. 1/2, p. 59-73, 1997.

BEVAN, S. C. AND THORBURN, S. A new quantitative integral detector for gas chromatography: The brunel mass integral detector. **Journal of Chromatography**, v. 11, n. 11, p. 301-306, 1965.

BICCHI, C. *et al.* Direct resistively heated column gas chromatography (Ultrafast module-GC) for high-speed analysis of essential oils of differing complexities. **Journal of Chromatography A**, v.1024, n. 1/2, p. 195-207, 2004.

BICCHI, C. *et al.* Impact of phase ratio, polydimethylsiloxane volume and size, and sampling temperature and time on headspace sorptive extraction recovery of some volatile compounds in the essential oil field. **Journal of Chromatography A**, v. 1071, n. 1/2, p. 111-118, 2005.

BLUMBERG, L. M.; KLEE, M. S. Characteristic thermal constant and dimensionless heating rate. The links to optimum heating rate in GC. **Analytical Chemistry**, v. 72, n. 17, p. 4080-4089, 2000.

BORGES, E.M.; BOTOLLI, C.B.G., COLLINS, C.H. Possibilidades e limitações no uso da temperatura em cromatografia líquida de fase reversa. **Química Nova**, v. 33, n. 4, p. 945-953, 2010.

BRAITHWAITE, A.; SMITH, F. J. Gas chromatography. *In*: BRAITHWAITE, A; SAMITH, F. J. **Chromatographic methods**. 5. ed. Doradrecht: Academic Publishres, 1999, Cap. 5, p. 165-254.

BRANDAO, L. F. P.; BRAGAB, J. W. B.; SUAREZ, P. A. Z. Determination of vegetable oils and fats adulterants in diesel oil by high performance liquid chromatography and multivariate methods, **Journal Chromatography A**, v. 1225, p. 150-157, 2012.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento (MAPA). **Guia de Validação e Controle de Qualidade Analítica**. Brasília, DF. 2011, 78 p.

BRASIL. Ministério de Minas e Energia (MME). Grupo de Trabalho para Testes com Biodiesel. **Relatório de consolidação dos testes para validação da utilização de misturas com Biodiesel B10 em motores e veículos**. 2018. 163p. Disponível em: <http://www.mme.gov.br/documents/1138769/0/MME+Relatorio+B10+30-04-2018+final+v01-com+Anexos.pdf/6bed5a6b-7d04-41df-a5f3-649f31afd3e9>. Acesso em: 10 jan. 2019.

BRASIL. Presidência da República. Casa Civil. **Lei nº 11.097**. Dispõe sobre a adição obrigatória de biodiesel ao óleo diesel comercializado com o consumidor final; altera as Leis nºs 9.478, de 6 de agosto de 1997, e 8.723, de 28 de outubro de 1993; revoga dispositivos da Lei nº 11.097, de 13 de janeiro de 2005; e dá outras providências. Diário Oficial da União, Brasília, DF, 24 set. 2014a. Disponível em: <http://www2.camara.leg.br/legin/fed/lei/2014/lei-13033-24-setembro-2014-779372-norma-atualizada-pl.html>. Acesso em: 07 jan. 2019.

BRASIL. Conselho Nacional de Política Energética (CNPE). **Resolução nº 3 de 21 de setembro de 2015**. Autoriza e define diretrizes para comercialização e uso voluntário de biodiesel. Diário Oficial da União, Brasília, DF. 14 out. 2015. Disponível em: [http://www.mme.gov.br/documents/10584/2431527/Resolu%C3%A7%C3%A3o\\_3\\_CNPE\\_Biodiesel.pdf/d7480b33-c6b4-45fe-a20b-19b0e23f3293](http://www.mme.gov.br/documents/10584/2431527/Resolu%C3%A7%C3%A3o_3_CNPE_Biodiesel.pdf/d7480b33-c6b4-45fe-a20b-19b0e23f3293). Acesso em: 09 jan. 2019.

BRASIL. Conselho Nacional de Política Energética (CNPE). **Resolução nº 23 de 9 de novembro de 2017**. Estabelece a adição obrigatória, em volume, de dez por cento de biodiesel ao óleo diesel vendido ao consumidor final. Diário Oficial da União, 30 nov. 2017. Disponível em: <https://www.legisweb.com.br/legislacao/?id=353004>. Acesso em: 11 jan. 2019.

BRASIL. Conselho Nacional de Política Energética (CNPE). **Resolução nº 16, de 29 de outubro de 2018**. Dispõe sobre a evolução da adição obrigatória de biodiesel ao óleo diesel vendido ao consumidor final, em qualquer parte do território nacional. Diário Oficial da União, Brasília, DF, 7 nov. 2018. Disponível em: <https://ubrabiio.com.br/wp-content/uploads/2018/11/2018-11-08-Resolu%C3%A7%C3%A3o-CNPE-16-2018-Fixa-B11-a-B15.pdf> Acesso em: 24 jan. 2019

BRASIL. Presidência da República. Casa Civil. **Lei nº 13.033**. Dispõe sobre a adição obrigatória de biodiesel ao óleo diesel comercializado com o consumidor final; altera as Leis nos 9.478, de 6 de agosto de 1997, e 8.723, de 28 de outubro de 1993; revoga dispositivos da Lei no 11.097, de 13 de janeiro de 2005; e dá outras providências. Diário Oficial da União, Brasília, DF. 25 set. 2014b. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2011-2014/2014/Lei/L13033.htm](http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2014/Lei/L13033.htm). Acesso em: 11 jan. 2019.

BRESSOLE, F.; BROMET-PETIT; M.; AUDRAN, M. Validation of liquid chromatographic and gas chromatographic methods. Applications to pharmacokinetics. **Journal of Chromatography B**, v. 686, n. 1, p. 3-10, 1996.

BREITKREITZ, M. *et al.* Determination of total sulfur in diesel fuel employing NIR spectroscopy and multivariate calibration. **Analyst**, v. 128, p. 1204-1207, 2003.

BRITISH PETROLEUM (BP). **Statistical Review of World Energy 2018**. 67<sup>th</sup> Edition, 2018. 56 p. Disponível em: <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>. Acesso em: 10 jan. 2019.



- BUENO, L. S. R. **Estudo da influência da composição do óleo vegetal sobre algumas propriedades do biodiesel**. Orientador: José Antonio Andrés Velásquez Alegre 2007. 103 f. Dissertação (Mestrado em Engenharia Mecânica) - Centro de Ciências Exatas e Tecnológicas, Pontifícia Universidade Católica do Paraná, Curitiba.
- BYLESJÖ, M., O. CLOAREC, AND M. RANTALAINEN. Normalization and closure. *In*: BROWN, S. D.; SARABIA, L. A.; TRYGG, J. **Comprehensive chemometrics**. Amsterdam: Elsevier, 2009. Cap. 2.07, p. 110-124.
- CALADO, V.; MONTGOMERY, D. **Planejamento de experimentos usando o Statistica**. Rio de Janeiro: E-papers, 2003. 260 p.
- CAMPOS, M.C.V. **Estudo dos ácidos naftênicos do gasóleo pesado derivado do petróleo marlim**. Orientador: Elina Bastos Caramão 2005. 158 f Tese (Doutorado em Química) – Instituto de Química, Universidade Federal do Rio Grande do Sul, Porto Alegre.
- CERQUEIRA, E. O. *et al.* Utilização de filtro de transformada de Fourier para minimização de ruídos em sinais analíticos. **Química Nova**, v. 23, n. 5, p. 609-698, 2000.
- CHERPITSA, S. V. *et al.* Determination of inspection parameters of diesel fuels. **Chemistry and Technology of fuels and oils**, v. 39, n. 6, p. 364-370, 2003.
- CNT. **Os impactos da má qualidade do óleo diesel brasileiro**. Brasília, 2012. Disponível em: <http://cms.cnt.org.br/Imagens%20CNT/Site%202015/Pesquisas%20PDF/Os%20impactos%20da%20m%C3%A1%20qualidade%20do%20%C3%B3leo%20diesel%20brasileiro.pdf>. Acesso em: 18 dez 2019.
- COLLINS, C. H. I Mikhael Tswett e o nascimento da cromatografia. **Scientia Chromatographica**, v. 1, p. 1-21, 2009.
- COLLINS, C. H. 2006 - Cem anos das palavras cromatografia e cromatograma. **Química Nova**, v. 29, n. 4, p. 889-890, 2006.
- COLLINS, C. H.; BRAGA, G. L.; BONATO, P. S. **Fundamentos de cromatografia**. 4. ed. Campinas: Editora Unicamp, 2006. 453 p.
- CORGOZINHO, C.N.C.; PASA, V.M.D.; BARBEIRA, P.J.S. Determination of residual oil in diesel oil by spectrofluorimetric and chemometric analysis, **Talanta**, v. 76, n. 2, p. 479-484, 2008.
- CORREIA, P. R. M.; FERREIRA, M. M. C. Non-supervised pattern recognition methods: exploring chemometrical procedures for evaluating analytical data. **Química Nova**, v. 30, n. 2, p. 481-487, 2007.
- CORRÊA, O.L.S. **Petróleo: Noções sobre Exploração, Perfuração, Produção e Microbiologia**. 1. ed. Rio de Janeiro: Editora Interciência, 2003. 102 p.
- COSTA, A. C. A.; PEREIRA, N.; ARANDA, D. A. G. The situation of biofuels in Brazil: New generation technologies. **Renewable and Sustainable Energy Reviews**, v. 14, p. 3041-3049, 2010.
- CRAMER, J. A. *et al.* Novel data abstraction strategy utilizing gas chromatography-mass spectrometry data for fuel property modeling. **Energy Fuels**, v. 28, n. 3, p. 1781-1791, 2014.
- CRAMERS, C. A. *et al.* High-speed gas chromatography: an overview of various concepts. **Journal of Chromatography A**, v. 856, n. 1/2, p. 315-329, 1999.
- CRAMERS, C. A.; LECLERCQ, P. A. Strategies for speed optimisation in gas chromatography: an overview. **Journal of Chromatography A**, v. 842, n. 1/2, p. 3-13, 1999.

CUNHA, D. A. *et al.* NMR in the time domain: a new methodology to detect adulteration of diesel oil with kerosene, **Fuel**, v. 166, p. 79-85, 2016.

DAGAN, S.; AMIRAV, A. Fast gas chromatography/mass spectrometry analysis of drugs in urine with hyperthermal surface ionization in supersonic molecular beams. **European Journal of Mass Spectrometry**, v. 4, n. 1, p. 15-21, 1998.

DAGAN, S.; AMIRAV, A. Electron impact mass spectrometry of alkanes in supersonic molecular beams. **Journal of the American Society for Mass Spectrometry**, v. 6, n. 2, p. 120-131, 1995.

DASZYKOWSKI, M.; WALCZAK, B. Use and abuse of chemometrics in chromatography. **Trends in Analytical Chemistry**, v. 25, n. 11, p. 1081-1096, 2006.

DASZYKOWSKI, M.; WALCZAK, B.; MASSART, D. L. Representative subset selection. **Analytica Chimica Acta**, v. 468, n. 1, p. 91-103, 2002.

DAVID, F. *et al.* Instrumentation and applications of fast high-resolution capillary gas chromatography. **Journal of Chromatography A**, v. 842, n. 1/2, p. 309-319, 1999.

DAVIS, S. C.; MAKAROV, A. A.; HUGHES, J. D. Ultrafast gas chromatography using time-of-flight mass spectrometry. **Rapid Communication Mass Spectrometry**, v. 13, n. 4, 237-241, 1999.

DECLERCK, S. *et al.* Calibration curves in chromatography. *In*: KOMSTA, L.; HEYDEN, Y. V.; SHERMA, J. **Chemometrics in chromatography**., Boca Raton: CRC Press, 2018, Seção 2 – Univariate analysis, Cap. 4, p. 46-63.

DEJAEGHER, B.; VANDER HEYDEN, Y. Ruggedness and robustness testing, **Journal of Chromatography A**, v. 1158 n. 1/2, p. 138-157, 2007.

DEUTSCHES INSTITUT FÜR NORMUNG. **DIN EN 14078**: Liquid petroleum products - Determination of fatty acid methyl ester (FAME) content in middle distillates - Infrared spectrometry method. Berlin, 2014. 17p.

DORMAN, F. L. *et al.* Gas Chromatography. **Analytical Chemistry**, v. 80, n. 12, p. 4487-4487, 2008.

DOS SANTOS, B.C.D.B. **Desenvolvimento, aperfeiçoamento e validação de método cromatográfico para previsão da qualidade e propriedades físico-químicas do óleo diesel tipo B**. Orientador: José Eduardo de Oliveira. 2009. 188f. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Estadual Paulista, Araraquara.

DOS SANTOS, B.C.D.; FLUMIGNAN, D.L.; OLIVEIRA, J.E. Development, Optimization and Validation of Gas Chromatographic Fingerprinting of Brazilian Commercial Diesel Fuel for Quality Control. **Journal of Chromatography Science**, v. 50, n. 9, p. 756-762, 2012.

EHRMANN, E. U. *et al.* Novel column heater for fast capillary gas chromatography. **Journal of Chromatographic Science**, v. 34, n. 12, p. 533-539, 1996.

EICEMAN, G. A. **Instrumentation of Gas Chromatography**. Chichester: John Wiley & Sons, Inc., 2006, 9p.

EICEMAN, G. A. *et al.*, Gas chromatography. **Analytical Chemistry**. p. 76, n. 12, p. 3387-3394, 2004.

EMPRESA DE PESQUISA ENERGÉTICA - EPE. **Panorama do refino e da petroquímica no Brasil**. Rio De Janeiro, 2018. Disponível em: Acesso em: [http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-topico-412/NT%20Refino%20e%20Petroqu%C3%ADmica\\_2018.11.01.pdf](http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-topico-412/NT%20Refino%20e%20Petroqu%C3%ADmica_2018.11.01.pdf). Acesso em: 18 dez 2019.

EURACHEM. Guide: **The fitness for purpose of analytical methods**: a laboratory guide to method validation and related topics. Magnusson, B. and U. Örnemark (Ed.), 2nd Edition, 2014. 70p. Disponível em: [https://www.eurachem.org/images/stories/Guides/pdf/MV\\_guide\\_2nd\\_ed\\_EN.pdf](https://www.eurachem.org/images/stories/Guides/pdf/MV_guide_2nd_ed_EN.pdf). Acesso em: 23 ago. 2018

ERMER, J.; LANDY, J. S. Validation of analytical procedures. *In*: SWARBICK, J. **Encyclopedia of Pharmaceutical Technology**. 3 ed. Pinehurst: Marcel Dekker, Inc. 2007. volume 1, Cap. 2, p. 92–113.

ETTRE, L. S. M. S. Tswett and the Invention of Chromatography Part I: Life and Early Work (1872–1903). *In*: ETTRE, L. S. **Chapters in the evolution of chromatography**. London: Imperial College Press, 2008a. Cap. 4, p. 49-58.

ETTRE, L. S. Early Petroleum Chromatographers. *In*: ETTRE, L. S. **Chapters in the evolution of chromatography**. London: Imperial College Press, 2008b. Cap. 3, p. 31-44.

ETTRE, L. S. M. S. Tswett and the Invention of Chromatography Part II: Completion of the Development (1903–1910). *In*: ETTRE, L. S. **Chapters in the evolution of chromatography**. Londres: Imperial College Press, 2008c. Cap. 6, p. 60-74.

ETTRE, L. S. Evolution of capillary columns for gas chromatography. **LCGC**, v.19, n. 1, p. 48-59. 2001.

FARAH, M.A. **Petróleo e seus derivados: definição, constituição, aplicações, especificações, características de qualidade**. Rio de Janeiro: Editora LTC, 2012. 255 p.

FARIA, R.C.M.; REZENDE, M.J.C.; PINTO, A.C. Desenvolvimento e validação de metodologia de análise de misturas biodiesel:diesel utilizando cromatografia gasosa-espectrometria de massas. **Química Nova**, v. 30, n. 8, p. 1900-1905, 2007.

FERREIRA, M. M. C. **Quimiometria**: conceitos, métodos e aplicações. Campinas: Editora da Unicamp, Campinas, 2015. 496 p.

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

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S. Chemometrics I: multivariate calibration, a tutorial. **Química Nova**, v. 22, n. 5, p. 724-731, 1999.

FERRARI, R.A.; OLIVEIRA, V.S.; SCABIO, A. Biodiesel de soja – Taxa de conversão em ésteres etílicos, caracterização físico-química e consumo em gerador de energia. **Química Nova**, v. 28, n. 1, 19-23, 2005.

FELIZZOLA, J. F.; COSTA, C. F. G. da; VASCONCELOS, S. S. **Passo a passo para o uso do cromatógrafo gasoso modelo GC-CP3800 Varian para análises de gases de efeito estufa (GEEs)**. Belém: Embrapa Amazônia Oriental, 2014, 70 p.

FLUMIGNAN, D. **Caracterização da qualidade e interpretação das propriedades físico-químicas de gasolinas C brasileiras através de cromatografia gasosa e métodos quimiométricos**. Orientador: José Eduardo de Oliveira. 2005. 140f. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Estadual Paulista, Araraquara.

FREYE, C. E. *et al.* Partial least squares analysis of rocket propulsion fuel data using diaphragm valve-based comprehensive two-dimensional gas chromatography coupled with flame ionization detection. **Talanta**, v. 153, p. 203-213, 2016.

GALDAMEZ, E.V.C. **A aplicação das técnicas de planejamento e análise de experimentos na melhoria da qualidade de um processo de fabricação de produtos plásticos.**

Orientador: Luiz César Ribeiro Carpinetti. 2002. 133f. Dissertação (Mestrado em Engenharia de Produção) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.

GASPARINI, F. **Avaliação e adaptação das condições da EN 14103 para quantificação de ésteres em biodieseis etílicos puros de soja, babaçu, palma e sebo bovino e em suas misturas.**

Orientador: José Eduardo de Oliveira. 2010. 99 f. Dissertação (Mestrado em Química) - Instituto de Química, Universidade Estadual Paulista, Araraquara, 2010.

GELADI, P.; KOWALSKI, B. R.; Partial least-squares regression: a tutorial. **Analytica Chimica Acta**, v. 185, p. 1-17, 1986.

GODING J. C. *et al.* Comparison of GC stationary phases for the separation of fatty acid methyl esters in biodiesel fuels. **Analytical and Bioanalytical Chemistry**, 405, n. 18, p. 6087-94, 2013.

GROB, R. L. **Theory of gas chromatography.** In: GROB, R. L. Modern practice of gas chromatography. 4. ed. Hoboken: John Wiley & Sons, 2004, Cap. 2, p. 25-64.

HARLEY, J.; NEL, W.; PRETORIUS, V. Flame ionization detector for gas chromatography, **Nature**, v. 181, N. 4603, P. 177-178, 1958.

HINSHAW, J. V. Handling Fast Peaks. **LCGC**, v. 21, n. 3, p. 268-272, 2003.

HINSHAW, J. V. How fast is fast enough? **LCGC**, v. 19, n. 2, p. 170-171, 2001.

HOLANDA, A. **Biodiesel e inclusão social.** Brasília: Câmara dos Deputados: Coordenação de Publicações, 2004. 200 p.

HORVATH, C., ETTRE, L. S.; PURCELL, J. E. Support coated open tubular columns: decade of progress. **American Laboratory**, v. 6, n. 8, p. 75-86, 1974.

HUBER, L., Validation of analytical methods: review and strategy. **LC-GC**, v. 11, p. 96-105, 1998.

HUPP, A. B. *et al.* Analysis of biodiesel-diesel blends using ultrafast gas chromatography (UFGC) and chemometric methods: Extending ASTM D7798 to biodiesel. **Fuel**, v. 231, p. 264-270, 2018.

HUPP, A. M. *et al.* Chemometric analysis of diesel fuel for forensic and environmental applications. **Analytica Chimica Acta**, v. 606, p. 159-171, 2009.

INFOMETRIX. **Manual Pirouette user guide. Version 4.51.** Woodinville: 2018. Disponível em: <https://chemometrix.files.wordpress.com/2016/07/pirouette.pdf>. Acesso em 15 mai. 2018.

INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL - INMETRO. **Orientação sobre validação de métodos analíticos.** Rio de Janeiro: 2018.

Disponível em:

[http://www.inmetro.gov.br/Sidoq/pesquisa\\_link.asp?seq\\_tipo\\_documento=5&cod\\_uo\\_numeracao=00581&num\\_documento=008.pdf](http://www.inmetro.gov.br/Sidoq/pesquisa_link.asp?seq_tipo_documento=5&cod_uo_numeracao=00581&num_documento=008.pdf) Acesso em: 28 ago. 2018.

IUPAC. **Compendium of chemical terminology:** the golden book, version 2.3.3. 2014. Disponível em: <http://goldbook.iupac.org/CT06948.html>. Acesso em: 20 out. 2018.

KANDUC, K.R.; ZUPAN, J.; MAJCEN, N.; Separation of data on the training and test set for modelling: a case study for modelling of five colour properties of a white pigment. **Chemometric Intelligent Laboratory System**, v. 65, n. 2, p. 221-229, 2003.

KENNARD, R. W.; STONE, L. A. Computer aided design of experiments. **Technometrics**. v. 11, n. 1, p.137-148, 1969.

KOMSTA, L.; HEYDEN, Y. V. Linear supervised techniques. *In*: KOMSTA, L.; HEYDEN, Y. V.; SHERMA, J. **Chemometrics in chromatography**., Boca Raton: CRC Press, 2018, Seção 4 – Applications overviews, Cap. 14, p. 255-266.

KORYTÁR, P. *et al.* Practical fast gas chromatography: methods, instrumentation and applications. *Trends in Analytical Chemistry*, v. 21, n. 9/10, p. 558-572, 2002.

KRAMER, R. **Chemometric techniques for quantitative analysis**. New York: Marcel Dekker Inc., 1998, 203 p.

LANÇAS, F. M. **Fundamentos da cromatografia gasosa**. Campinas: Editora Átomo: 2017. 237 p.

LECLERCQ, P. A.; CRAMERS, C. A. Optimum performance of capillary GC columns as a function of tube diameter and film thickness under various operating conditions. Computer program for calculation of H- $\bar{u}$  curves and minimum analysis times. **Journal of High Resolution Chromatography, Chromatography Communications**, v. 8, n. 11, p. 764-771, 1985.

LEITE, F. **Validação em análise química**. 5 ed. Campinas: Editora Átomo, 2008, 357 p.

LIMA, A. S. O Programa de Monitoramento da Qualidade dos Combustíveis – PMQC. **Séries Temáticas ANP**, n. 5, p. 9, 2012. Disponível em: <http://www.anp.gov.br/publicacoes/livros-e-revistas/2379-o-programa-de-monitoramento-da-qualidade-dos-combustiveis-pmqc>. Acesso em: 9 jan. 2019.

LUONG, J.; GRAS, R.; JENNINGS, W.; An advanced solventless column test for capillary GC columns. **Journal of Separation Science**, v. 30, n. 15, p. 2480-2492, 2007.

MACEDO, R. M. P. de. **Aplicação da metodologia de planejamento de experimentos para formulação de massas cerâmicas para telhas**. orientador: Rubens Maribondo do Nascimento. 2007. 103f. tese (doutorado em ciência e engenharia de materiais) – Universidade Federal do Rio Grande do Norte, Natal.

MALINOWSKI, E. R. **Factor Analysis in Chemistry**. 3. ed. New York: John & Wiley, Inc., 2002. 432 p.

MARIANO, J.B. **Impactos ambientais do refino do petróleo**. Orientador: Roberto Schaeffer 2001. 216f. Dissertação (Mestrado em Ciências em Planejamento Energético) Universidade Federal do Rio de Janeiro, Rio de Janeiro.

MARSHALL, L. J. *et al.* Association and discrimination of diesel fuels using chemometric procedures. **Analytical and Bioanalytical Chemistry**, 394, n. 8, p. 2049-59, 2009.

MASSART, D. L. *et al.* **Handbook of chemometrics and qualimetrics**: data handling in science and technology., Amsterdam: Elsevier, 1997. Volumes 20A, 867 p.

MASSART, D. L.; BUYDENS, L. **Journal of Pharmaceutical and Biomedical Analysis**. v. 6, n. 6/8, p. 535-454, 1988.

MASOOD, M. A. SALEM, N. High-throughput analysis of plasma fatty acid methyl esters employing robotic transesterification and fast gas chromatography. **Lipids**, v. 43, n. 2, p. 171-180, 2008).

MATISOVÁ, E.; DOMOTOROVÁ, M. Fast gas chromatography and its use in trace analysis. **Journal of Chromatography A**, v.1000, n. 1/2, p. 199-221, 2003.

MASTOVSKA, K.; LEHOTAY, S. J. Practical approaches to fast gas chromatography-mass spectrometry. **Journal of Chromatography A**, v. 1000, n. 1/2, p. 153-180, 2003.

MATOS, G. D. *et al.* Análise exploratória em química analítica com emprego de quimiometria: PCA e PCA de imagens. **Revista Analytica**, n. 6, p. 38-50, ago. 2003.

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**, v. 180, p. 239-47, 2018.

MENEZES, E.W.; CATALUÑA, R. Amostragem do material particulado e fração orgânica volátil das emissões em motor ciclo diesel sem a utilização de túnel de diluição. **Química Nova**, v. 31, n. 8, p. 2027-2030, 2008.

MELO, F. O. T.; PAULILLO, VIAN, C. E. F. O biodiesel no Brasil: panorama, perspectivas e desafios. **Informações Econômicas**, v.37, n.1, p.1-13, 2007.

MCWILLIAM, I. G.; DEWAR, R. A. Flame Ionization Detector for Gas Chromatography. **Nature**, v. 181, n. 4611, p. 760, 1958.

MCNAIR, H. M.; MILLER, J. M. **Basic gas chromatography**. 2. ed. Hoboken: John Wiley & Sons, Inc. 2009. 227 p.

MONDELLO, L. *et al.* Rapid micro-scale separation and very fast gas chromatographic separation of cod liver fatty acid methyl esters. **Journal Pharmaceutical Biomedical Analysis**, v. 41, n. 5, p. 1566-1570, 2006.

MOITA NETO, J. M. M.; MOITA, G. C. Uma introdução à análise exploratória de dados multivariados. **Química Nova**, v. 21, n. 4, p. 467-469, 1998.

MORAIS, J.M. **Petróleo em águas profundas**. Brasília: Ipea: Petrobrás. 2013, 424 p.

MUNHOZ, J.F.V.L. **Método Alternativo para Detecção de Não Conformidade em Gasolina Comercial Brasileira do Tipo C: Uso da Cromatografia Ultrarrápida Aliada a Ferramentas Quimiométricas**. Orientador: José Eduardo de Oliveira. 2013. 200f. Dissertação (Mestrado em Química) – Instituto de Química Universidade Estadual Paulista, Araraquara.

NAES, T. *et al.* **A user-friendly guide to multivariate calibration and classification**, Chichester: NIR Publications, 2002.

NEIVA, J. **Conheça o petróleo**, Rio de Janeiro: Editora Ao Livro Técnico, 5ª Edição, 1986. 306 p.

NOMAN, M.T., ASHRAF, M.A., JAMSHAD, H., ALI, A. A novel green stabilization of TiO<sub>2</sub> nanoparticles onto cotton. **Fibers and Polymers** 19(11): 2268-2277, 2018.

NOVAES, C. G. *et al.* Otimização de métodos analíticos usando metodologia de superfícies de resposta - parte i: variáveis de processo. **Revista Virtual de Química**, v. 9, n. 3, p. 1184-1215, 2017.

PARREIRA, T. F. **Utilização de métodos quimiométricos em dados de natureza multivariada**. 2003. 91 f. Dissertação (Mestrado em Química) -Instituto de Química, Universidade Estadual de Campinas, Campinas, 2003.

PACHECO, S. *et al.* History of liquid chromatography. **Revista Virtual de Química**, v. 7, n. 4, p. 1225-1271, 2015.

PAULS, R. E. A Review of chromatographic characterization techniques for biodiesel and biodiesel blends. **Journal of Chromatographic Science**, v. 49, n. 5, p. 384-396, 2011a.

- PAULS, R. E. Fast Gas Chromatographic Separation of Biodiesel. **Journal of Chromatographic Science**, v. 49, n. 5, p. 370-374, 2011b.
- PEREIRA-FILHO, E.R. **Planejamento fatorial em química: maximizando a obtenção de resultados**. São Carlos: Edufscar, 2015. 90 p. (Série Apontamentos)
- PHILLIPS, J. B.; JAIN, V. On-column temperature programming in gas chromatography using temperature gradients along the capillary column. **Journal of Chromatographic Science**, v. 33, n. 10, p. 541-550, 1995.
- PIANOVISKI JUNIOR, G. **Utilização de misturas de óleo diesel e ésteres etílicos de óleo de soja como combustíveis alternativos: análise do desempenho e do processo de combustão no motor diesel**. Orientador: José Antonio Andrés Velásquez Alegre 2002. 132 f. Dissertação (Mestrado em Engenharia Mecânica) - Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Paraná, Curitiba.
- PIROUETTE. Register Software. License number 29d8e5782747b0b36f7. Serial number, 0506P311007. Woodinville. Infometrix. 2003.
- POVOLO, M.; PELIZZOLA, V.; CONTARINI, G. Directly resistively heated-column gas chromatography for the evaluation of cow milk fat purity. **European Journal of Lipid Science and Technology**, v. 110, n. 11, p. 1050-1057, 2008.
- QUINTELLA, C. M. **Química Nova**, v. 19, n. 6, p. 660-667, 1996.
- RAMOS, L.P. *et al.* Biodiesel: Um projeto de sustentabilidade econômica e sócio-ambiental para o Brasil. **Revista Biotecnologia Ciência & Desenvolvimento**, v. 31, p. 28-37, 2003.
- ROCHA, W. F. C.; SHEEN, D. A. Determination of physicochemical properties of petroleum derivatives and biodiesel using GC/MS and chemometric methods with uncertainty estimation. **Fuel**, v. 243, p. 413-422, 2019.
- RIBANI, M. *et al.* Validacao em métodos cromatográficos e eletroforéticos. **Química Nova**, v. 27, n. 5, p. 771-780, 2004.
- RIBEIRO, R.A. **Análise espacial de reservatórios de petróleo usando DFA de dados geofísicos**. Orientador: Liacir dos Santos Lucena. 2014. 113f. Tese (Doutorado em Ciências e Engenharia do Petróleo) – Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, Natal.
- ROCHA, D.Q. *et al.* Determinação da matéria-prima utilizada na produção de biodiesel adicionado ao diesel mineral através de monitoramento seletivo de íons. **Química Nova**, v. 31, n. 5, p. 1062-1066, 2008.
- ROSA, T.; RODRIGUES, J.G.A.; FERREIRA, R.Q. Uso do planejamento experimental para otimização de um procedimento voltamétrico para determinação simultânea das concentrações dos metais Zn, Cd, Pb e Cu livres em água de coco. **Química Nova**, v. 39, n. 2, p. 221-228, 2016.
- SALES, F. *et al.* Standardization of a multivariate calibration model applied to the determination of chromium in tanning sewage, **Talanta**, v. 52, n. 2, p. 329-336, 2000.
- SCANLON, J. T., WILLIS, D. E. Calculation of flame ionization detector relative response factors using the effective carbon number concept. **Journal of Chromatographic Science**, v. 23, n. 8, p. 333-340, 1985.
- SCOTT, R. P. W. **Detectors: design, function and operation**. New York: Marcel Dekker, Inc., 1996, 514 p.

- SEQUINEL, R. *et al.* Cromatografia gasosa ultrarrápida: uma visão geral sobre parâmetros, instrumentação e aplicações. **Química Nova**, v.33, n.10, p. 2226-2232, 2010.
- SHARAF, M. A.; ILLMAN, D. L.; KOWALSKI, B. R. Exploratory data analysis. *In*: SHARAF, M. A.; ILLMAN, D. L.; KOWALSKI, B. R. **Chemometrics**. New York: John Wiley & Sons, 1986. v. 82, cap. 6, p. 219-227.
- SHIBAHARA, A.; YAMAMOTO, K.; KINOSHITA, A. High-speed analysis of the major components of fatty acid methyl esters by capillary gas chromatography. **Lipid Technology**, V. 20, n. 4, p. 88-90, 2008.
- SHELL INTERNATIONAL PETROLEUM COMPANY LIMITED. **The petroleum handbook**. 6 ed. Amsterdam: Elsevier, 1983. 729 p.
- SILVA, A.S. **Avaliação da secagem de bagaço de cajá usando planejamento fatorial composto central**. Orientador: Edson Leandro de Oliveira. 2008. 70f. Dissertação (Mestrado em Engenharia Química) – Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, Natal.
- SILVÉRIO, F. O.; BARBOSA, L. C. A. A pirólise como técnica analítica. **Química Nova**, v. 31, n. 6, p. 1543-1552, 2008.
- SIMON, Jr. R. K. Optimization of a flame ionization detector for capillary chromatography. *Journal of Chromatographic Science*, v. 23, n. 7, p. 313-319, 1985.
- SMOLKOVÁ-KEULEMANSOVÁ, E. A Few Milestones on the Journey of Chromatography. **Journal of Separation Science**, v.23, n. 7/8, p. 497-501, 2000.
- SMOLIŃSKI, A.; WALCZAK, B.; EINAX, J. W. Hierarchical clustering extended with visual complements of environmental data set, **Chemometrics and Intelligent Laboratory Systems**, v. 64, n. 1, p. 45–54, 2002
- SOUZA, L. C. *et al.* Desenvolvimento de modelos de calibração nirs para minimização das análises de madeiras de *Eucalyptus* spp. **Ciência Florestal**, v. 21, n. 3, p. 591-599, 2011.
- SOUZA, A. M.; POPPI, R. J. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial, parte I. **Química Nova**, v. 35, n. 1, p. 223-229, 2012.
- SPEIGHT, J.G. **The chemistry and technology of petroleum**. 4. ed. Boca Raton: CRC Press, 2006. 954 p.
- SLOAN, K. M.; MUSTACICH, R. V.; ECKENRODE, B. A. Development and evaluation of a low thermal mass gas chromatograph for rapid forensic GC–MS analyses. **Field Analytical Chemistry & Technology**, v. 5, n. 6, p. 288-301, 2001
- SLUSZZ, T.; MACHADO, J.A.D. **Características das potenciais culturas matérias-primas do biodiesel e sua adoção para agricultura familiar**. *IN*: Encontro de Energia no Meio Rural, 6. Campinas, 2006. Disponível em: <http://www.proceedings.scielo.br/scielo.php?> Acesso em: 24 jan. 2019.
- SZKLO, A.; ULLER, V. C. **Fundamentos do Refino de Petróleo**. Rio de Janeiro: Editora Interciência, 2008. 346 p.
- TERCINI, A. C. B. *et al.* Ultrafast gas chromatographic method for quantitative determination of total FAMES in biodiesel: An analysis of 90-s. **Fuel**, v. 222, p. 792-799, 2018.
- THERMO SCIENTIFIC. **Trace determination of organo-phosphorous pesticides in olive oil by gc analysis through PTV backflush/FPD**. Application



Note: 10049. Milan: 2007. Disponível em:

<https://tools.thermofisher.com/content/sfs/brochures/Trace-Determination-of-Organophosphorous-Pesticides-in-Olive-Oil-by-GC-Analysis.pdf>. Acesso em: 22 abr. 2018.

THERMO SCIENTIFIC. **Trace gc ultra ufm ultra fast module device**. Instruction Manual PN 31709398. Milan: 2010. Disponível em: <https://assets.thermofisher.com/TFS-Assets/CMD/manuals/Man-31709398-TRACE-GC-Ultra-UFM-Man31709398-EN.pdf>. Acesso em: 20 abr. 2018.

THOMAS, J. E. **Fundamentos de Engenharia do Petróleo**. Rio de Janeiro: Editora Interciência, 2001. 281 p.

TRYGG, J.; WOLD, S. Orthogonal projections to latent structures (o-pls). **Journal of Chemometrics**, v. 16, p. 119–128, 2002.

TSWETT, M. S. Physikalisch-chemische Studien fiber das Chlorophyll. **Die Adsorption. Berichte der Deutschen Botanischen Gesellschaft**, v. 24, p. 316-323, 1906a.

TSWETT, M. S. Adsorptions analyse und chromatographische Methode. **Anwendung auf die Chemic des Chlorophylls Berichte der Deutschen Botanischen Gesellschaft**, v. 24, p. 384-393, 1906b.

TRANCHIDA, P. Q. *et al.* Evaluation of use of a very short polar microbore column segment in high-speed gas chromatography analysis. **Journal Separation Science**, v. 32, n. 14, p. 2634-2639, 2008.

UNIÃO BRASILEIRA DO BIODIESEL E BIOQUEROSENE - UBRABIO. **CNPE publica resolução do B15**. Disponível em: <https://ubrablo.com.br/2018/11/08/cnpe-publica-resolucao-do-b15/>. Acesso em: 24 jan. 2019.

van DEURSEN, *et al.* Evaluation of time-of-flight mass spectrometric detection for fast gas chromatography. **Journal of Chromatography A**, v. 878, n. 2, p. 205-213, 2000.

van DEURSEN, M. *et al.* **Journal of Microcolumn Separations**, v. 13, n. 8, p. 337-345, 2001.

van DEURSEN, M. *et al.* Possibilities and Limitations of Fast Temperature Programming as a Route towards Fast GC. **Journal High Resolution Chromatographic**, v. 22, n. 9, p. 509-513, 1999.

VANDER-HEYDEN, Y.; NIJHUIS, A.; SMEYERS-VERBEKE, J. Guidance for robustness/ruggedness tests in method validation, **Journal of Pharmaceutical and Biomedical Analysis**, v. 24, n. 5/6, p. 723-753, 2001.

VIAENE, J.; HEYDEN, Y. V. Introduction to herbal fingerprinting by chromatography. *In*: KOMSTA, L.; HEYDEN, Y. V.; SHERMA, J. **Chemometrics in chromatography**., Boca Raton: CRC Press, 2018, Seção 6 – Applications overviews, Cap. 19, p. 351-370.

VEMPATAPU, B. P.; KANAUIA, P. K. Monitoring petroleum fuel adulteration: a review of analytical methods. **Trends Analytical Chemistry**, v. 92, p. 1–11, 2017.

WEBER B. M.; WALSH P.; HARYNUK, J. J. Determination of hydrocarbon group-type of diesel fuels by gas chromatography with vacuum ultraviolet detection. **Analytical Chemistry**, v.88, n. 11, p. 5809-5817, 2016.

WOLD, S. Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models. **Technometrics**, v. 20, n. 4, p. 397-405, 1978.

WOLD, S.; ESBENSEN, K.; GELADI, P. Principal component analysis. **Chemometrics and Intelligent Laboratory Systems**, v. 2, n. 1/3, p. 37–52, 1987.

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 58, n. 2, p. 109–130, 2001.

ZANÃO, L. R. *et al.* Prediction of relative density, distillation temperatures, flash point and cetane number of S500 diesel oil using multivariate calibration of gas chromatographic profiles. **Energy Fuels**, v. 32, n. 8, p. 8108-8114, 2018.

## APÊNDICE A – TESTE DE COCHRAN

O teste de Cochran compara as variâncias de cada ponto da curva de calibração com a maior variância encontrada com o objetivo de avaliar a suposição de que os erros experimentais tenham variâncias iguais. A primeira etapa consiste em calcular a estatística (C), dada por:

$$C = \frac{s_{max}^2}{k} = \frac{\text{maior variância}}{\sum_{i=1}^k s_i^2},$$

- onde **k** representa o número de níveis do fator;
- $s_i^2$ : representa a variância amostral  $s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$ ;
- **n**: representa o número de medidas em cada nível do fator.

Os valores obtidos são analisados com 5% de significância através do valor de P.

As análises foram realizadas pelo *software* estatístico *ActionStat*, de onde também estas informações foram retiradas.

### Avaliação da Homocedasticidade: Teste de Cochran – C<sub>8</sub>H<sub>18</sub>

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
0,5803	3	0,0784

Como P-valor (0,0784) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

### Avaliação da Homocedasticidade: Teste de Cochran - C<sub>10</sub>H<sub>22</sub>

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
0,6121	3	0,0527

Como P-valor (0,0527) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

### **Avaliação da Homocedasticidade: Teste de Cochran - C<sub>12</sub>H<sub>26</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
0,5786	3	0,0798

Como P-valor (0,0798) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

### **Avaliação da Homocedasticidade: Teste de Cochran - C<sub>14</sub>H<sub>30</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
0,1667	3	1,00

Como P-valor (1,00) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

### **Avaliação da Homocedasticidade: Teste de Cochran - C<sub>16</sub>H<sub>34</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
0,5637	3	0,0948

Como P-valor (0,0948) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

### **Avaliação da Homocedasticidade: Teste de Cochran - C<sub>18</sub>H<sub>38</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
--------------------	-----------	----------------

0,5887	3	0,0706
--------	---	--------

Como P-valor (0,0706) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

#### **Avaliação da Homocedasticidade: Teste de Cochran – C<sub>20</sub>H<sub>42</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<u>Estatística</u>	<u>GL</u>	<u>P.valor</u>
0,4388	3	0,3341

Como P-valor (0,3341) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

#### **Avaliação da Homocedasticidade: Teste de Cochran – C<sub>22</sub>H<sub>46</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<u>Estatística</u>	<u>GL</u>	<u>P.valor</u>
0,5182	3	0,1558

Como P-valor (0,1558) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

#### **Avaliação da Homocedasticidade: Teste de Cochran – C<sub>24</sub>H<sub>50</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<u>Estatística</u>	<u>GL</u>	<u>P.valor</u>
0,5693	3	0,0889

Como P-valor (0,0889) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

### **Avaliação da Homocedasticidade: Teste de Cochran – C<sub>26</sub>H<sub>54</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
0,5409	3	0,1224

Como P-valor (0,1224) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

### **Avaliação da Homocedasticidade: Teste de Cochran – C<sub>28</sub>H<sub>58</sub>**

Para avaliarmos a homocedasticidade da variância realizamos o seguinte teste de hipóteses:

H0: Variâncias dos níveis são iguais;

H1: Pelo menos uma variância diferente.

A seguir, apresentamos o teste de Teste de Cochran.

<b>Estatística</b>	<b>GL</b>	<b>P.valor</b>
0,6073	3	0,0561

Como P-valor (0,0561) do Teste de Cochran é maior que 0,05 (conforme proposto), não rejeitamos a hipótese de igualdade das variâncias ao nível de significância de 5%. Logo, temos um modelo homocedástico.

## APÊNDICE B - TABELA ANOVA

TABELA ANOVA -  $C_8H_{18}$

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	319.977.724.406.374,00000	1	319.977.724.406.374,00000	4,290	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 955$
Resíduo Sqres	1.193.471.106.841,600000	16	74.591.944.177,600100		<b>Modelo confiável para previsões.</b>
Total SQT	321.171.195.513.216,00000	17	-	-	-
Falta de ajuste SQFaj	481.847.564.325,602000	4	120.461.891.081,400000	2,03	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,62$
Erro Puro SQEP	711.623.542.516,000000	12	59.301.961.876,333300		<b>Não apresenta falta de ajuste.</b>
r	0,9992				
R <sup>2</sup>	0,9985				

TABELA ANOVA -  $C_{10}H_{22}$

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	285.509.323.157.618,00000	1	285.509.323.157.618,00000	14,801	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 3.293$
Resíduo Sqres	308.645.330.024,038000	16	19.290.333.126,502400		<b>Modelo confiável para previsões.</b>
Total SQT	285.817.968.487.642,00000	17	-	-	-
Falta de ajuste SQFaj	126.572.031.264,704000	4	31.643.007.816,176100	2,09	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,64$
Erro Puro SQEP	182.073.298.759,333000	12	15.172.774.896,611100		<b>Não apresenta falta de ajuste.</b>
r	0,9998				
R <sup>2</sup>	0,9996				

TABELA ANOVA -  $C_{12}H_{26}$

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	337.185.731.277.204,00000	1	337.185.731.277.204,00000	4,658	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 1.037$
Resíduo Sqres	1.158.138.921.182,120000	16	72.383.682.573,882600		<b>Modelo confiável para previsões.</b>
Total SQT	338.343.870.198.386,00000	17	-	-	-
Falta de ajuste SQFaj	577.931.360.819,454000	4	144.482.840.204,864000	2,99	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,92$
Erro Puro SQEP	580.207.560.362,667000	12	48.350.630.030,222200		<b>Não apresenta falta de ajuste.</b>
r	0,9991				
R <sup>2</sup>	0,9982				

TABELA ANOVA -  $C_{14}H_{30}$ 

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	303.830.584.263.801,00000	1	303.830.584.263.801,00000	12,523	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 2,787$
Resíduo Sqres	388.196.952.248,107000	16	24.262.309.515,506700		<b>Modelo confiável para previsões.</b>
Total SQT	304.218.781.216.049,00000	17	-	-	-
Falta de ajuste SQFaj	83.036.137.180,773800	4	20.759.034.295,193500	0,82	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,25$
Erro Puro SQEP	305.160.815.067,333000	12	25.430.067.922,277800		<b>Não apresenta falta de ajuste.</b>
r	0,9998				
R <sup>2</sup>	0,9997				

TABELA ANOVA -  $C_{16}H_{34}$ 

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	257.994.806.616.242,00000	1	257.994.806.616.242,00000	103,075	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 22,936$
Resíduo Sqres	40.047.553.423,315400	16	2.502.972.088,957210		<b>Modelo confiável para previsões.</b>
Total SQT	258.034.854.169.666,00000	17	-	-	-
Falta de ajuste SQFaj	10.958.786.276,648800	4	2.739.696.569,162200	1,13	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,35$
Erro Puro SQEP	29.088.767.146,666700	12	2.424.063.928,888890		<b>Não apresenta falta de ajuste.</b>
r	0,9995				
R <sup>2</sup>	0,9990				

TABELA ANOVA -  $C_{18}H_{38}$ 

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	248.617.393.071.736,00000	1	248.617.393.071.736,00000	11,754	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 2,615$
Resíduo Sqres	338.430.034.330,069000	16	21.151.877.145,629300		<b>Modelo confiável para previsões.</b>
Total SQT	248.955.823.106.066,00000	17	-	-	-
Falta de ajuste SQFaj	137.487.092.498,070000	4	34.371.773.124,517500	2,05	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,63$
Erro Puro SQEP	200.942.941.832,000000	12	16.745.245.152,666700		<b>Não apresenta falta de ajuste.</b>
r	0,9997				
R <sup>2</sup>	0,9994				



TABELA ANOVA -  $C_{20}H_{42}$ 

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	204.007.838.397.598,00000	1	204.007.838.397.598,00000	3,745	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 833$
Resíduo Sqres	871.567.310.422,153000	16	54.472.956.901,384600		<b>Modelo confiável para previsões.</b>
Total SQT	204.879.405.708.020,00000	17	-	-	-
Falta de ajuste SQFaj	361.790.206.656,820000	4	90.447.551.664,205000	2,13	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,65$
Erro Puro SQEP	509.777.103.765,333000	12	42.481.425.313,777800		<b>Não apresenta falta de ajuste.</b>
r	0,9991				
R <sup>2</sup>	0,9982				

TABELA ANOVA -  $C_{22}H_{46}$ 

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	254.872.891.627.578,00000	1	254.872.891.627.578,00000	10.983	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 2.444$
Resíduo Sqres	371.314.194.888,785000	16	23.207.137.180,549000		<b>Modelo confiável para previsões.</b>
Total SQT	255.244.205.822.467,00000	17	-	-	-
Falta de ajuste SQFaj	100.220.201.584,117000	4	25.055.050.396,029400	1,11	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,34$
Erro Puro SQEP	271.093.993.304,667000	12	22.591.166.108,722200		<b>Não apresenta falta de ajuste.</b>
r	0,9998				
R <sup>2</sup>	0,9996				

TABELA ANOVA -  $C_{24}H_{50}$ 

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	174.156.837.677.138,00000	1	174.156.837.677.138,00000	11.950	$F_{Tab}: 4,49$ $F_{Cal}/F_{Tab}: 2.659$
Resíduo Sqres	233.178.294.955,691000	16	14.573.643.434,730700		<b>Modelo confiável para previsões.</b>
Total SQT	174.390.015.972.094,00000	17	-	-	-
Falta de ajuste SQFaj	97.270.759.496,357500	4	24.317.689.874,089400	2,15	$F_{Tab}: 3,26$ $F_{Cal}/F_{Tab}: 0,66$
Erro Puro SQEP	135.907.535.459,333000	12	11.325.627.954,944400		<b>Não apresenta falta de ajuste.</b>
r	0,9997				
R <sup>2</sup>	0,9994				

TABELA ANOVA - C<sub>26</sub>H<sub>54</sub>

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	156.775.867.379.666,00000	1	156.775.867.379.666,00000	12,271	F <sub>Tab</sub> : 4,49 F <sub>Cal</sub> /F <sub>Tab</sub> : 2.731
Resíduo Sqres	204.411.285.340,566000	16	12.775.705.333,785400		<b>Modelo confiável para previsões.</b>
Total SQT	156.980.278.665.007,00000	17	-	-	-
Falta de ajuste SQFaj	92.473.942.198,566000	4	23.118.485.549,641500	2,48	F <sub>Tab</sub> : 3,26 F <sub>Cal</sub> /F <sub>Tab</sub> : 0,76
Erro Puro SQEP	111.937.343.142,000000	12	9.328.111.928,500000		<b>Não apresenta falta de ajuste.</b>
r	0,9995				
R <sup>2</sup>	0,9991				

TABELA ANOVA - C<sub>28</sub>H<sub>58</sub>

	Soma quadrática	Graus de liberdade	Média quadrática	Teste F	F tabelado (95%)
Regressão SQR	120.416.253.144.964,00000	1	120.416.253.144.964,00000	6,385	F <sub>Tab</sub> : 4,49 F <sub>Cal</sub> /F <sub>Tab</sub> : 1.421
Resíduo Sqres	301.738.630.640,535000	16	18.858.664.415,033400		<b>Modelo confiável para previsões.</b>
Total SQT	120.717.991.775.604,00000	17	-	-	-
Falta de ajuste SQFaj	146.049.634.357,201000	4	36.512.408.589,300300	2,81	F <sub>Tab</sub> : 3,26 F <sub>Cal</sub> /F <sub>Tab</sub> : 0,86
Erro Puro SQEP	155.688.996.283,333000	12	12.974.083.023,611100		<b>Não apresenta falta de ajuste.</b>
r	0,9994				
R <sup>2</sup>	0,9988				

## **ANEXO A – CÁLCULOS DE REPETIBILIDADE E REPRODUTIBILIDADE DOS PARÂMETROS ESTUDADOS DE ACORDO COM OS GUIAS NORMATIVOS**

### **Parâmetro massa específica – Norma ABNT NBR 14065**

Repetibilidade = 0,0001

Reprodutibilidade = 0,0005

### **Parâmetro ponto de fulgor – Norma ASTM D93**

Repetibilidade =  $0,029 * X$ , onde X é a média dos valores do conjunto

Reprodutibilidade =  $0,071 * X$ , onde X é a média dos valores do conjunto

### **Parâmetro temperatura de destilação – Norma ABNT NBR 9619**

#### **T10**

Repetibilidade =  $0,0094 * T$ , onde T é a média dos valores do conjunto

Reprodutibilidade =  $0,022 * T$ , onde T é a média dos valores do conjunto

#### **T50**

Repetibilidade = 0,94

Reprodutibilidade = 2,97

#### **T90**

Repetibilidade =  $0,0041 * T$ , onde T é a média dos valores do conjunto

Reprodutibilidade =  $0,015 * T$ , onde T é a média dos valores do conjunto

### **Parâmetro teor de biodiesel – Norma EN 14078**

Repetibilidade =  $0,016 * X - 0,0195$ , onde X é a média dos valores do conjunto

Reprodutibilidade =  $0,0793 * X - 0,413$ , onde X é a média dos valores do conjunto

### **Parâmetro teor de enxofre**

#### **- S10 – Norma ASTM D5453**

Repetibilidade = 1,0

Reprodutibilidade = 3,3

#### **- S500 – Norma ABNT NBR 14533**

Repetibilidade =  $0,4347 * (X^{0,6446})$ , onde X é a média dos valores do conjunto

Reprodutibilidade =  $1,9182 * (X^{0,6446})$ , onde X é a média dos valores do conjunto