

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

**Determinação do ponto de fulgor e pressão de vapor Reid
em petróleo por HTGC e DHA associados à Quimiometria**
Determination of flash point and Reid vapor pressure in petroleum by HTGC
and DHA associated to chemometrics

Márcia Helena Cassago Nascimento

Dissertação de Mestrado em Química

**Vitória
2018**

Márcia Helena Cassago Nascimento

Determinação do ponto de fulgor e pressão de vapor Reid em petróleo por HTGC e DHA associados à Quimiometria

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

Área de Concentração: Química

Linha de Pesquisa:

Química do Petróleo e Biocombustíveis.

Orientador: Prof. Dr. Paulo Roberto Filgueiras

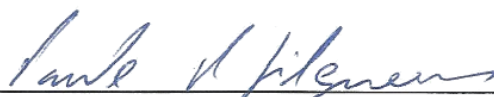
**Vitória
2018**

"Determinação do ponto de fulgor e pressão de vapor Reid em petróleo por HTGC e DHA associados à Quimiometria."

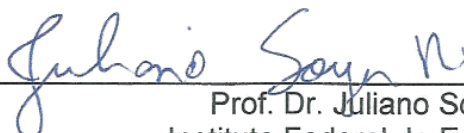
Márcia Helena Cassago Nascimento

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

Aprovada em 16/03/2018 por:



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



Prof. Dr. Juliano Souza Ribeiro
Instituto Federal do Espírito Santo
Examinador Externo



Prof. Dr. Wanderson Romão
Instituto Federal do Espírito Santo/UFES
Examinador Interno

Universidade Federal do Espírito Santo
Vitória, março de 2018

Dedico este trabalho ao meu filho, Luiz Fellype.

AGRADECIMENTOS

Agradeço primeiramente a Deus por tudo!

Agradeço à minha família pelo apoio e pela compreensão nos muitos momentos de ausência. Especialmente à minha mãe, Maria, que tanto me apoiou e ao meu filho Luiz Fellype, que além de ser minha fonte de inspiração para continuar tentando a cada obstáculo, ainda segurou minhas lágrimas e não me permitiu desistir, quando era eu quem deveria estar cuidando-o e amparando-o.

Ao meu orientador Prof. Dr. Paulo Roberto Filgueiras, pela honrosa dedicação, pelo apoio, e pelos imprescindíveis ensinamentos.

Aos colegas do Laboratório de Quimiometria, em especial à Karla Rainha e Betina Pires, as quais o companheirismo e o apoio foram determinantes para a conclusão deste trabalho.

Ao coordenador do Programa de Pós-Graduação em Química e a todos os professores pela oportunidade de ingressar e de conviver durante este período. Em especial à Prof.^a Dr^a Geisamanda Athayde por todo o carinho e confiança com que me acolheu.

Aos professores membros das bancas de Qualificação e de Defesa da Dissertação, pelas valiosas considerações que muito enriqueceram este trabalho.

Ao Centro de Pesquisa em Petróleo, Petrobras, pelo fornecimento das amostras e ensaios necessários para a realização deste trabalho.

Ao LabPetro-UFES, pela estrutura técnica para conclusão de diferentes etapas deste trabalho.

Por último, mas não menos importante, agradeço aos meus alunos, que me incentivam e me impulsionam a querer sempre ser uma pessoa melhor.

“A persistência é o caminho do êxito.”

Charles Chaplin

LISTA DE FIGURAS

FIGURA 1.1	Representação de linha de tempo de principais eventos iniciais no desenvolvimento da cromatografia gasosa.	20
FIGURA 3.1	Organograma de distribuição da composição do petróleo.....	31
FIGURA 3.2	Procedimentos empregados para destilação PEV, com representações de curvas PEV de amostras reais de três diferentes tipos de óleo (leve, médio e pesado).....	35
FIGURA 3.3	Faixa de temperatura de ebulição de derivados de petróleo	36
FIGURA 3.4	Temperatura de cortes e rendimentos de frações básicas de refino sobre curva PEV.	37
FIGURA 3.5	Desenho esquemático de copo teste para medição de ponto de fulgor pelo método Abel (<i>Abel closed-cup method</i>): (A) nível; (B) chama para teste; (C) agitador; (D) termômetro.	40
FIGURA 3.6	Equipamento para ensaio de Ponto de fulgor.....	40
FIGURA 3.7	Copo teste para ensaio de ponto de fulgor. (A) Visualização de termômetro, agitador e dispositivo de ignição chama-teste; (B) Posicionamento instrumentos em copo teste; (C) Abertura de inserção de chama-teste; (D) Copo teste acoplado ao equipamento para ensaio de ponto de fulgor.....	40
FIGURA 3.8	Desenho esquemático de um aparelho para ensaio de pressão de vapor Reid: (A) manômetro; (B) câmara de vapor; (C) Câmara de líquido.....	42
FIGURA 3.9	Esquema de cromatógrafo a gás: (A) Fonte de gás de arraste; (B) controlador de vazão e regulador de pressão; (C) sistema de injeção de amostra; (D) forno; (E) coluna cromatográfica; (F) sistema de detecção; (G) sistema de registro e tratamento dos dados.	43
FIGURA 3.10	Representação esquemática de sistema de detecção e resposta em cromatograma após registro de dados.....	44
FIGURA 3.11	Desenho esquemático do funcionamento de um cromatógrafo a gás para método de DHA com reversão de fluxo, colunas em série e detector FID.	45
FIGURA 3.12	Desenho esquemático do funcionamento de um cromatógrafo a gás para método de cromatografia gasosa de alta temperatura (HTGC) com detector FID e geração de curva de destilação simulada SIMDIS.....	47
FIGURA 3.13	Representação esquemática de organização de dados espectrais em matriz.	49
FIGURA 3.14	Representação gráfica de procedimento de calibração multivariada em análise cromatográfica de petróleo.	50
FIGURA 3.15	Demonstração gráfica do princípio da PCA.	52

FIGURA 3.16 Exemplo gráfico de projeção de <i>loadings</i> e escores na PCA de uma componente principal para duas variáveis. (A) peso (<i>loadings</i>) são os cosenos dos ângulos do vetor de direção; (B) escores são as projeções das amostras (pontos) na direção da componente principal (PC).	53
FIGURA 3.17 Demonstração gráfica do princípio do PLS.	56
FIGURA 3.18 Gráfico demonstrativo de RMSECV em função de número de variáveis latentes na etapa de otimização de modelo de regressão por PLS, método de validação cruzada k-fold.....	59
FIGURA 4.1 Fluxograma de elaboração dos conjuntos de dados instrumentais a partir dos cromatogramas DHA e HTGC para construção dos modelos pelo PLS.	74
FIGURA 4.2 Diagrama de etapas para construção de modelos de calibração multivariada.....	77
FIGURA 4.3 Aspecto visual de curvas de destilação simulada DHA e SIMDIS para identificação de ponto de junção, correções e mesclagem de dados	78
FIGURA 4.4 Aspecto gráfico de componentes para mesclagem de curvas DHA e SIMDIS.....	79
FIGURA 5.1 Cromatogramas de três amostras de petróleo pelas técnicas DHA e HTGC. (A) Cromatograma DHA, óleo leve; (B) Cromatograma HTGC, óleo leve; (C) Cromatograma DHA, óleo médio; (D) Cromatograma HTGC, óleo médio; (E) Cromatograma DHA, óleo pesado; (F) Cromatograma HTGC, óleo pesado.	80
FIGURA 5.2 Efeito visual do alinhamento dos cromatogramas HTGC pelo Icoshift. (A) Antes do alinhamento; (B) Após alinhamento.	82
FIGURA 5.3 Histogramas com distribuição da propriedade ponto de fulgor medida pelo método padrão. (A) 66 amostras analisadas pelo DHA; (B) 87 amostras analisadas pelo HTGC.....	82
FIGURA 5.4 Gráficos de correlação entre valores de ponto de fulgor (°C) de referência versus previsto pelo modelo PLS. (A), Modelo 1, DHA ₁ ; (B) Modelo 2, HTGC ₁ ; (C) Modelo 3, DHA ₂ ; (D) Modelo 4, HTGC ₂ ; (E) Modelo 5, FUSÃO ₁	84
FIGURA 5.5 Representação do cromatograma pelo DHA, das variáveis latentes (VL), em função dos <i>loadings</i> do modelo 3 – DHA ₂ na previsão de ponto de fulgor.....	85
FIGURA 5.6 Representação do cromatograma pelo HTGC, das variáveis latentes (VL), em função dos <i>loadings</i> do modelo 4 – HTGC ₂ na previsão de ponto de fulgor.	86
FIGURA 5.7 Resultados avaliação teste significância estatística Modelos preditivos PF. (A) Modelo 1 - DHA ₁ ; (B) Modelo 2 - HTGC ₁ ; (C) Modelo 3 - DHA ₂ ; (D) Modelo 4 - HTGC ₂ ; (E) Modelo 5 - FUSÃO ₁	87

FIGURA 5.8 Histogramas com distribuição da propriedade pressão de vapor Reid medida pelo método padrão. (A) 58 amostras analisadas por DHA; (B) 101 amostras analisadas pelo HTGC.	88
FIGURA 5.9 Gráficos de correlação entre valores de pressão de vapor Reid (kPa) de referência versus previsto pelo modelo PLS. (A), Modelo 1, DHA ₁ ; (B) Modelo 2, HTGC ₁ ; (C) Modelo 3, DHA ₂ ; (D) Modelo 4, HTGC ₂ ; (E) Modelo 5, FUSÃO ₁	90
FIGURA 5.10 Representação do cromatograma pelo HTGC, das variáveis latentes (VL), em função dos <i>loadings</i> do modelo 2 - HTGC ₁ na previsão de pressão de vapor Reid.....	92
FIGURA 5.11 Resultados avaliação teste significância estatística Modelos preditivos PVR. (A) Modelo 1 - DHA ₁ ; (B) Modelo 2 - HTGC ₁ ; (C) Modelo 3 - DHA ₂ ; (D) Modelo 4 - HTGC ₂ ; (E) Modelo 5 - FUSÃO ₁	93
FIGURA 5.12 Efeito visual da correção das curvas para interpolação e fusão. (A) Curva DHA e SIMDIS, sem correção; (B) Curvas DHA, SIMDIS e SIMDIS' (com correção); (C) Curva final obtida DHA/SIMDIS'.....	94
FIGURA 5.13 Curva simulada desenvolvida pela fusão de dados de DHA com SIMDIS para um óleo médio.	95
FIGURA 5.14 Comparação entre as curvas de amostra de petróleo médio para as técnicas de DHA, SIMDIS e fusão dos dados DHA/SIMDIS. (A) Curva DHA; (B) Curva SIMDIS; (C) Curva fusão DHA/SIMDIS.	95
FIGURA 5.15 Gráfico de comparação entre erros relativos das curvas PEV, SIMDIS1 e SIMDIS2 de amostra de óleo pesado (API 17,7; PF 57,0 °C; PVR 10,1 kPa).....	96

LISTA DE TABELAS

TABELA 3.1. Composição básica de petróleo típico.	32
TABELA 3.2. Classificações do petróleo de acordo com o °API.	34
TABELA 3.3 Determinações vigentes da ANP para Frações de Petróleo de acordo com curva PEV, grau API e pontos de corte	37
TABELA 5.1 Principais parâmetros estatísticos dos modelos PLS para predição de ponto de fulgor.....	83
TABELA 5.2 Principais parâmetros dos modelos PLS para previsão de pressão de vapor Reid.	89

LISTA DE ABREVIATURAS E SIGLAS

ANN	Redes Neurais Artificiais, do inglês <i>Artificial Neural Networks</i>
ANP	Agência Nacional de Petróleo, Gás natural e Biocombustíveis.
API	Instituto Americano de Petróleo, do inglês <i>American Petroleum Institute</i> .
ASTM	Sociedade americana de testes de materiais, do inglês <i>American Society for Testing and Materials</i> .
ATR-FTIR	Espectroscopia no infravermelho por transformada de Fourier com reflectância total atenuada, do inglês <i>Fourier Transform Infrared – Attenuated Total Reflectance Spectroscopy</i> .
BSW	Parâmetro de presença de Água e sedimentos, do inglês <i>Basic Sediments and Water</i> .
CENPES	Centro de Pesquisa e Desenvolvimento da Petrobras.
CNTP	Condições Normais de Temperatura e Pressão.
COW	Método de correlação otimizada, do inglês <i>correlation optimized warping</i> .
CV	Validação cruzada, do inglês <i>cross-validation</i> .
DHA	Análise detalhada de hidrocarbonetos, do inglês <i>Detailed Hydrocarbons Analysis</i>
DSC	Calorimetria diferencial de varredura, do inglês <i>Differential Scanning Calorimetry</i> .
FFT	Transformada de Fourier rápida, (do inglês, <i>Fast Fourier Trasnformation</i> .
FID	Detector por ionização em chama, do inglês <i>Flame Ionization Detector</i> .
FT-ICR-MS	Ressonância de íon ciclotron com transformada de Fourier, do inglês <i>Fourier transform ions ciclotron ressonanse-Mass Spectroscopy</i> .
FT-IR	Espectroscopia no infravermelho por Transformada de Fourier, do inglês <i>Fourier Transform Infrared</i> .
GC	Cromatografia gasosa.
GLP	Gás liquefeito de petróleo
HCA	Análise hierárquica de agrupamentos, do inglês <i>Hierarchical Cluster Analysis</i>
HTGC	Cromatografia gasosa de alta temperatura, do inglês <i>High Temperature Gas Chromatography</i>
IEA	Agência Internacional de Energia, do inglês <i>International energy Agency</i> .

ISO	- Organização INternacional de Padronização, do inglês <i>International Organization for Standardization</i>
kPa	Unidade de pressão Pascal.
MSC	Correção multiplicativa de sinal, do inglês <i>Multiplicative Scatter Correction</i> .
NAT	Número de acidez total.
°C	Graus Celsius.
°F	Graus fahrenheit
PCA	Análise por componentes principais, do inglês <i>Principal Component Analysis</i>
PCR	Regressão por componentes principais, do inglês <i>Principal Component Regression</i>
PEV	Ponto de Ebulição Verdadeiro
PF	Ponto de fulgor
PFE	Ponto Final de ebulição.
PIONA	Método cromatográfico de separação em classes Parafinas, Isoparafinas, Olefinas, Naftenos e Aromáticos.
PLS	Regressão por Mínimos Quadrados Parciais, do inglês <i>Partial Least Squares</i>
PVR	Pressão de vapor Reid
QSPR	Estudos quantitativos da relação estrutura-propriedade, do inglês <i>Study of Quantitative Structure-property Relationship</i> .
R ²	Coeficiente de determinação.
RMN	Ressonância magnética nuclear.
RMSEC	raiz quadrada do erro quadrático médio do conjunto de calibração, do inglês <i>Root Mean Square Error of Calibration</i> .
RMSECV	raiz quadrada do erro quadrático médio de validação cruzada, do inglês <i>root mean square error of cross validation</i> .
RMSEP	Raiiz quadrada do erro quadrático médio do conjunto de previsão, do inglês <i>Root Mean Square Error of Prediction</i> .
SARA	Método cromatográfico de separação em classes: Saturados, Aromáticos, Resinas e Asfaltenos.
SIMCA	Modelagem independente flexível por analogia de classe, do inglês <i>Soft Independent Modeling of Class Analogy</i>

SIMDIS	Destilação simulada por cromatografia de alta temperatura, do inglês <i>Simulated Distillation</i> .
SNV	Padronização normal de sinal, do inglês, <i>Standard Normal Variate</i> .
SVR	Regressão por vetor de suporte, do inglês <i>Support Vector Regression</i> .
TIAC	Temperatura inicial de aparecimento de cristais.
TJ	Temperatura de junção.
VL	Variável latente.

RESUMO

Ponto de fulgor (PF) e pressão de vapor Reid (PVR) são importantes propriedades físico-químicas do petróleo, utilizadas como parâmetros de segurança e meio ambiente. Os ensaios para medição de PF e de PVR são realizados em petróleo bruto e algumas frações, fornecendo parâmetros de qualidade do óleo e de derivados. Propôs-se neste trabalho, uma metodologia inovadora para determinação de PF e PVR em óleo bruto, utilizando cromatografia gasosa de alta temperatura (HTGC) e análise detalhada de hidrocarbonetos (DHA), associados com regressão pelos mínimos quadrados parciais (PLS). A HTGC, geralmente, é utilizada para a técnica de destilação simulada do petróleo. Porém, existe uma limitação nesta técnica na parte inicial da curva, *light-end*, equivalente à destilação até *n*-C10. A técnica DHA fornece eluição dos componentes mais leves do petróleo, portanto, as informações dessas técnicas têm caráter complementar. Para a realização deste trabalho, utilizou-se 128 amostras de óleo bruto (com densidade API entre 17,0 a 54,0) oriundas da bacia sedimentar brasileira. Para determinação das propriedades físico-químicas, PF e PVR, de petróleo bruto, por meio dessas técnicas cromatográficas, realizou-se a associação desses dados à calibração multivariada pelo PLS. Aplicou-se uma fusão de dados de alto nível, aos dados cromatográficos das técnicas DHA e HTGC na determinação de PF e PVR. Para melhor descrever a curva PEV, aplicou-se a composição de uma fusão de médio nível entre os dados das curvas simuladas, pelas duas técnicas, para gerar uma curva PEV completa pela destilação simulada. Devido à relação direta com os constituintes voláteis no petróleo, foi possível prever com eficiência essas duas propriedades a partir dos dados cromatográficos das técnicas HTGC e DHA. O principal resultado foi obtido por meio dos cromatogramas de HTGC na determinação da pressão de vapor Reid, com erro médio (RMSEP) de 0,4 kPa e coeficiente de determinação (R^2_p) de 0,99. Com a fusão de informações das duas técnicas cromatográficas foi possível estimar o ponto de fulgor com erro médio de 5,3 °C e coeficiente de determinação de 0,89. A fusão de dados representa ganho de informações na estimativa de ponto de fulgor, entretanto, para estimar a pressão de vapor Reid, os cromatogramas pela HTGC são suficientes. Os modelos desenvolvidos foram avaliados estatisticamente quanto à presença de erros nos resíduos. Teste F foi aplicado para avaliar a melhora na exatidão pela fusão de dados, enquanto que a significância estatística foi avaliada pelo teste de permutação dos valores da propriedade de interesse. Os modelos foram ainda avaliados pela significância química. Os resultados apontam para modelos quimicamente e estatisticamente significativos. Além disso, a fusão dos dados forneceu importante melhora na curva PEV simulada pelo HTGC, com maior abrangência das informações de hidrocarbonetos leves. O fornecimento da curva PEV com maiores informações e a determinação das propriedades por meio dos dados cromatográficos com quimiometria, foram concluídos de forma satisfatória.

Palavras-chave: Óleo bruto; Ponto de fulgor; Pressão de vapor Reid; HTGC; DHA; PLS.

ABSTRACT

Flash Point (FP) and Reid vapor pressure (RVP) are important physico-chemical properties of petroleum, used as safety parameters and environment. The essay for the measurement of FP and RVP are carried out in crude oil and in some fractions, providing the quality parameters of oil and derivatives. It was proposed in this work, an innovative methodology for determining the FP and RVP in crude oil, using the high temperature gas chromatography (HTGC) and detailed hydrocarbon analysis (DHA), associated with regression by the minimum partial squares (PLS). The HTGC is usually used for the simulated distillation of oil. However, there is a limitation on this technique in the initial part of the curve, light-end, equivalent to distillation until nC_{10} . The DHA technique provides elution of the lighter oil components, therefore, the information of these techniques is of the complementary character. For this work, 128 samples of crude oil (with API density between 17.0 to 54.0) were used, from the Brazilian sedimentary basin. For the determination of the physico-chemical properties, FP and RVP, of crude oil, by means of these chromatographic techniques, was held the association of these data with the multivariate calibration, by PLS. A high-level data fusion was applied to the data chromatographic of the techniques DHA and HTGC in the determination of FP and RVP. To better describe the TBP curve, the composition of a medium-level data fusion was applied between the simulated curves, by the two techniques, to generate a complete TBP curve by simulated distillation. Due to the direct relationship with the volatile constituents in petroleum, it was possible to effectively predict these two properties from the chromatographic data of the HTGC and DHA techniques. The main result was obtained by means of the chromatograms of HTGC in determining the Reid vapor pressure, with average error (RMSEP) of 0.4 kPa and coefficient of determination (R^2_p) of 0.99. With the fusion of information of the two chromatographic techniques it was possible to estimate the flash point with average error of 5.3 °C and coefficient of determination of 0.89. Data fusion represents information gain in the flash point estimation, however, to estimate the Reid vapor pressure, the chromatograms by HTGC are sufficient. The models developed were statistically evaluated as to the presence of errors in the residues. Test F was applied to evaluate the improvement in accuracy by the data fusion, while the statistical significance was evaluated by the permutation test of the property values of reference. The models were also evaluated by the chemical significance. The Results they point to chemically and statistically significant models. In addition, a data merger provided significant improvement in the TBP curve simulated by the HTGC, with greater scope of information of the light hydrocarbons. The provision of the TBP curve with more information and the determination of the properties by means of the chromatographic data with chemometrics, have been satisfactorily completed.

Keyword: Crude oil; Flash point; Reid vapor pressure; HTGC; DHA; PLS.

SUMÁRIO

1	INTRODUÇÃO	17
2	OBJETIVOS	29
2.1	Objetivo Geral.....	29
2.2	Objetivos específicos.....	29
3	FUNDAMENTAÇÃO TEÓRICA	30
3.1	Petróleo	30
3.2	Cromatografia gasosa	42
3.2.1	Análise detalhada de hidrocarbonetos (DHA)	44
3.2.2	Destilação simulada por cromatografia gasosa de alta temperatura (SIMDIS, ou HTGC)	46
3.3	Quimiometria	47
3.3.1	Calibração multivariada	49
3.3.2	Análise de Componentes Principais (PCA)	51
3.3.3	Regressão por componentes principais (PCR)	53
3.3.4	Regressão por mínimos quadrados parciais (PLS)	54
3.3.5	Pré-tratamento e pré-processamento dos dados	60
3.3.5.1	Correção de deslocamentos ao longo da linha de base	61
3.3.5.2	Correção de linha de base	62
3.3.5.3	Derivada	62
3.3.5.4	Normalização	62
3.3.5.5	Correção multiplicativa de espalhamento (MSC).....	63
3.3.5.6	Padronização normal de sinal (SNV).....	63
3.3.5.7	Centragem na média.....	64
3.3.5.8	Autoescalamento	64
3.3.6	Avaliação dos modelos.....	65
3.3.6.1	Avaliação de erros sistemáticos.	67
3.3.6.2	Avaliação de erros de tendência pelo teste não paramétrico	68
3.3.7	Fusão de dados.....	69
4	METODOLOGIA	71
4.1	Análise por cromatografia gasosa.....	71
4.2	Medições das propriedades físico-químicas	72
4.2.1	Ponto de fulgor	73

4.2.2	Pressão de vapor Reid	73
4.3	Modelagem dos dados	73
4.3.1	Determinação das propriedades físico-químicas	74
4.3.1.1	Ponto de fulgor.....	75
4.3.1.2	Pressão de vapor Reid.....	75
4.3.1.3	Avaliação dos modelos PLS.....	75
4.4	Fusão de dados para obtenção de curva de destilação simulada	77
5	RESULTADOS E DISCUSSÕES	80
5.1	Perfil cromatográfico.....	80
5.2	Modelagem dos dados	81
5.2.1	Estimativa de ponto de fulgor	82
5.2.2	Estimativa pressão de vapor Reid	88
5.3	Fusão de dados para estimativa da curva PEV	93
6	CONCLUSÕES	97
7	REFERÊNCIAS	98
ANEXOS		108

1 INTRODUÇÃO

O petróleo apresenta uma ampla variação em sua composição química e propriedades físico-químicas.¹⁻³ Essas propriedades influenciam na exploração, produção, refino e qualidade do produto final obtido. O conhecimento de sua composição e do comportamento desses constituintes durante o processamento do petróleo é de suma importância para a indústria. A partir dos parâmetros analíticos de avaliação das propriedades físico-químicas podem-se estimar os componentes predominantes no óleo, as melhores condições para escoamento e processamento, além de produtos a serem obtidos pelo refino.

Um dos processos de maior importância na determinação do perfil de composição química do petróleo é a destilação. A partir da destilação do óleo, são obtidos os subprodutos, denominados frações, que possuem maior valor econômico em relação ao óleo bruto.^{3,4,20} A destilação fornece a curva por Ponto de Ebulição Verdadeiro (PEV). Essa curva é um dos parâmetros mais importantes de valoração do petróleo.⁴ A partir do petróleo bruto e das frações obtidas pela destilação são realizadas as aferições das propriedades físico-químicas, em uma série de análises para caracterização do petróleo.³⁻⁵ Para a realização da caracterização completa do óleo bruto e/ou de suas frações, os ensaios analíticos demandam tempo e quantidades de amostras que nem sempre vão ao encontro da necessidade de otimização da indústria petrolífera. O desenvolvimento de metodologias de análises rápidas, robustas e com menores recursos tornou-se um desafio importante neste cenário.¹⁻⁵

Com o interesse da indústria petrolífera pela aplicação da cromatografia gasosa na obtenção da composição do petróleo, por meios eficientes, rápidos e com menor volume de amostra, o método de destilação simulada pela cromatografia gasosa de alta temperatura (HTGC, do inglês *High Temperature Gas Chromatography*) foi desenvolvido.²¹ Este método passou a ser utilizado para obtenção de uma curva PEV com menor volume de amostra e menor tempo de análise.²¹⁻²⁴ Porém, na etapa de amostragem, os componentes mais leves da mistura podem ser perdidos e a parte inicial da curva PEV simulada pode não refletir o ensaio da curva PEV, com maior volume de amostra. Em contrapartida, a técnica de cromatografia gasosa para análise detalhada de hidrocarbonetos (DHA, do inglês *Detailed Hydrocarbons Analysis*) realiza ensaio com alta resolução em temperatura mais branda e promove eluição de

componentes mais voláteis, até *n*-C14.^{7-8,25-26} Este método é geralmente utilizado para avaliação de composição de frações voláteis como a nafta e gasolina, pela técnica de análise detalhada de hidrocarbonetos. A união de informações das técnicas cromatográficas que contemplem todas as frações e forneçam a destilação simulada do óleo bruto de forma completa, pode ser complementar e melhor descrever a curva PEV.²⁴⁻²⁵

Atualmente, análises que fornecem informações sobre a composição de frações e produtos de petróleo são possíveis pelas diferentes técnicas instrumentais que incluem a ressonância magnética nuclear (RMN), espectrometria de massa (MS) e espectroscopia de infravermelho (IR) e várias técnicas cromatográficas. A cromatografia gasosa fornece o perfil de eluição com informações sobre a composição do óleo e de suas frações de forma rápida. Dessa forma, entre as técnicas instrumentais, a cromatografia gasosa tornou-se uma das técnicas analíticas de maior importância na indústria do petróleo. A extração de informações a partir do perfil de eluição de um petróleo pode ser aplicada para inferir sobre o comportamento deste óleo e sua correlação com diferentes propriedades físico-químicas. Principalmente aquelas relacionadas com os fatores que influenciam a eluição, como, por exemplo, temperatura de ebulição e pressão de vapor, ou seja, relacionadas à volatilidade, como as propriedades de interesse neste trabalho: o ponto de fulgor e a pressão de vapor.²⁴⁻²⁹

Entre as propriedades físico-químicas do petróleo, o ponto de fulgor e a pressão de vapor estão relacionados com a capacidade de emissão evaporativa e com riscos de incêndio, durante transporte e processamento do óleo e de seus derivados. São propriedades importantes para caracterizar o petróleo a respeito de questões de segurança e meio ambiente. Também são importantes sob o aspecto de escoamento do óleo e de derivados, estando relacionadas às condições de ocorrência de cavitação de bombas. O ensaio ponto de fulgor (PF) é aplicado aos derivados médios e pesados, querosene de aviação, óleo diesel, óleos combustíveis, asfalto e óleos lubrificantes. A determinação de pressão de vapor é fornecida pelo ensaio de pressão de vapor Reid (PVR). O ensaio PVR é aplicado para petróleo e frações mais voláteis como gasolina e gás liquefeito de petróleo (GLP).⁶ A estimativa dessas propriedades por meio de técnicas instrumentais tem sido estudada e proposta em frações de petróleo, como óleo diesel e gasolina.^{6, 9-19} Porém, para petróleo bruto ainda são poucos os trabalhos com metodologias para estimativas dessas propriedades.

A associação de técnicas cromatográficas, bem como a extração dos dados químicos pelos cromatogramas e suas relações com as propriedades físico-químicas do petróleo, são possíveis por meio de técnicas desenvolvidas e consolidadas pela quimiometria. Esta área da química permite o processamento de informações químicas por meio de ferramentas estatísticas e matemáticas.²⁷⁻³⁵

A quimiometria surgiu com a modernização dos equipamentos de análise instrumental, mas ficou consolidada com a popularização dos microcomputadores.³² Com o aumento na quantidade de informações obtidas pelas análises instrumentais e processamento de dados, muitas ferramentas quimiométricas foram aplicadas aos tratamentos de dados químicos, dentre as quais, podemos citar, análise de componentes principais (PCA, do inglês *Principal Component Analysis*), análise hierárquica de agrupamentos (HCA, do inglês *Hierarchical Cluster Analysis*), modelagem independente flexível por analogia de classe (SIMCA, do inglês *Soft Independent Modeling of Class Analogy*), redes neurais artificiais (ANN, do inglês *Artificial Neural Networks*), regressão por componentes principais (PCR, do inglês *Principal Component Regression*) e regressão pelos mínimos quadrados parciais (PLS, do inglês *Partial Least Squares*), entre outros.^{32,34-35}

Entretanto, durante um longo tempo as ferramentas quimiométricas eram utilizadas em cromatografia apenas visando melhorar a resolução dos métodos cromatográficos.²⁷ Com a consolidação da calibração multivariada com métodos de regressão, principalmente o método de regressão pelos mínimos quadrados parciais (PLS), a associação da quimiometria com técnicas instrumentais analíticas, tais como infravermelho com transformada de *Fourier* (FT-IR), RMN e a cromatografia, vem sendo amplamente divulgada para determinação de propriedades físico-químicas de petróleo e de suas frações.^{5,7-11,13-19,45-46,48-53}

A associação da quimiometria à cromatografia gasosa amplia a importância analítica do método, uma vez que associa os dados de composição do petróleo para inferir sobre suas propriedades físico-químicas, permitindo a caracterização do petróleo com redução de tempo e de recursos de análise. Daszykowski e Walczak (2006)²⁷ discutiram as diversas aplicações, vantagens e limitações na relação da quimiometria com cromatografia. Sendra-Bosque *et al.* (2012)²⁸ descreveram vários métodos de aplicação de diferentes técnicas quimiométricas em dados cromatográficos.

1.1 Cromatografia gasosa na indústria do petróleo

A cromatografia gasosa (GC) teve um papel essencial na análise de petróleo nos últimos 60 anos.^{26,28-29} Pela necessidade de técnicas eficientes de separação e identificação dos componentes químicos do petróleo, projetos desenvolvidos pelo Instituto Americano de Petróleo (API, do inglês *American Petroleum Institute*) no período entre 1928 a 1930, indicavam o interesse pela cromatografia gasosa na indústria de petróleo.^{26,29,36} Neste período técnicas de extração em fase sólida e cromatografia em coluna por adsorção eram testadas para separação de petróleo. Em 1956, comprovou-se a utilização da cromatografia gasosa para a indústria de petróleo, no 1º Simpósio Internacional sobre Cromatografia em fase de vapor (gás), oficializando-se o primeiro cromatógrafo a gás comercial.^{21-22,26,29}

O desenvolvimento inicial e a consolidação da cromatografia gasosa ocorreram entre os anos de 1941 a 1958, numa série de eventos, conforme descrito na Figura 1.1. A cromatografia gasosa teve o despertar da técnica em torno de 1952, e de seus equipamentos a partir de 1958 com a introdução das colunas capilares e detectores por ionização em chama (FID, do inglês, *Flame Ionization Detector*) o que ampliou a seletividade e a capacidade analítica do método.^{26,29,36-38}

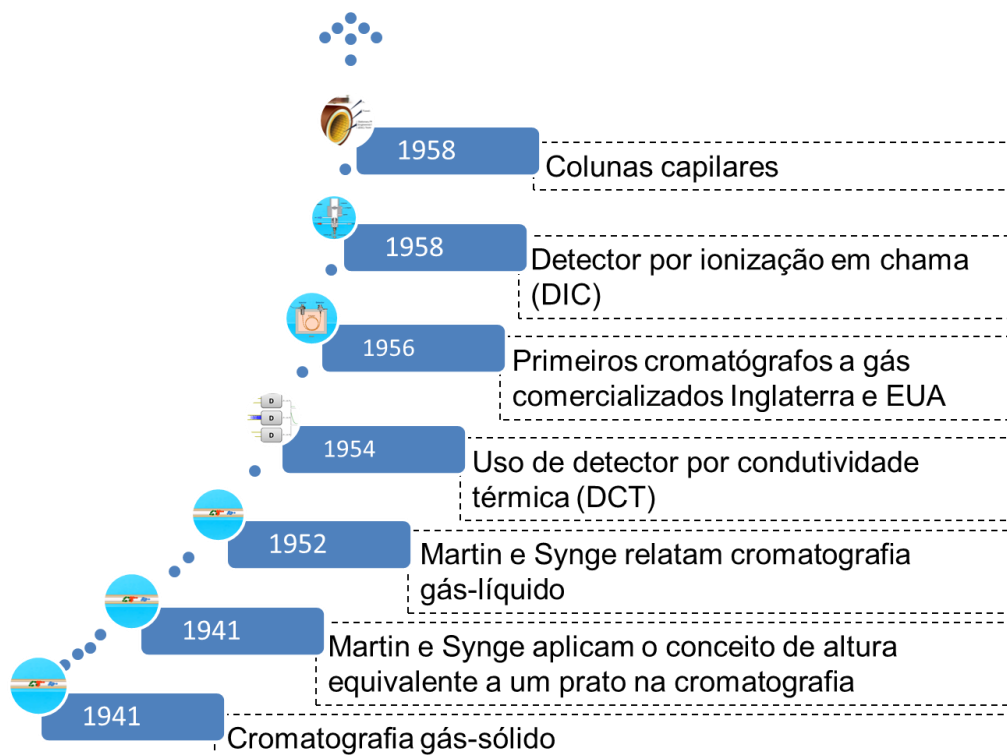


Figura 1.1 Representação de linha de tempo de principais eventos iniciais no desenvolvimento da cromatografia gasosa.³⁶
Fonte: Autor.

Schneider e Bruderreck (1964)³⁸ relataram anomalias entre tempo de retenção e temperaturas de ebulição em uma mesma fase estacionária nas cromatografias utilizadas para análise de petróleo. Os autores avaliaram a separação de hidrocarbonetos de C1 a C8 com cromatografia gasosa, utilizando coluna capilar de vidro com suporte de grafite adsorvido de *n*-decano ($C_{10}H_{22}$) comparando-se com adsorção por esqualano ($C_{30}H_{62}$). Relataram maior eficiência de separação, maior resolução, maior simetria dos picos e redução no efeito de cauda (efeito causado pela interação desigual de um mesmo substrato com a fase estacionária), com suporte de grafite adsorvido com esqualano.

Com o desenvolvimento de sistemas de válvulas de refluxo e diferentes fases estacionárias, o método de cromatografia gasosa foi testado pela primeira vez para a nafta de petróleo por Pesek e Blair (1979)³⁹. Os autores utilizaram uma metodologia padronizada pela sociedade americana de testes de materiais (ASTM, do inglês *American Society for Testing and Materials*) para análise de gasolina de aviação com pontos de ebulição abaixo de 177 °C, para determinação de aromáticos e alquilbenzenos de cadeia *n*-C8. Dessa forma surgiu o método de cromatografia gasosa de alta resolução, denominado análise detalhada de hidrocarbonetos (DHA). Este método é aplicado para frações voláteis do petróleo, como a nafta, misturas de matéria-prima para gasolina, gasolina e análise que determina o teor de parafinas, isoparafinas, olefinas, naftenos e aromáticos (PIONA) em produtos derivados de petróleo.^{4,25-26}

Em contrapartida, com o desenvolvimento das colunas capilares de materiais mais resistentes a altas temperaturas, como colunas metálicas e de sílica fundida, também na década de 1960, foi possível desenvolver a cromatografia gasosa de alta temperatura.²¹⁻²⁴ Esta técnica proporcionou uma relação direta da separação de componentes pela cromatografia gasosa com a separação pela destilação do petróleo. Com isso a cromatografia gasosa passou a fornecer um ensaio de destilação simulada por cromatografia gasosa de alta temperatura para petróleo bruto.²¹⁻²² Com o desenvolvimento deste método a análise de petróleo passou a alcançar informações confiáveis com maiores temperaturas de ebulição e menor tempo de análise.²¹⁻²⁴

Com o despertar desta técnica como procedimento laboratorial de análise de petróleo, a destilação por pontos de ebulição verdadeiros passou a ser chamada de destilação física, enquanto que o método de destilação pela cromatografia gasosa de alta temperatura passou a ser denominada de destilação simulada. A ASTM passou

a padronizar métodos para destilação simulada pela cromatografia gasosa a partir de 1974.²⁴

A curva PEV por meio da destilação física é obtida pelos métodos ASTM D2892⁴⁰ e D5236⁴¹. Entretanto, estes métodos demandam dias para o ensaio e torno de 16 litros de volume de amostra.⁴⁰⁻⁴¹ A destilação simulada apresenta algumas vantagens em relação à destilação física: menor tempo de análise - em torno de 47 minutos cada análise; menor volume de amostra - 1 mL a cada réplica; maior reprodutibilidade devido à maior automação do método e maior informação da composição da amostra, devido à ampliação do alcance de temperatura de ebulição. Dessa forma, a técnica tem sido amplamente utilizada como análise laboratorial de destilação simulada na indústria de petróleo.^{23-24,42-46}

Espinosa, Figueroa e Jiménez (2004)⁴² compararam o método de destilação simulada com o método de destilação física pela primeira vez para óleos pesados mexicanos. Os autores encontraram concordância entre os métodos, mas atribuíram à destilação simulada um maior grau de confiança e maior potencialidade para óleos crus pesados. Austrich, Buenetro e Galeana (2015)⁴³ compararam dois métodos de destilação simulada, ASTM D5307 e ASTM D7169 citando e correlacionando seus resultados, afirmando estarem de acordo aos de Espinosa, Figueroa e Jiménez (2004)⁴².

Roussis e Fitzgerald (2000)⁴⁴ utilizaram o método de destilação simulada com cromatografia gasosa acoplada a espectrômetro de massas para determinação de curva de pontos de ebulição e recuperação de eluição em porcentagem de massas e de volume a partir de amostras referência de óleo cru. Compararam os resultados obtidos em cinco amostras de óleo cru com a curva obtida pela destilação física, obtendo média máxima de diferença no intervalo de 295°C-510°C para o cálculo da massa e de 319°C-510°C para o cálculo do volume. Os autores atribuíram as discrepâncias aos ajustes matemáticos na destilação física para a combinação dos métodos de destilação atmosférica e a vácuo, e dessa forma consideraram a destilação simulada com maior grau de confiança por não utilizar essa correção, sendo realizada em ensaio único. Obtiveram coeficientes de correlação (R^2) entre a destilação simulada e a destilação física de 0,9993 na determinação de recuperação em massa e de 0,9986 para a conversão em volume.

Coto *et al.* (2011)⁴⁵ utilizaram o método de destilação simulada por cromatografia gasosa de alta temperatura em petróleo para determinação de

distribuição de *n*-parafinas comparando com métodos padronizados. Os autores propuseram metodologia de melhoria na integração de picos de cromatogramas de HTGC. Avaliaram a influência da distribuição de parafinas nos modelos termodinâmicos de estudo de precipitação de parafinas. Espada et al. (2013)⁴⁶ utilizaram metodologia proposta por Coto et al (2011)⁴⁵ para obtenção de curva de distribuição de *n*-parafinas em frações (105°C-535°C) de petróleos parafinicos (leves), para o qual compararam as técnicas de calorimetria diferencial de varredura (DSC, do inglês *Differential Scanning Calorimetry*) e curva de destilação simulada pela cromatografia gasosa. Eles encontraram resultados semelhantes entre as técnicas para cortes leves e médios, porém discrepantes para cortes pesados, sendo atribuídas as discrepâncias às limitações de DSC para esses cortes.

Grecsek, Osenbach e Tipler (2010)²⁵ utilizaram os dados da técnica DHA com o objetivo de corrigir o efeito de interferência do solvente e limitações do início da curva obtida pela destilação simulada pelo HTGC. Os autores utilizaram a cromatografia gasosa de petróleo bruto com um sistema de comutação de fluxo e uma pré-coluna, de sílica fundida - 60 cm, entre o comutador de fluxo e a coluna analítica, de 100 m. Os dados das curvas das duas técnicas foram mesclados em um software próprio para correção e mesclagem destas técnicas. Como resultado reportaram mais precisão nas caracterizações e uma curva de pontos ebulição completa, abrangendo de *n*-C1 a *n*-C100.

Nielsen et al. (2012)³³ propuseram um método de classificação de petróleo com utilização de cromatografia gasosa e técnicas quimiométricas. Os autores realizaram testes com amostras de petróleo bruto de diferentes tipos, amostras de derramamento de petróleo e simulações de intemperismo. O objetivo do trabalho foi desenvolver um método aplicável aos derramamentos reais de óleo, classificar a origem de um óleo mesmo após ocorrer intemperismo. Com a aplicação de análises por componentes principais (PCA) identificaram amostras de óleos biodegradáveis. Os resultados foram satisfatórios, para identificação de variáveis afetadas pelo intemperismo e aplicação dos modelos em amostras reais de derramamento para validação do método.

Apesar de se obter informação quantitativa a partir de um cromatograma, a identificação de componentes de uma mistura baseado apenas no tempo de retenção pode não ser suficiente para garantir a acurácia do procedimento. O desenvolvimento de materiais para colunas capilares e de fases estacionárias de menor espessura proporcionaram outros avanços da cromatografia como, o despertar das técnicas

chamadas “hifenizadas”. Estas técnicas compreendem o acoplamento de técnicas cromatográficas a diversos espectrômetros, permitindo a identificação de uma substância pura. Entre estas técnicas, a cromatografia gasosa acoplada à espectroscopia de massas (GC/MS) ganhou importantes aplicações em diversas áreas da química analítica assim como na análise composicional do petróleo.^{26,36,44}

Outro desafio na análise composicional de misturas complexas como petróleo, é a coeluição de componentes, principalmente pelos possíveis isômeros de cada constituinte. Por essa razão, as técnicas de cromatografia gasosa multidimensional foram introduzidas como uma forma de efetuar uma melhor separação de misturas complexas. Essas técnicas promovem a separação em diferentes dimensões, com colunas capilares em paralelo, de diferentes tamanhos e polaridades. Entre tais técnicas encontram-se a cromatografia gasosa bidimensional (GC-GC), a cromatografia gasosa bidimensional abrangente (GCXGC) ou bidimensional abrangente com detector de ionização em chama (GCXGC-FID), ou ainda, a bidimensional abrangente com detector de massas (GCXGC/MS).²⁶⁻³¹ Tendo em vista que estas técnicas não estão no escopo deste trabalho, não serão descritas com maiores detalhes no decorrer deste, podendo ser encontradas maiores informações em Beens e Brinkman (2000)²⁹; Mühlen, Zini e Caramão (2007)³⁰ e Nizio, McGinitie e Harynuk (2012)³¹.

Para o presente trabalho, foram avaliadas as técnicas de HTGC e de DHA, para que, associadas às ferramentas quimiométricas, possam fornecer as relações de seus dados instrumentais na caracterização de algumas propriedades físico-químicas do petróleo.

1.2 Caracterização de propriedades físico-químicas do petróleo

Desde o início dos processos de refino do petróleo, eram realizadas análises para caracterizar os produtos. As análises geralmente compreendiam medições físicas, sendo a faixa de ebulição a mais comum, seguida pelas medições de propriedades como densidade, índice de refração, viscosidade, ponto de fulgor e ponto de fluidez, além das análises composicionais que eram vistas como maior desafio para a indústria.²⁹ Atualmente, o petróleo bruto e suas frações são caracterizados utilizando-se técnicas analíticas e dados laboratoriais, com os quais são calculados parâmetros para determinar suas propriedades físico-químicas. As análises pelos métodos instrumentais geralmente estão associadas às ferramentas

da quimiometria, principalmente calibração multivariada, tendo como valores de referência as medições padronizadas.

Tozzi *et al.* (2015)⁴⁸ utilizaram métodos-padrão para medições de propriedades físico-químicas: acidez, medida como número de acidez total (NAT); teor de enxofre, ponto de fluidez, viscosidade cinemática e gravidade API. Essas propriedades inferem sobre as condições de processamento e de transporte do óleo. Petróleos com alta acidez, associando-se a outros fatores e interações, causam problemas sérios de corrosão de equipamentos. Petróleos com alta viscosidade apresentam dificuldades na produção, transporte e processamento. Com o objetivo de avaliar a melhoria de propriedades físico-químicas os autores compararam os parâmetros medidos de quatro amostras de óleo entre médio e pesado, realizaram misturas de diferentes petróleos e compararam as propriedades físico-químicas das misturas e dos óleos originais. Os autores propuseram ainda modelo preditivo das propriedades físico-químicas de misturas como função da propriedade no óleo original pelas suas respectivas composições.

Morales-Medina e Guzmán (2012)⁴⁹ utilizaram destilação simulada pela cromatografia gasosa de alta temperatura (seguindo a norma ASTM D5307) para estimar propriedades físico-químicas: densidade API e viscosidade cinemática de petróleo bruto colombiano, utilizando análise de componentes principais (PCA). Em seguida os autores utilizaram a regressão por componentes principais (PCR) e regressão pelos mínimos quadrados parciais (PLS). O trabalho discutiu a respeito da influência dos compostos voláteis, presentes nos cinco primeiros minutos dos cromatogramas, obtendo diferentes resultados para algumas das propriedades ao suprimir esses dados dos cromatogramas.

Filgueiras *et al.* (2014)⁵ determinaram propriedades físico-químicas: densidade API, viscosidade cinemática e teor de água, em petróleos utilizando espectroscopia no infravermelho médio com transformada de Fourier (ATR-FTIR) associada à calibração multivariada pelos métodos de regressão de vetor de suporte (SVR, do inglês *Support Vector Regression*) e regressão PLS. Terra *et al.* (2015)⁵⁰ utilizaram ressonância de íon cíclotron com transformada de Fourier (FT-ICR-MS) associado à técnica PLS para determinar nitrogênio básico e teor de aromáticos em petróleo. Duarte *et al.* (2016)⁵¹ utilizaram técnica de ressonância magnética nuclear de ¹H e regressão PLS para determinação das propriedades físico-químicas (densidade API,

resíduo de carbono, temperatura de aparecimento de cristais de parafinas, e teor de nitrogênio total) em óleo bruto de petróleo brasileiro.

Flumignan *et al.* (2007)⁸ descreveram o primeiro trabalho com análise pela cromatografia gasosa e aplicação de métodos quimiométricos para identificação de adulterantes na gasolina comercial. Utilizaram técnica quimiométrica de análise exploratória de dados, análise hierárquica de cluster (HCA), para selecionar entre 2400 amostras coletadas de gasolina comercial o conjunto de amostras para análise pela cromatografia gasosa. As amostras analisadas foram classificadas com ferramenta quimiométrica de modelagem independente de analogia de classe (SIMCA) em duas classes: de acordo ou não de acordo com a regulamentação vigente, obtendo-se 95% de classificação correta para conjunto de teste. Concluíram com a indicação do método SIMCA para análise de qualidade de gasolina e detecção de adulterante.

Entre as propriedades físico-químicas do petróleo, algumas são relacionadas diretamente à volatilidade, possuem peculiaridades que dificultam suas medições pelas técnicas espectroscópicas e, com isso, tornam escassas as publicações de suas determinações em amostras de petróleo bruto. Dentre essas propriedades, destacam-se o ponto de fulgor e a pressão e vapor Reid que estão diretamente relacionadas à volatilidade do óleo e são importantes parâmetros de segurança e meio ambiente. Essas propriedades fornecem informações sobre o nível de emissões evaporativas e risco de incêndio durante transporte e processamento do óleo ou do produto derivado do petróleo.^{2-3,9-13}

Butler *et al.* (1956)⁹ publicaram o primeiro trabalho com previsão de ponto de fulgor por meio de métodos matemáticos em frações de petróleo. Os autores utilizaram a correlação de informações experimentais como peso molecular, ponto de ebulição e pressão de vapor para a previsão de ponto de fulgor. A partir deste período foram desenvolvidos diversos métodos matemáticos teóricos, ou seja, sem realização de medidas experimentais da propriedade de interesse, para previsão do ponto de fulgor, em hidrocarbonetos puros, misturas de hidrocarbonetos e de frações de petróleo.⁹⁻¹¹

Estes métodos teóricos, geralmente correspondem às correlações empíricas, estudos quantitativos da relação estrutura-propriedade (QSPR, do inglês *Study of Quantitative Structure-property Relationship*) e métodos de contribuição de grupos. Entretanto, os métodos de QSPR e de contribuição de grupos não são aplicáveis para o petróleo, uma vez que sua aplicação utiliza informações exatas de composição e

estrutura da molécula. Dessa forma somente são aplicáveis em substâncias puras ou misturas de composição conhecida.¹¹

Montin (1998)¹² propôs substituição de equipamentos de medição de pontos de fulgor “*on line*”, instalados nas unidades de destilação, de refinaria de uma instituição na Suécia, por método teórico sistematizado nas unidades de controle de torres de destilação. Pelo método proposto, o ponto de fulgor foi estimado pela relação entre viscosidade, pontos de corte e rendimento, obtidos pela unidade de destilação. O modelo teórico era calibrado/validado, frequentemente, por meio de coleta de amostra e medição da propriedade predita pelos ensaios de ponto de fulgor e cromatografia gasosa utilizando correlação da propriedade com a região de componentes voláteis, correspondente ao teor de 5% de volume recuperado na destilação.

DAI *et al.* (2015)¹⁰ utilizaram modelos de contribuição de grupos para determinar pontos de fulgor de dados de uma amostragem de 50 hidrocarbonetos de composição conhecida, com técnicas quimiométricas de regressão linear, obtendo desvio médio de 2,5% em relação aos dados de pontos de fulgor obtidos em outros métodos teóricos e empíricos coletados em base de dados.

Alqaheem e Riazi (2017)¹¹ avaliaram diversos métodos teóricos existentes para previsão de ponto de fulgor em hidrocarbonetos puros, misturas e frações de petróleo. Os autores utilizaram dados coletados do instituto API de 64 hidrocarbonetos de C5 a C22, como ponto de fulgor, peso molecular, temperatura de ebulição, densidade específica e pressão de vapor. Também utilizaram dados de 73 amostras de frações de petróleo como, querosene, diesel, nafta, gasóleo e resíduo de destilação a vácuo. Os métodos existentes foram avaliados quanto à aplicação nos hidrocarbonetos puros e nas diferentes frações, que compõem misturas de maior complexidade, obtendo-se desvios médios de 1,6 a 4,9% entre dados experimentais coletados e calculados pelos métodos disponíveis para hidrocarbonetos puros e de 3,7% a 11,1% para frações de petróleo. Os autores propuseram método teórico de predição de ponto de fulgor por meio de relações deste parâmetro com ponto de ebulição, com desvio médio de 1,7% para hidrocarbonetos puros e 2,8% para frações de petróleo.

Outros trabalhos para previsão de ponto de fulgor e de pressão e vapor Reid foram realizados com derivados de petróleo como diesel e gasolina, utilizando-se medições analíticas e associação de métodos instrumentais com algumas técnicas de regressão multivariada.¹³⁻¹⁹ Alves, Henriques e Poppi (2012)¹⁵ utilizaram análise de óleo diesel por infravermelho próximo (NIR) associado ao método de regressão não

linear SVR e regressão linear pelo PLS na estimativa de ponto de fulgor. Côco, Yamamoto e Meien (2005)¹⁸ determinaram a pressão de vapor Reid, densidade API e curva de destilação, de amostras de gasolina comercial brasileira, pelas correlações entre a composição química, obtida pela análise com cromatografia gasosa DHA, associada ao método de regressão de redes neurais artificiais (ANN), alcançando erro médio de 1%. Ré-Poppi *et al.* (2009)⁶ compararam a determinação de pressão de vapor Reid e outras propriedades físico-químicas da gasolina comercial pelos métodos padronizados pela Agência Nacional de Petróleo e Biocombustíveis (ANP) à determinação pela cromatografia gasosa com ferramentas quimiométricas. Alcançaram resultados com maior discrepância apenas no teor de olefinas. Mendes, Aleme e Barbeira (2017)¹⁹ aplicaram dados de curva de destilação física, associados à regressão por mínimos quadrados parciais (PLS), para determinação de pressão de vapor Reid de oitenta amostras de gasolina automotiva dentro do intervalo de 49,1 kPa a 67,4 kPa, com coeficiente de determinação entre valores de referência e estimados pela regressão, R^2 de 0,9743.

Propriedades relacionadas à volatilidade, como ponto de fulgor e pressão de vapor Reid em petróleo, podem apresentar limitações em técnicas que requeiram preparos de amostras ou medições em condições que impliquem perda de componentes mais leves. A cromatografia gasosa pode ser aplicada para encontrar uma relação linear com essas propriedades físico-químicas, pela característica de sua técnica estar intrinsicamente relacionada aos parâmetros de temperatura de ebulição e pressão de vapor da mistura. Entretanto, devido à maior complexidade no comportamento de petróleo bruto, são escassos os trabalhos com estimativas das propriedades ponto de fulgor e pressão de vapor Reid nesse tipo de amostra.

2 OBJETIVOS

2.1 Objetivo Geral

Utilizar informações das técnicas cromatográficas: destilação simulada por cromatografia gasosa de alta temperatura (HTGC) e análise detalhada de hidrocarbonetos (DHA) para estimar a curva de pontos de ebulição verdadeiros de petróleo, ponto de fulgor e pressão de vapor Reid em petróleo por calibração multivariada.

2.2 Objetivos específicos

- Determinar a curva de pontos de ebulição verdadeiros de petróleo a partir da técnica de destilação simulada por cromatografia gasosa de alta temperatura (HTGC);
- Associar informações do HTGC com DHA para estimar a curva PEV de petróleos;
- Desenvolver modelos de calibração multivariada baseados na regressão PLS para estimar as propriedades físico-químicas de petróleo bruto: ponto de fulgor e pressão de vapor Reid utilizando a associação das informações das técnicas HTGC e DHA.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 Petróleo

Do latim *petra* (pedra) e *oleum* (óleo) o petróleo no estado líquido é uma substância oleosa, inflamável, menos densa que a água e de cheiro característico e cor variando entre negro e castanho claro. Quando a mistura possui uma maior quantidade de moléculas pequenas, seu estado físico é gasoso, quando contêm moléculas maiores seu estado físico é líquido nas Condições Normais de Temperatura e Pressão (CNTP).¹ Petróleo, segundo a Agência Nacional de Petróleo, Gás natural e Biocombustíveis (ANP) é “todo e qualquer hidrocarboneto líquido em seu estado natural, ou obtido através de processamento, a exemplo do óleo cru, condensado e óleo sintético”.⁵²

O petróleo é uma mistura complexa, cuja composição básica apresenta diferentes tipos de hidrocarbonetos, como os compostos parafínicos, que são alcanos de cadeia normal e de cadeia ramificada; naftênicos, que são compostos cíclicos (cicloalcanos); aromáticos, compostos com pelo menos um anel benzeno; resinas e asfaltenos, que são cadeias policíclicas, de alta massa molar contendo ainda átomos de oxigênio, enxofre e nitrogênio; além de sais orgânicos contendo metais.¹⁻³

Os óleos obtidos em diferentes reservatórios possuem características diversas, sendo que alguns são pretos, densos, viscosos, liberando pouco ou nenhum gás, enquanto que outros são castanhos ou bastante claros, com baixa viscosidade e densidade, liberando quantidade apreciável de gás.¹ Outros reservatórios podem produzir somente gás. Entretanto, todos eles produzem análises elementares semelhantes, com maior porcentagem de hidrogênio e carbono. Dessa forma, a composição do petróleo pode ser descrita pela presença de hidrocarbonetos, constituinte principal, e não-hidrocarbonetos, sendo os mais comuns o nitrogênio, o enxofre e o oxigênio.¹ A distribuição da composição do petróleo pelos hidrocarbonetos e não-hidrocarbonetos presentes, está esquematizada na Figura 3.1.⁶

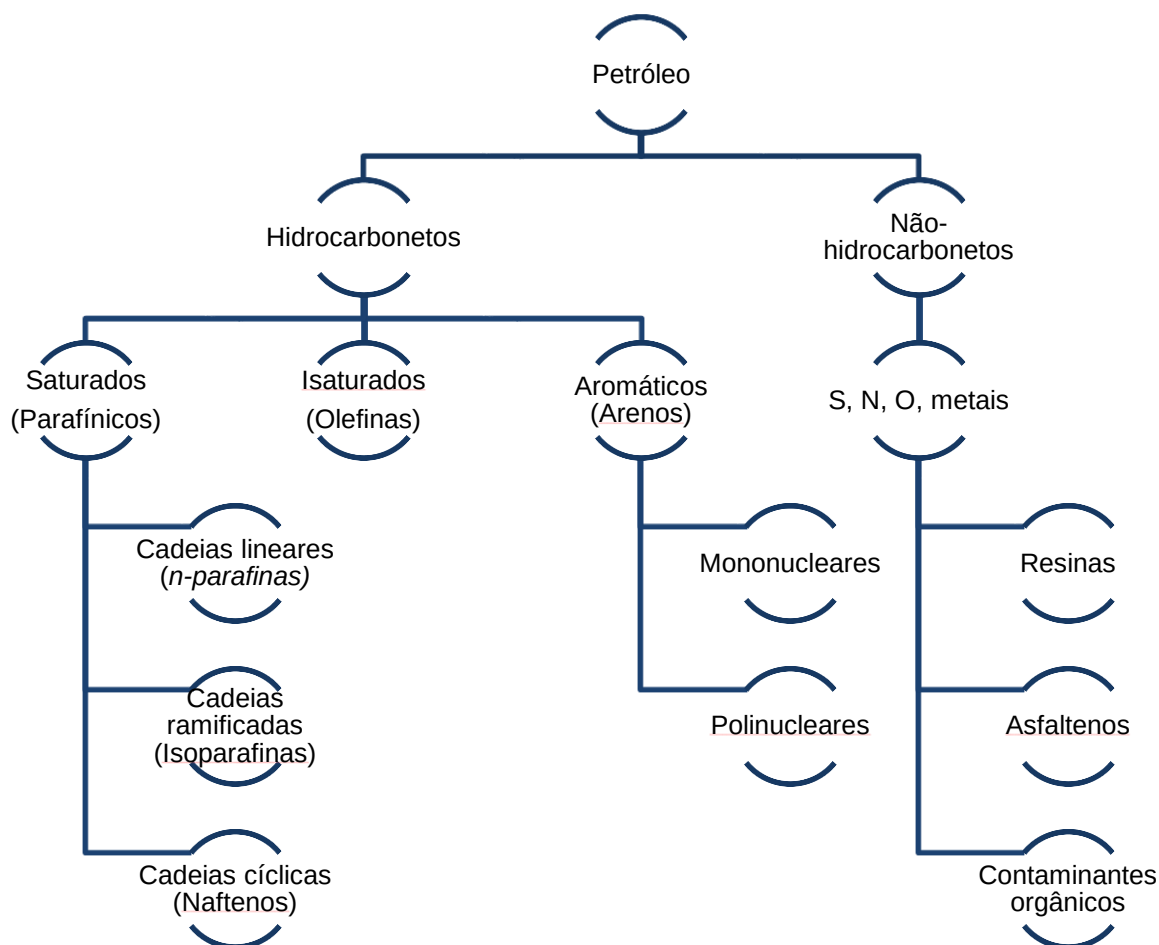


Figura 3.1. Organograma de distribuição da composição do petróleo.
Fonte: Adaptado de Farah (2012)⁶.

A qualificação do petróleo por meio de sua composição química pode classificar o petróleo de acordo com o tipo de hidrocarboneto predominante. Para os geoquímicos a classificação do óleo é importante para relacioná-lo à rocha-mãe, inferindo assim sobre sua origem e condições de produção. Para os refinadores a classificação é importante para inferir sobre o comportamento desse óleo e de suas frações durante exploração e processo de refino, assim como os possíveis produtos finais e suas características.¹

A classificação do petróleo, de acordo com a composição química, baseia-se nos teores de parafinas, naftenos, aromáticos, enxofre, asfaltenos e resinas, nas frações de temperatura de ebulição acima de 210 °C.⁶ A classificação segundo Tissot e Welte (1978)^{6,55} classifica o petróleo em: parafínicos, parafínicos-naftênicos, naftênicos, aromáticos intermediários, aromáticos-naftênicos e aromáticos asfálticos.

Uma segunda classificação, proposta por Speight (2001)² não considera os teores de asfaltenos e de resinas.

Compostos insaturados (olefinas) quase não são encontrados no petróleo bruto devido a sua alta reatividade a reações de hidrogenação, com a consequente formação de compostos saturados, que normalmente são encontrados em maior quantidade, exceto em óleos biodegradados. Compostos saturados podem compreender cerca de 60% em óleos do tipo parafínicos-naftênicos, cerca de 40-45% em óleos ditos aromáticos e de 20-25% em óleos biodegradados.^{6,55-57} Uma composição básica de um petróleo considerado típico, segundo Thomas (2004)¹ é apresentada na Tabela 3.1.

Tabela 3.1. Composição básica de petróleo típico.

<i>Composição</i>	<i>Porcentagem em massa (%m/m)</i>
Parafinas normais	14%
Parafinas ramificadas	16%
Parafinas cíclicas	30%
Aromáticos	30%
Resinas e asfaltenos	10%

Fonte: Adaptado de Thomas (2004)¹.

Nos estudos sobre a constituição do petróleo, o Instituto API realizou análises em vários petróleos de diferentes origens, concluindo-se que em todos continham os mesmos tipos de hidrocarbonetos, variando a quantidade relativa de cada grupo. Porém, de mesma ordem de grandeza a quantidade relativa de compostos individuais dentro de cada grupo de hidrocarbonetos. Por conseguinte, essas variações implicam nas características físico-químicas do petróleo.¹

Muitas propriedades dos óleos como cor, densidade, viscosidade, podem ser influenciadas pela presença dos não hidrocarbonetos. As impurezas primárias como a água, os sais, compostos nitrogenados, sulfurosos, oxigenados e metálicos, são fatores preocupantes no processo industrial, pois podem causar problemas como corrosão, contaminação de catalisadores, estabilidade de emulsões, alterações nos produtos obtidos, entre outros. De uma forma geral, os contaminantes são mais encontrados em frações mais pesadas do petróleo, são fatores determinantes nas estratégias de produção, no transporte e refino e, frequentemente, determinam os

produtos que podem ser obtidos. Dessa forma, não existe um esquema universal de refino adaptável a qualquer tipo de óleo bruto.²⁻⁴

Outro aspecto a ser considerado é que os óleos apresentam essa composição heterogênea, não somente em diferentes poços, mas também de mesmo poço a diferentes profundidades.²⁻⁴ Torna-se importante o conhecimento de sua composição química para atender às especificações da refinaria. As características do óleo que é encaminhado para o refino, são dadas por um conjunto de parâmetros obtidos por uma série de métodos analíticos que podem informar sobre o potencial de qualidade e derivados esperados, bem como indicar possíveis dificuldades de refino e transporte.^{2-4,6,55-58}

O Instituto API desenvolveu uma forma de classificar o petróleo denominada gravidade API, derivada da gravidade específica. Essa forma de classificação, também descrita como grau API, geralmente é utilizada como um fator da qualidade do petróleo e se tornou um importante critério para estabelecer os valores comerciais do petróleo bruto. Compreende uma escala hidrométrica que representa uma forma de expressar a densidade relativa de um óleo ou derivado. Essa escala API, medida em graus API, varia inversamente com a densidade relativa (Equação 3.1). Dessa forma, quanto maior o grau API, menor densidade do produto, e, normalmente, maior sua qualidade e seu valor comercial.^{1,3-4,6}

$$^{\circ}API = \frac{141,5}{\rho(60^{\circ}F/60^{\circ}F)} - 131,5 \quad (3.1)$$

Onde a letra grega ρ (*rho*) é a densidade 60 °F/60°F que é a massa específica do óleo a 60°F comparada com a da água na mesma temperatura. O valor do °API é calculado com a densidade do petróleo a 15,6°C ou 60°F, sendo classificado o petróleo de acordo com os valores de °API.

Existem diferentes classificações do petróleo de acordo com seu °API, que diferem de acordo com critérios específicos de instituições de pesquisa, exploração e regulamentadoras. Neste trabalho destacam-se duas diferentes classificações (Tabela 3.2): de acordo com Petróleo Brasileiro (Petrobras)⁶ e a Agência Internacional de Energia (IEA, do inglês *International Energy Agency*), descrita por Neto (2017)⁵⁹.

Tabela 3.2. Classificações do petróleo de acordo com o °API.

Tipo de Petróleo	Classificações pelo Grau API	
	Petrobras	IEA
Extra-leve	API > 40,0	----
Leve	40 > API ≥ 33	API > 25
Médio	33 > API ≥ 27	25 ≥ API > 20
Pesado	27 > API ≥ 19	20 ≥ API > 10
Extra-pesado	19 > API ≥ 15	API < 10
Asfáltico	API < 15,0	----

Fonte: Adaptado de Farah (2012)⁶ e adaptado de Neto (2017)⁵⁹

A avaliação para valoração do óleo bruto inicia-se com a construção da curva pelos Pontos de Ebulição Verdadeiros (PEV), por meio de técnica laboratorial de destilação especificada nas normas da associação americana de normalização de materiais (ASTM) D2892⁴⁰ - Método Padrão para Destilação de Petróleo Cru, a destilação atmosférica; e ASTM D5236⁴¹ - Método Padrão para Destilação de Misturas de Hidrocarbonetos Pesados, a destilação a vácuo. Essa curva fornece as frações evaporadas de um dado tipo de petróleo em função da temperatura.^{4,6,24}

O processo de destilação laboratorial do petróleo possui o princípio de funcionamento das torres atmosférica e a vácuo das refinarias, utilizadas na etapa de separação do óleo cru. Porém, diferem na complexidade do sistema, forma de operação e tamanho da carga a ser destilada. Com os valores de rendimento de massa e volume a cada faixa de corte do petróleo nos processos de destilação atmosférica e a vácuo, é determinada a curva PEV, que é representada por um gráfico de percentual acumulado *versus* temperatura de ebulição (Figura 3.2).

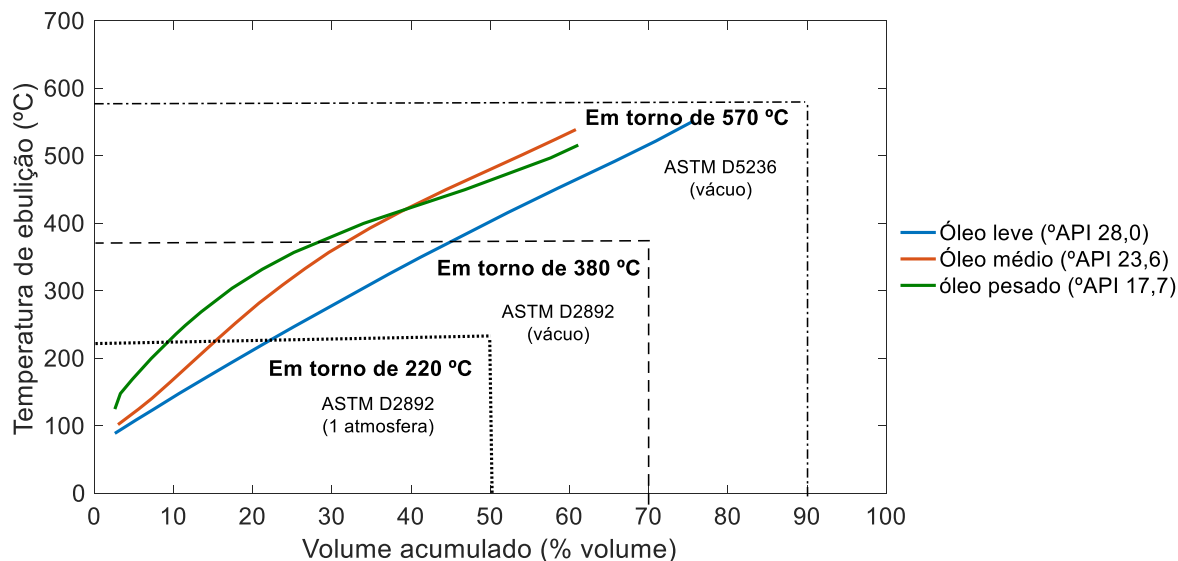


Figura 3.2 Procedimentos empregados para destilação PEV, com representações de curvas PEV de amostras reais de três diferentes tipos de óleo (leve, médio e pesado).
Fonte: elaboração própria a partir de Farah (2012)⁶ e dados de Laboratório de Núcleo de Competências em Química do Petróleo - LabPetro/UFES.

A curva PEV permite determinar o rendimento dos produtos, nortear a valoração de petróleo bruto e ainda controle de qualidade dos cortes que estão sendo destilados, servindo estas análises de parâmetros para inferir os processos que sejam necessários para enquadramento às especificações do mercado.^{4,20} Se a curva for mais inclinada horizontalmente ela terá uma quantidade maior de um determinado produto para um intervalo menor de temperatura e se ela for mais inclinada verticalmente ela terá menos produto para aquela faixa de temperatura.⁴

Com a obtenção dos percentuais das frações por cada faixa de temperatura é possível estimar o rendimento dos produtos de refino. Uma relação dos principais produtos derivados do refino de petróleo com as faixas de ebulição e tamanho da cadeia carbônica é apresentada na Figura 3.3.⁴

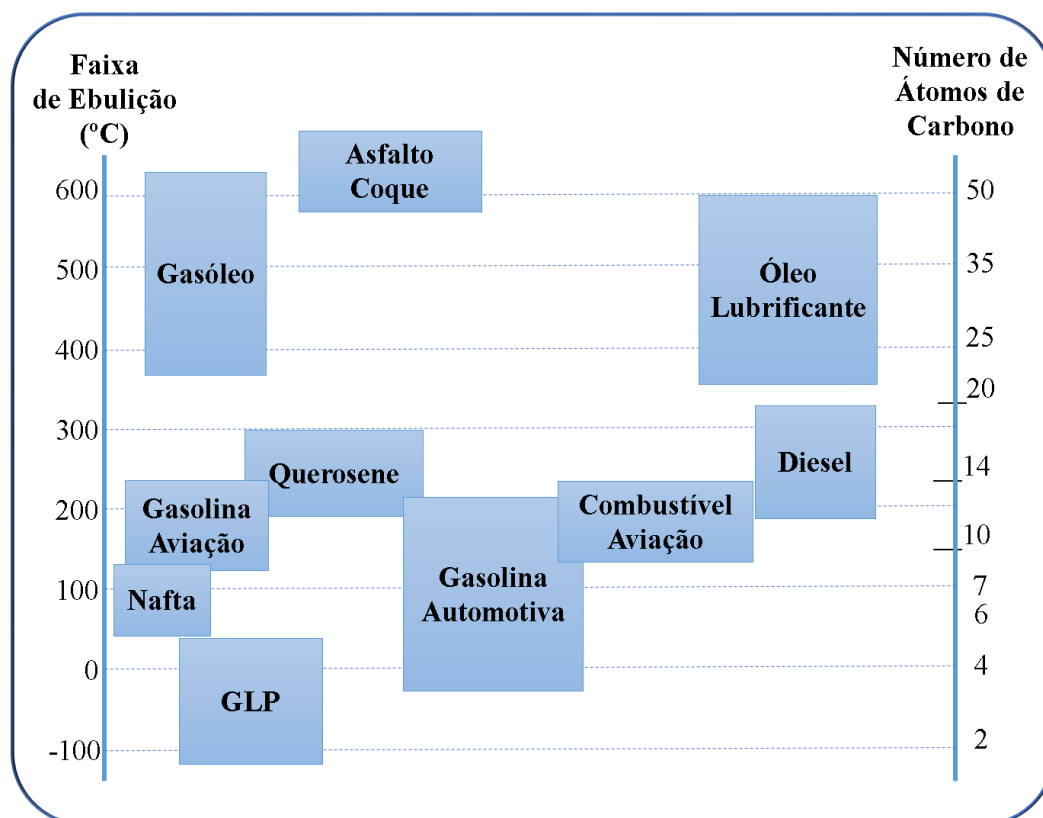


Figura 3.3 Faixa de temperatura de ebulição de derivados de petróleo
 Fonte: Adaptado de Fahim, Al-Sahhaf e Elkilani (2010)⁴.

Devido às características dos componentes as faixas de ebulição dos derivados podem ser sobrepostas. Dessa forma, a definição das temperaturas de cortes das frações básicas tende a atender a demanda de produção de derivados. Essa definição varia de acordo com a região geoeconômica e a época do ano, sendo atendidos demanda e qualidade do derivado a ser obtido. Considerando-se uma fixação de valores de temperaturas de corte de cada fração básica de refino sobre a curva PEV do petróleo, é definido o rendimento teórico dessas frações.^{3,4,6} Um exemplo de diferentes faixas de cortes de frações básicas de refino e rendimento teórico é dado na Figura 3.4.

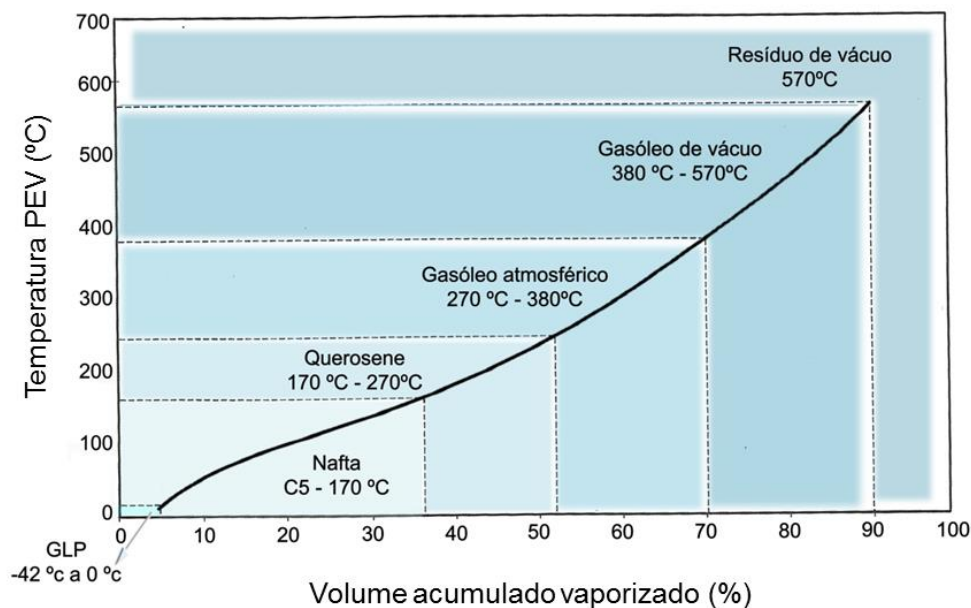


Figura 3.4 Temperatura de cortes e rendimentos de frações básicas de refino sobre curva PEV.
Fonte: adaptado de Farah (2012).⁶

A ANP (Resolução n.703/2017)⁶⁰ reformulou a forma de estipulação de valores de petróleo, de acordo com propriedades físico-químicas e os pontos de corte (temperatura de ebulição entre duas frações), da destilação pela PEV (Tabela 3.3). Até o ano 2017, os pontos de corte variavam em função do grau API (Portaria 206/2000)⁶¹. A partir de 2018 a Resolução 703/2017⁶⁰ passou a ser implementada de forma gradual, até que seja completamente vigorada em 2021, na qual o ponto de corte é estipulado uniformemente para todos os tipos de petróleo (Tabela 3.3).

Tabela 3.3 Determinações vigentes da ANP para Frações de Petróleo de acordo com curva PEV, grau API e pontos de corte

Grau API	Pontos de Corte			Regulamento ANP
	Destilados Leves	Destilados Médios	Resíduos Pesados	
< 27	Até 290°C	290°C a 380°C	Acima de 380°C	Portaria Nº 206, de 29/08/2000.
≥ 27 e < 36	Até 270°C	270°C a 450°C	Acima de 450°C	
≥ 36	Até 210°C	210°C a 500°C	Acima de 500°C	
Não se aplica	Até 180°C	180°C a 350°C	Acima de 350°C	Resolução Nº 703, de 26/09/2017.

Fonte: Adaptado de ANP. Portaria n. 206, de 29/08/2000;⁶¹ Resolução 703/2017.⁶⁰

Além do óleo bruto, as frações obtidas na destilação do petróleo também passam por diferentes ensaios físico-químicos, envolvendo ainda procedimentos de separação e análises de acordo com as especificações dos derivados que essas frações irão produzir. Algumas propriedades físico-químicas são peculiares da fase exploratória, outras da produção e de refino. A densidade relativa ($^{\circ}\text{API}$), a viscosidade e o ponto de fluidez são considerados as principais características do petróleo, por apresentarem grande importância econômica. Servem como parâmetros para dimensionamento de bombas de escoamento do óleo e de previsão de comportamento do petróleo durante todo o processo de produção.²⁻⁵

Outros parâmetros são utilizados para determinar a presença de diversas impurezas, como: água e sedimentos (BSW, do inglês *Basic Sediments and Water*), teor de sal, teor de enxofre; acidez naftênica (TAN, do inglês *total acidity number*), teor de nitrogênio.²⁻⁵ Outros são relacionados à compreensão da composição química, como: teor de Saturados, Aromáticos, Resinas e Asfaltenos (SARA); precipitação de parafinas, ou formação de cristais de parafinas medida como temperatura inicial de cristais (TIAC), entre outros. O teor de enxofre ($\%\text{m.m}^{-1}$), a acidez naftênica (TAN) em mgKOH.g^{-1} e o teor de nitrogênio ($\%\text{m.m}^{-1}$) são utilizados nos cálculos de preço de referência de petróleo pela Resolução 703/2017 da ANP.^{2-5,17,51, 60}

Entre as propriedades físico-químicas do petróleo, utilizou-se neste trabalho ponto de fulgor (PF) e a pressão de vapor Reid (PVR) que são parâmetros de qualidade, segurança e meio ambiente do petróleo ou do produto derivado de petróleo.^{3,6,9-13}

O ponto de fulgor representa a menor temperatura na qual o produto se vaporiza em quantidade suficiente para formar com o ar uma mistura capaz de se inflamar momentaneamente quando sobre ela incide uma centelha. No ponto de fulgor, a combustão não é suficientemente sustentada pelos vapores formados. O parâmetro que é utilizado para medir essa sustentação da combustão é o ponto de combustão.⁶

O ensaio de ponto de fulgor é utilizado como parâmetro de qualidade de derivados médios e pesados, como querosene de aviação, óleo diesel, óleos combustíveis, asfaltos e lubrificantes. Correlaciona-se de modo inverso ao teor de leves, uma vez que quanto maior o teor de componentes leves na mistura, maior sua pressão de vapor e a facilidade de vaporização, logo, menor é o ponto de fulgor. Geralmente, com o aumento do ponto de ebulição também aumenta o ponto de fulgor.

É utilizado como critério de segurança no transporte e manuseio dos produtos inflamáveis e combustíveis, derivados e não derivados de petróleo, pela sua medição fornecer dados de riscos de incêndio e perdas pela volatilização.^{6,11}

As normas regulamentadoras de prevenção e segurança utilizam ponto de fulgor como referência principal para caracterizar um determinado líquido como inflamável ou combustível. Como parâmetro de limites de inflamabilidade, utilizado em normas de segurança e prevenção, ponto de fulgor relaciona-se ao limite inferior de explosividade.⁶

A classificação de líquidos inflamáveis e combustíveis difere pouco entre normas regulamentadoras de países que seguem a orientação das Nações Unidas para a classificação. De acordo com a norma regulamentadora brasileira, NR 20,⁶² é classificado como inflamável produto que possua ponto de fulgor menor que 60,5 °C, e combustível, ponto de fulgor maior que 60,5 °C e menor que 93°C.⁶²

Existem diversos ensaios para medição de ponto de fulgor, divididos em duas classes de técnicas - recipiente (copo) aberto e copo fechado - que variam de acordo com a amplitude de faixa. Geralmente, a medição consiste em acondicionar a amostra previamente resfriada, mantida entre -35°C ou até 17°C abaixo do ponto de fulgor previsto, em um recipiente próprio inserido em um banho de resfriamento, seguido de agitação e aquecimento constante. Uma centelha de chama de teste é aplicada em intervalos de 1°C e 2°C, dependendo do método e do equipamento utilizado. A menor temperatura em que ocorre um primeiro flash pela aplicação da centelha, indicando uma combustão não sustentada na superfície do líquido, é registrada como ponto de fulgor em °C. O método denominado copo fechado Abel, pela ISO 13736 (2008)⁶³ é aplicada para ensaio em óleo cru e resíduos de petróleo, com amplitude de -30 °C a 70 °C, mas com precisão para a faixa entre -5 °C a 66,5 °C. Este método utiliza um recipiente fechado, denominado copo de teste, que também fica previamente em banho de arrefecimento a -35 °C. Uma representação esquemática do frasco pode ser observada na Figura 3.5. Da mesma forma, fotos do equipamento e de copo-teste, para ensaio de ponto de fulgor são apresentados em Figura 3.6 e Figura 3.7.

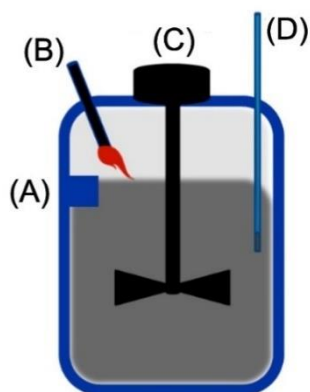


Figura 3.5 Desenho esquemático de copo teste para medição de ponto de fulgor pelo método Abel (*Abel closed-cup method*): (A) nível; (B) chama para teste; (C) agitador; (D) termômetro.
Fonte: Adaptado de Farah (2012).⁶



Figura 3.6 Equipamento para ensaio de Ponto de fulgor.
Fonte: Laboratório de Núcleo de Competências em Químicas do Petróleo – LabPetro/UFES.

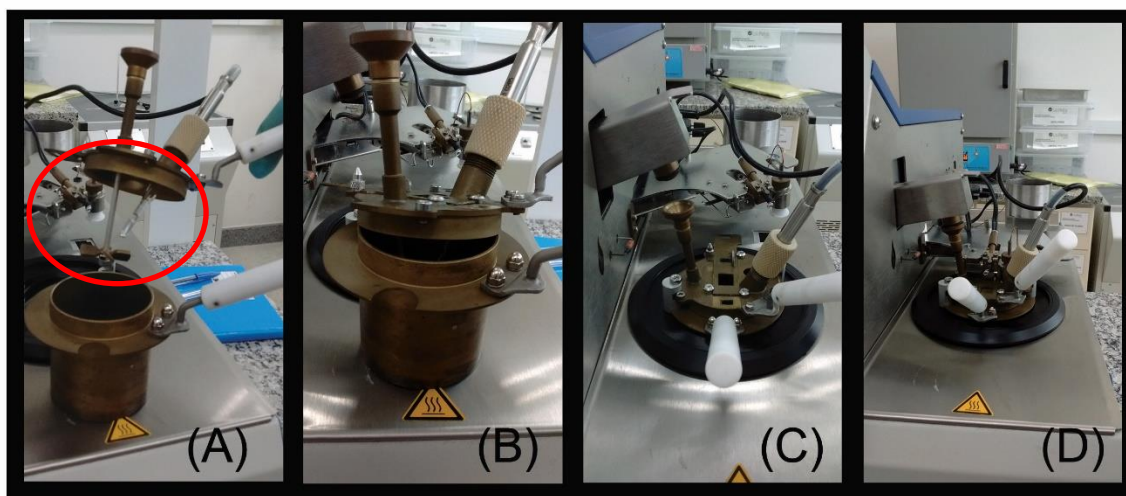


Figura 3.7 Copo teste para ensaio de ponto de fulgor. (A) Visualização de termômetro, agitador e dispositivo de ignição chama-teste; (B) Posicionamento instrumentos em copo teste; (C) Abertura de inserção de chama-teste; (D) Copo teste acoplado ao equipamento para ensaio de ponto de fulgor.
Fonte: Laboratório de Núcleo de Competências em Químicas do Petróleo – Labpetro/UFES.

A pressão de vapor é um dos mais importantes parâmetros de frações voláteis combustíveis, como o GLP, a gasolina automotiva e a gasolina de aviação. É analisada pelo ensaio de pressão de vapor Reid. Este ensaio é aplicado para GLP, gasolina e petróleo, sendo indicativo de presença de componentes leves.⁶

A ANP, pela resolução conjunta ANP/Inmetro n.1 (2013)⁶⁴, classifica como petróleo estabilizado o óleo cuja pressão de vapor seja inferior a 70 kPa, na temperatura de medição. A gasolina automotiva comercial possui vapor de pressão Reid entre 45 kPa a 69 kPa, enquanto que a gasolina de aviação apresenta pressão de vapor Reid entre 38 kPa a 49 kPa e frações como nafta e gás condensado pressão de vapor maiores que 70 kPa.⁶⁴

A vaporização ocorre em temperatura e pressão constantes, quando se trata de uma substância pura, até que toda a substância esteja vaporizada. Mas, tratando-se de uma mistura, o equilíbrio entre líquido e vapor irá sofrer influência da composição da mistura. Inicialmente, na fase vapor obtida predominam os componentes mais voláteis, mas podem estar presentes nessa fase outros componentes da mistura. Dessa forma, a temperaturas constantes, a pressão de vapor de uma mistura é dada como a pressão em que ocorre a formação da primeira bolha de vapor.⁶

Devido à dificuldade em se medir com precisão o início da vaporização de uma mistura complexa, como petróleo e derivados, é utilizada a medição em uma condição intermediária de vaporização, em que há maior estabilidade do sistema. O método de pressão de vapor Reid é definido como um ensaio que determina a pressão no interior de um cilindro apropriado, à temperatura de 37,8 °C (100 °F), em condição intermediária de vaporização do derivado, com relação líquido-vapor de $\frac{1}{4}$ para petróleo, gasolina e igual a 4 para GLP, obtendo a pressão de equilíbrio, que pode ser correlacionada com a pressão de vapor verdadeira da mistura.⁶

Para a determinação de pressão de vapor Reid de gasolina e de petróleo é utilizada a norma ASTM D323.⁶⁵ A câmara de amostra, correspondente a 20% do volume total, é cheia com a substância após ter sido passada por um resfriamento. A câmara de amostra é conectada a uma câmara de vapor. O aparelho é imerso em um banho termostático a 37,8 °C e agitado periodicamente até atingir o equilíbrio. Após medições consecutivas no manômetro a pressão de vapor Reid é dada pela pressão

em que o sistema apresenta estar em equilíbrio. Uma representação esquemática do aparelho utilizado para a medição pode ser visto na Figura 3.8.⁶

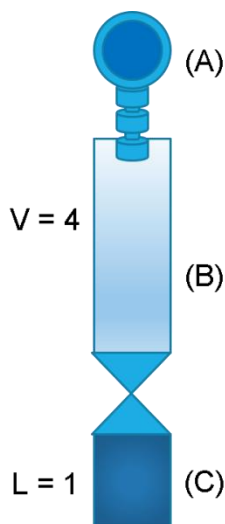


Figura 3.8 Desenho esquemático de um aparelho para ensaio de pressão de vapor Reid: (A) manômetro; (B) câmara de vapor; (C) Câmara de líquido. Fonte: adaptado de Farah (2012)⁶.

3.2 Cromatografia gasosa

A cromatografia gasosa é um método físico-químico de separação que compreende a distribuição das substâncias de uma mistura complexa entre uma fase estacionária (sólida ou líquida) e uma fase móvel (gasosa). A técnica é empregada na análise de substâncias voláteis e estáveis termicamente. O princípio da técnica cromatográfica é semelhante ao princípio físico utilizado na destilação do petróleo, ou seja, um dos fatores de separação dos componentes da mistura é pelo ponto de ebulição.^{36-37, 65-67}

A amostra é injetada por meio de um sistema de injeção na coluna contendo a fase estacionária. A temperatura adequada na injeção e na coluna causa a vaporização da amostra. De acordo com as propriedades da fase estacionária e as propriedades das substâncias contidas na amostra, essas substâncias são retidas por tempos determinados pela fase estacionária, eluem pela coluna e chegam ao final da coluna em instantes diferentes.^{37,66-67}

A técnica utilizada no desenvolvimento da análise é de eluição. Uma corrente de gás de arraste (fase móvel) passa continuamente pela coluna. A amostra vaporizada é inserida nessa corrente de gás e arrastada pela coluna. Os componentes que são retidos mais fortemente na fase estacionária movem-se mais lentamente na

fase móvel. Os componentes que interagem mais fracamente com a fase estacionária movem-se mais rapidamente. Com essa diferença de migração os componentes da amostra são separados em bandas ou zonas discretas.^{37,66-67}

Os componentes presentes na mistura, depois de separados, chegam ao detector, que gera um sinal para um sistema de registro e tratamento de dados. Uma representação esquemática de cromatógrafo a gás é apresentado na Figura 3.99.

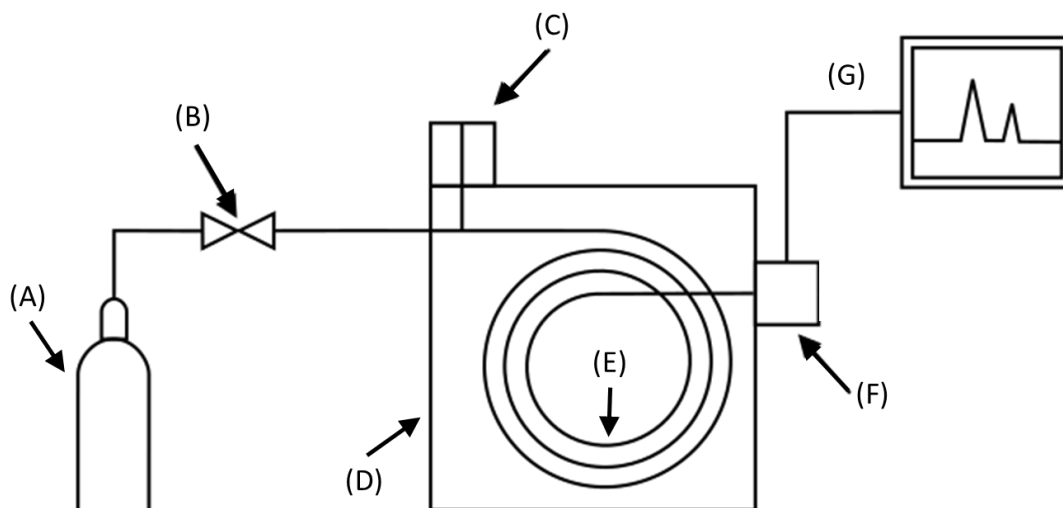


Figura 3.9 Esquema de cromatógrafo a gás: (A) Fonte de gás de arraste; (B) controlador de vazão e regulador de pressão; (C) sistema de injeção de amostra; (D) forno; (E) coluna cromatográfica; (F) sistema de detecção; (G) sistema de registro e tratamento dos dados. Fonte: adaptado de Bonato (2006)³⁷.

Como gás de arraste, alguns gases comumente utilizados são: hélio (He), nitrogênio (N₂) e argônio (Ar), entre outros. A escolha é determinada de acordo com a disponibilidade, pureza, consumo e tipo de detector empregado. Existem diversas técnicas de injeção da amostra. O sistema de injeção comumente empregado para amostras líquidas é uma microseringa com agulha hipodérmica.^{37,66-67}

Na saída da coluna utiliza-se um detector adequado para codificação de sinal e quantificação dessas substâncias. O detector responde à concentração ou massa do soluto. Seu sinal é registrado em função do tempo, obtendo-se o cromatograma com uma série de picos gaussianos (Figura 3.10). As posições dos picos no eixo do tempo podem ser usadas para identificar os componentes da amostra e a área sob os picos fornecem uma medida da quantidade desses componentes.^{37,66-67}

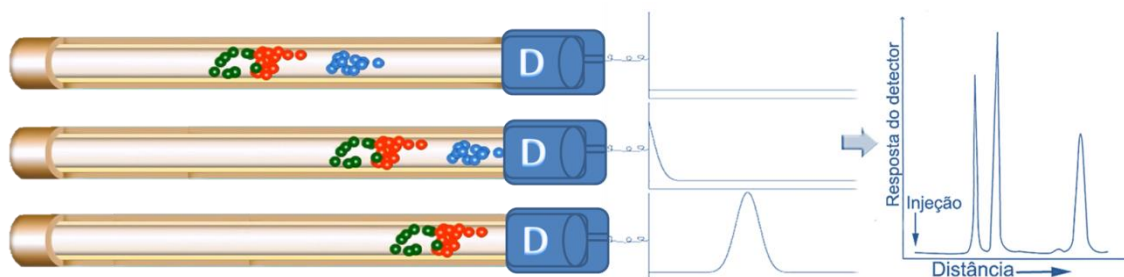


Figura 3.10 Representação esquemática de sistema de detecção e resposta em cromatograma após registro de dados.

Fonte: adaptado de Bonato (2006)³⁷

O tipo de detector empregado dependerá da natureza do analito e do nível de concentração a ser medido. Esses detectores funcionam como transdutores, transformando as moléculas que são detectadas em sinais elétricos. Os detectores podem ser universais, que geram sinal para qualquer substância eluída; seletivos, para substâncias com determinada propriedade físico-química e específicos, que geram sinal para substâncias que contenham determinado grupo ou elemento em sua composição.^{37,66-67}

A separação das diferentes moléculas de uma mistura deve ser eficientemente realizada. Isso apresentaria o que se denomina de cromatograma ideal, com os picos finos, separados e simétricos. Na realidade, em alguns casos, ocorrem sobreposições de picos e assimetrias. O tempo de retenção e o alargamento dos picos ao longo da coluna são utilizados na avaliação da eficiência da distribuição dos componentes da mistura. O alargamento do pico está diretamente relacionado à eficiência de separação, ou eficiência da coluna.^{37,66}

O alargamento é proveniente do espalhamento dos componentes por difusão no meio fluido. À medida que a amostra é carregada pelo gás de arraste ao longo da coluna, entra em sucessivos equilíbrios de distribuição entre as fases estacionária e móvel, semelhante ao que ocorre nos pratos das torres de destilação. A velocidade de migração dos componentes pela coluna é função do equilíbrio de partição e depende da passagem entre as fases.^{37,66}

3.2.1 Análise detalhada de hidrocarbonetos (DHA)

O método de análise detalhada de hidrocarbonetos (DHA) promove a eluição até *n*-C14. Atualmente, é muito utilizado em frações mais voláteis do petróleo e em derivados como gasolina e gasolina de aviação. Este método utiliza pré-fracionamento

e colunas capilares em série. É amplamente empregado na determinação de parafinas, isoparafinas, olefinas, naftenos e aromáticos (PIONA) em análises de composição de frações e derivados.^{4,25-26,68} A ANP incluiu em 2017 o método de cromatografia gasosa de alta resolução com colunas capilares de 100 metros para análise detalhada de hidrocarbonetos, pela ASTM D6729⁶⁹, para determinação de benzeno e de componentes individuais em medição de especificações técnicas da gasolina comercial.⁶⁴

Os cromatógrafos normalmente apresentam auto-amostragem, injetor, colunas de preparo, válvulas de reversão de fluxo, programador de temperatura, coluna analítica, geralmente revestida com polidimetilsiloxano (fase estacionária apolar) e detector FID (Figura 3.11).⁶⁸

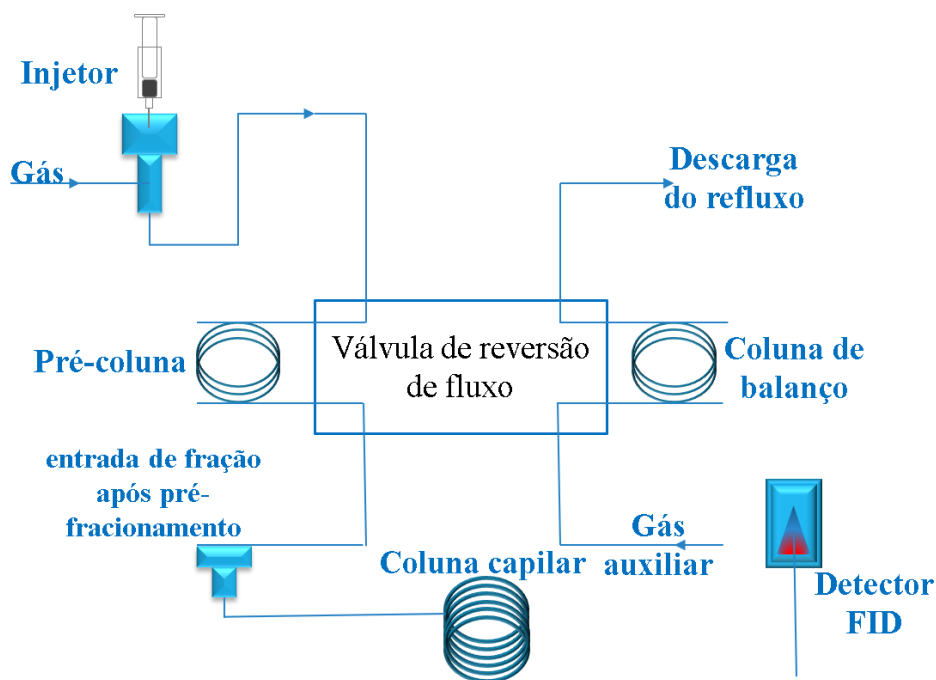


Figura 3.11 Desenho esquemático do funcionamento de um cromatógrafo a gás para método de DHA com reversão de fluxo, colunas em série e detector FID.⁶⁸
Fonte: Autor.

A pré-coluna é considerada coluna de preparo e a válvula de reversão é responsável pelo refluxo das frações acima de *n*-C14, por meio de um balanço de cargas e de pressão. A partir da calibração com misturas de *n*-parafinas são ajustados os parâmetros de temperatura, pressão e tempo de ativação da válvula de reversão de fluxo. Com a reversão do fluxo, os componentes abaixo de *n*-C14 são analisados na coluna analítica e detectados pelo detector FID.⁶⁸

A ASTM D5134-13⁷¹ padroniza método DHA para determinação de parafinas, naftenos e monoaromáticos (PNA) em nafta de petróleo, ou misturas de hidrocarbonetos líquidos, livres de olefinas, com separação e identificação de componentes até *n*-C9. A análise detalhada de hidrocarbonetos em amostras contendo olefinas é padronizada pelas ASTM D6729⁶⁹, D6730⁷¹, ou D6733⁷³. É padronizado pela norma ASTM D7900⁷⁴ para análise detalhada de óleo bruto com reversão de frações acima de *n*-C14.

O sistema de *software* acoplado ao equipamento utiliza o índice Kováts⁷⁵ e padrões externos com bibliotecas pré-definidas para identificar componentes eluídos. Esse índice é a retenção de componentes relativos aos compostos de parafinas.

3.2.2 Destilação simulada por cromatografia gasosa de alta temperatura (SIMDIS, ou HTGC)

A cromatografia gasosa de alta temperatura (HTGC), também conhecida como destilação simulada (SIMDIS, do inglês *Simulated Distillation*), foi estabelecida a partir do desenvolvimento de colunas capilares com materiais mais resistentes a altas temperaturas, como as colunas metálicas, de vidro e as de sílica fundida.²¹⁻²⁴ Os métodos disponíveis utilizam programação de temperatura, chegando-se atualmente à eluição de amostras de petróleo até 720 °C.⁷⁶ A destilação simulada baseia-se no princípio de componentes de uma amostra eluírem em ordem crescente de seus pontos de ebulição. Para isso utiliza-se coluna cromatográfica/fase estacionária, com injetor a frio ou injetor por vaporização com temperatura programável, fases estacionárias apolares (metil-polisiloxanas; fenilmetilpolisiloxanas) e detector FID.²¹⁻

24

Ferreira e Aquino Neto (2005)²⁴ realizaram comparação entre os métodos de destilação física e destilação simulada do petróleo, relatando as similaridades e as peculiaridades das técnicas. Os autores descreveram que: “[...] quando operado em condições apropriadas, um cromatógrafo a gás funciona como uma unidade de microdestilação altamente eficiente.” Dessa forma a análise cromatográfica pelo método deve ser capaz de fornecer uma curva PEV semelhante a que é gerada pelo método de destilação física, com o diferencial de que essa curva passa a ser obtida em um intervalo menor de tempo e são alcançados valores mais altos de Ponto Final de Ebulição (PFE) (Figura 3.12).²¹⁻²⁴

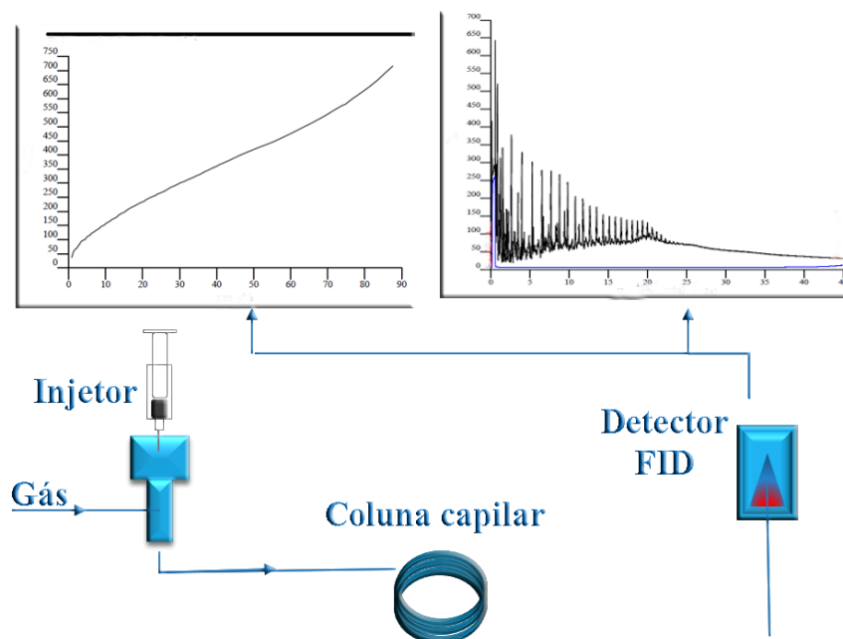


Figura 3.12 Desenho esquemático do funcionamento de um cromatógrafo a gás para método de cromatografia gasosa de alta temperatura (HTGC) com detector FID e geração de curva de destilação simulada SIMDIS.²⁴
Fonte: Autor.

A partir de 1973, a ASTM passou a estabelecer métodos padronizados de análise por destilação simulada, como os métodos D-2887⁷⁷, para análise de produtos de petróleo com ponto final de ebulição até 538 °C; D-5307⁷⁸, para análise de óleo cru até 538 °C; e D-7169⁷⁶, para análise de óleo cru até 720 °C (*n*-C120), complementando procedimentos de destilação física.

Devido à pequena quantidade de amostra utilizada na determinação da curva PEV o método de destilação simulada apresenta limitação na região de leves na curva, na altura de T10%, que é a temperatura na qual ocorre recuperação de 10%, em massa ou em volume, do material eluído ou destilado. Devido a esta limitação no método, a ASTM 7169⁷⁶ e D7900⁷⁴ orienta a junção dos resultados da técnica de destilação simulada com os resultados de análise detalhada de hidrocarbonetos para uma análise completa de petróleo bruto.

3.3 Quimiometria

Na química, principalmente na química analítica, com a modernização da instrumentação houve o aumento da quantidade e variedade dos dados experimentais, com produção de grande quantidade de informações a serem processadas. Dessa forma, surgiu a necessidade de utilizar ferramentas estatísticas e matemáticas para processamento dessas variáveis com o objetivo de extrair

informações de compreensões químicas desses dados. Assim surgiu a quimiometria, destinada à análise de dados químicos multivariados.^{29,34-35,79}

A quimiometria pode ser conceituada de forma resumida como “a administração e processamento de informações de natureza química”, com utilização de ferramentas matemáticas e estatísticas, aplicada a diversas áreas da química.²⁹ Os métodos quimiométricos podem ser divididos em três grandes áreas: planejamento e otimização de experimentos, métodos qualitativos e de reconhecimento de padrões e calibração multivariada.^{29,34} A química analítica teve grande impacto com o desenvolvimento das técnicas de calibração multivariada. Nesta é possível estimar uma propriedade de interesse a partir de outras medições, que normalmente são espectros obtidos por procedimentos analíticos mais simples, rápidos, menos onerosos e dependem de menor quantidade de amostra.

Essa aplicação de ferramentas envolve uma série de etapas e de procedimentos para interpretação de dados gerados após a medição de magnitude de um instrumento. Diferentes métodos são aplicados para pré-processamento de dados, em que envolve algoritmos para ajustes e redução de ruídos, visualização e classificação dos dados, redução da dimensionalidade e ferramentas de calibração multivariada para relacionar as variáveis medidas com as propriedades de interesse.^{29,34-35,79}

Para aplicação das técnicas quimiométricas, os dados reportados pelos instrumentos, como os cromatogramas ou os espectros, são organizados em um arranjo de linhas e colunas, constituindo uma matriz de dados $X_{(m \times n)}$. Cada linha representa uma amostra e contém as respostas instrumentais para a mesma, como pode ser observado na Figura 3.13. Cada variável de resposta instrumental (tempo de retenção, número de onda, deslocamento químico, etc.) é emparelhada na mesma coluna, sendo esse conjunto de dados geralmente compreendido em variáveis independentes. O outro conjunto de dados é constituído dos valores da propriedade de interesse a serem estimados, geralmente compreendidos como variáveis dependentes. É organizado em matriz $Y_{(m \times p)}$, caso se tenha mais de uma variável, ou vetor y , no caso de uma variável com a propriedade de interesse.^{29,34-35,79}

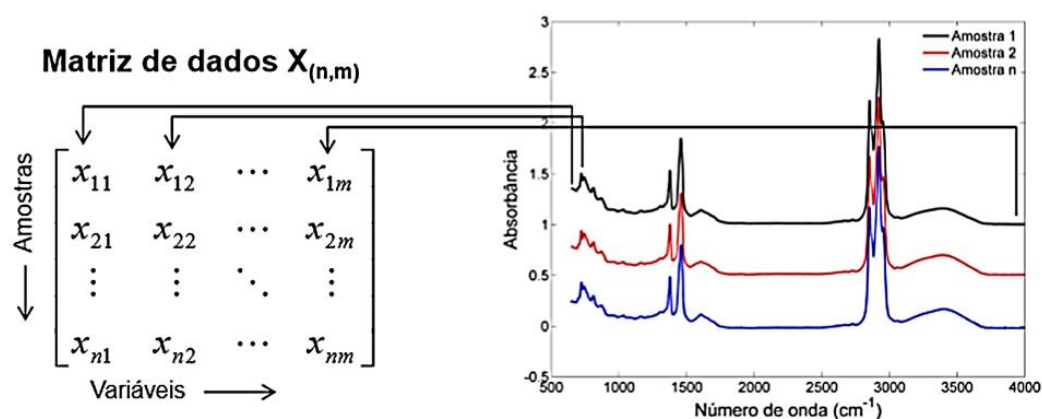


Figura 3.13 Representação esquemática de organização de dados espectrais em matriz.
Fonte: Filgueiras (2011)⁵⁸

Para associação de métodos quimiométricos com análises cromatográficas, os dados podem ser originados de duas fontes: direto do cromatograma, ou dados derivados do perfil cromatográfico, como intensidade do sinal, área do pico, composição ou porcentagem em massa.^{27,28}

Quando se estabelece uma relação matemática entre as variáveis dependentes e uma única variável independente, essa relação é dada pela calibração univariada. Quando a relação é dada entre as variáveis dependentes e várias variáveis independentes, é uma calibração multivariada, cuja maior parte é linear. Os métodos multivariados são adequados para os dados químicos por permitirem estudos com várias espécies presentes, independente de presença ou ausência de diferenças espectrais ou de correlações nos dados, assim como identificação de problemas como linha de base e interferentes, uma vez que sejam inseridos na calibração e na previsão.^{29,34-35,79-80}

3.3.1 Calibração multivariada

Calibração é um conjunto de procedimentos que visa encontrar um algoritmo matemático que produza informações respectivas à determinada propriedade de interesse, a partir de resultados registrados por um instrumento ao qual as amostras foram submetidas. Uma vez encontrado o algoritmo pelo procedimento de calibração, este poderá ser utilizado para previsão da propriedade em amostras de informações desconhecidas, submetidas à mesma análise. Um resumo gráfico do procedimento

de calibração com análise cromatográfica de amostras de petróleo está demonstrado no diagrama abaixo (Figura 3.14).⁷⁹

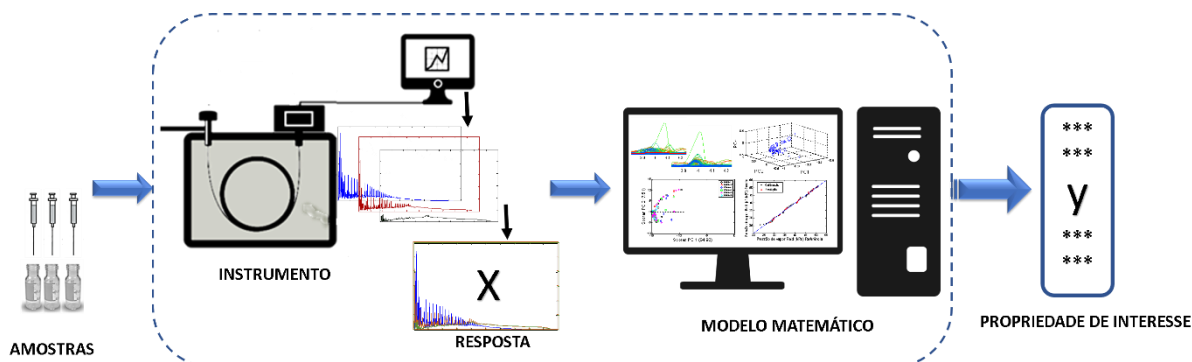


Figura 3.14 Representação gráfica de procedimento de calibração multivariada em análise cromatográfica de petróleo.

Fonte: adaptado de Ferreira et al., 1999.⁷⁹

As etapas principais do processo de calibração são modelagem e validação do modelo. A modelagem é o procedimento da relação matemática entre X e a matriz Y ou o vetor y . Em problemas quantitativos y_i representa o valor da propriedade de interesse da amostra, e em problemas qualitativos a classe à qual a amostra pertence.^{29,34-35,58,79-80} A validação é o procedimento de otimização do modelo no sentido de encontrar melhor descrição da propriedade de interesse. Com o modelo de calibração ajustado e validado pode ser utilizado para previsão de outras amostras. Etapas intermediárias de pré-processamento podem ser úteis uma vez que os dados originais podem não ter as distribuições adequadas, podendo dificultar a extração de informações químicas relevantes.^{29,34-35, 58,79-80}

As diferentes etapas que consistem na calibração multivariada estão resumidamente descritas abaixo:

- a) Preparação de conjunto de calibração, que compreende a obtenção de amostras que sejam representativas de toda a informação a ser modelada. As amostras que consistem no conjunto de calibração devem ter a informação da propriedade de interesse de forma confiável;
- b) Registro do sinal analítico, que compreende as análises instrumentais;
- c) Organização dos dados, análise exploratória e pré-processamento, que compreende a minimização de possíveis contribuições não desejadas dos sinais que possam diminuir a capacidade preditiva do modelo;

- d) Separação dos conjuntos de dados em subconjuntos de calibração do modelo e de previsão.
- e) Construção do modelo com seleção do modelo que melhor estabeleceu a relação matemática entre sinal instrumental e propriedade de interesse;
- f) Validação do modelo para assegurar a capacidade preditiva. A validação compreende o estudo quantitativo dos resultados de aplicação do modelo em amostras que não fizeram parte do conjunto de calibração;
- g) Aplicação no modelo para previsão de amostras cuja propriedade de interesse deve ser estimada;

Um dos métodos de regressão comumente aplicados é o de componentes principais (PCR). Neste método os dados da matriz $\mathbf{X}$ são decompostos em Componentes Principais (PC) e uma regressão linear múltipla é realizada relacionando os escores da matriz $\mathbf{X}$ com a matriz/vetor da propriedade de interesse. Todas as variáveis relevantes são incluídas neste método para a construção do modelo, o que faz com que a calibração seja eficiente mesmo com presença de interferentes. A base do método de regressão pelo PCR é a análise de componentes principais (PCA).^{29,34-35, 58,79-80}

3.3.2 Análise de Componentes Principais (PCA)

A PCA é principalmente utilizada pela sua capacidade de compressão dos dados em função da existência de correlação entre diversas variáveis medidas. É uma técnica estatística que transforma linearmente um conjunto original de variáveis em um conjunto substancialmente menor de variáveis não correlacionadas que contém a maior parte da informação do conjunto original.^{29,34-35, 58,79-80}

A PCA é basicamente a manipulação da matriz de dados com o objetivo de representar as variações presentes em muitas variáveis, por um número menor de fatores. Para isso é construído um sistema de eixos denominados componentes principais ou autovetores para representar as amostras, diminuindo a dimensão dos dados. A PCA é muito utilizada como método de análise exploratória dos dados, para identificação de amostras anômalas (*outliers*) e métodos de classificação e predição de resultados.^{29,34-35, 58,79-81}

No processamento dos dados pela PCA a matriz $\mathbf{X}$ é decomposta em duas matrizes que capturam os dados relevantes, conforme Equação 3.2 e Figura 3.15.

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E} \quad (3.2)$$

Onde $\mathbf{X}$ corresponde à matriz ($m \times n$) de dados instrumentais, $\mathbf{T}$ a matriz ($m \times h$) de escores para h componentes principais e $\mathbf{P}^T$ a matriz ($h \times n$) dos pesos (*loadings* ou autovetores). $\mathbf{T}$ descreve dados das amostras, onde m corresponde às linhas na matriz $\mathbf{X}$, ou seja, número de amostras. $\mathbf{P}^T$ descreve os dados da leitura instrumental, onde n corresponde às colunas na matriz $\mathbf{X}$, ou seja, variáveis instrumentais. h corresponde aos componentes principais utilizados para descrever $\mathbf{X}$. $\mathbf{E}$ é a matriz de resíduos, que contém a parte dos dados não explicados pelo modelo de componentes principais. As componentes principais descrevem o máximo de variância. ^{29,34-35, 58,79-81}

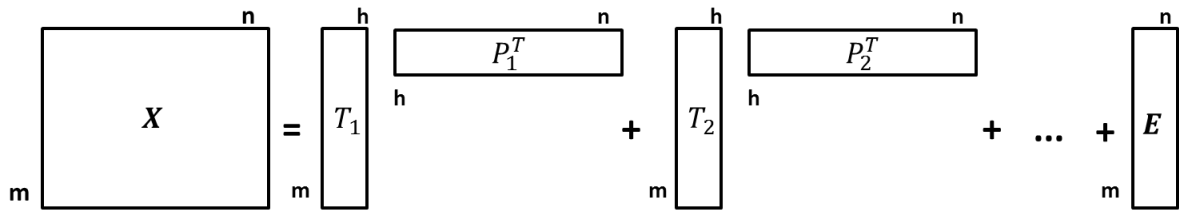


Figura 3.15 Demonstração gráfica do princípio da PCA.
Fonte: adaptado de Geladi e Kowalski (1986)⁸¹

A matriz de escores $\mathbf{T}$ tem as seguintes propriedades: número de linhas igual ao número de linhas dos dados originais; número de colunas igual ao número de fatores significantes nos dados originais (componentes principais); e soma dos quadrados dos elementos de cada coluna da matriz de escores, descrita frequentemente como autovalor. O maior valor de autovalor significa o componente mais significativo. A matriz de pesos $\mathbf{P}$ (*loadings*) tem as seguintes propriedades: número de colunas igual ao número de colunas dos dados originais; número de linhas igual ao número de componentes principais; e o valor 1 (um) como resultado da soma do quadrado dos elementos de cada coluna. ^{29,34-35, 58,79-81}

Dessa forma cada componente principal h é caracterizado por meio de um vetor-coluna da matriz de escores t_h , um vetor-linha da matriz de *loadings* p_h e um autovalor g_h . Este último é definido pela equação 3.3.

$$g_h = \sum_{i=1}^I t_{ih}^2 \quad (3.3)$$

As componentes principais geralmente são descritas geometricamente. Os dados são representados como pontos em espaços dimensionais e o padrão formado neste novo espaço indica as informações dos dados. A primeira PC pode ser definida como a linha reta de melhor ajuste neste novo espaço multidimensional. Os *loadings* representam a direção (ângulo) da linha reta (Figura 3.16(A)). Os escores representam a distância a esta linha (Figura 3.16(B)).^{29,34-35, 58,79-81}

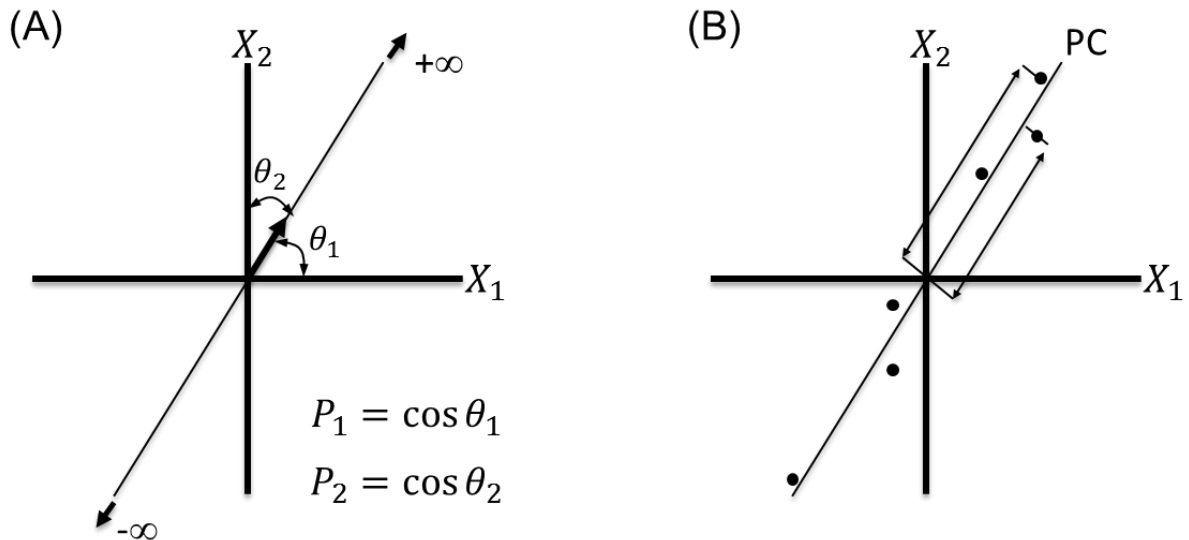


Figura 3.16 Exemplo gráfico de projeção de *loadings* e escores na PCA de uma componente principal para duas variáveis. (A) peso (*loadings*) são os cossenos dos ângulos do vetor de direção; (B) escores são as projeções das amostras (pontos) na direção da componente principal (PC).

Fonte: adaptado de Geladi e Kowalski (1986)⁸¹

3.3.3 Regressão por componentes principais (PCR)

A regressão pelo PCR utiliza a PCA para modelagem quantitativa em dados complexos, sendo considerado um modelo robusto, que significa que seus parâmetros não sofrem alteração quando são incluídas novas amostras no conjunto de calibração. Nesta regressão os dados da matriz $\mathbf{X}$ são representados pela sua matriz de escores. A aproximação em direção aos dados de propriedade de interesse é dada pela retirada dos vetores de escores correspondentes aos menores autovalores, resolvendo problemas de colinearidade.^{29,34-35, 58,79-81}

Porém, o método PCR tem como característica principal a construção das componentes principais utilizando somente a matriz $\mathbf{X}$, não considerando as informações da propriedade de interesse. Isso se caracteriza uma desvantagem do método uma vez que pode ser necessário um número maior de componentes principais para a construção do modelo.

3.3.4 Regressão por mínimos quadrados parciais (PLS)

A regressão pelos mínimos quadrados parciais, o PLS tem sido amplamente consolidado e divulgado em calibração multivariada na caracterização de petróleo e produtos de petróleo.^{5,10,34,45-19,58} Existem diversos algoritmos para PLS, a partir do desenvolvimento dos primeiros algoritmos.⁸⁰⁻⁸² Geladi e Kowalski (1986)⁸¹ propuseram um algoritmo baseado nos primeiros, sendo considerado mais completo com alguns pontos ajustados, principalmente na decomposição das matrizes de dados instrumentais e de propriedade de interesse de forma simultânea, sendo importante para aplicação em previsões.

O método de regressão pelos mínimos quadrados parciais, o PLS, também utiliza a modelagem pelas componentes principais, mas não apresenta a dificuldade apresentada acima do PCR, uma vez que o PLS utiliza as informações da propriedade de interesse para a obtenção dos fatores, das variáveis latentes. O modelo de regressão pelo PLS é bastante robusto, com a inclusão de novas amostras no conjunto de calibração sem alteração de seus parâmetros, desde que essas contenham características semelhantes aos dados originais. As variáveis latentes são combinações lineares das componentes principais calculadas pelo método PCR. A primeira variável latente descreve a direção de máxima variância que também se correlaciona com a propriedade de interesse.⁷⁹⁻⁸¹

PLS emprega análise de fatores, com o diferencial de utilizar a informação da matriz de resposta $\mathbf{Y}$ no cálculo das variáveis latentes (equivalentes às componentes principais na PCA). Os fatores Variáveis Latentes (VL) são definidos de modo a manter um compromisso entre a explicação da variância em $\mathbf{X}$ e a previsão da variável de interesse, $\mathbf{Y}$.^{29,35,79-81}

No método de regressão PLS há duas variantes denominadas PLS1 e PLS2. No método PLS1 a regressão é feita para uma única variável de interesse, dessa forma $\mathbf{y}$ é um vetor coluna, ou seja, para cada constituinte de interesse é calculado

um conjunto de escores e variáveis latentes. No método PLS2, $\mathbf{Y}$ é uma matriz e calcula-se um único conjunto de escores e variáveis latentes para todas as colunas dessa matriz, que é comum a todas as propriedades de interesse que estão sendo modeladas.^{29,35,79-81}

O modelo é obtido por um processo iterativo de otimização simultânea da projeção das amostras sobre os pesos para a determinação dos escores, seguido de ajuste por uma função linear dos escores da matriz $\mathbf{X}$ aos escores do vetor $\mathbf{y}$. Os pesos ou *loadings* correspondem ao peso (Influência) de cada variável na variável latente formada. Os escores correspondem aos valores de projeção nos novos eixos formados para cada amostra.^{29,35,79-81}

A matriz $\mathbf{X}$ de medidas instrumentais com m amostras e n variáveis, e a matriz $\mathbf{Y}$, ou vetor $\mathbf{y}$, com m medidas de p parâmetros, como propriedades físico-químicas, por exemplo, são simultaneamente decompostas em uma soma de h fatores ou variáveis latentes, representados nas equações 3.4 e 3.5 e graficamente na Figura 3.17.

$$\mathbf{X}_{(m \times n)} = \mathbf{T}_{(m \times h)} \cdot \mathbf{P}_{(h \times n)}^T + \mathbf{E}_{x(m \times n)} \quad (3.4)$$

$$\mathbf{y}_{(m \times p)} = \mathbf{U}_{(m \times h)} \cdot \mathbf{q}_{(h \times p)}^T + \mathbf{e}_{y(m \times p)} \quad (3.5)$$

Onde $\mathbf{T}_{(m \times h)}$ e $\mathbf{U}_{(m \times h)}$ são matrizes de escores da matriz $\mathbf{X}$ e vetor $\mathbf{y}$, respectivamente, $\mathbf{P}_{(h \times n)}^T$ e $\mathbf{q}_{(h \times p)}^T$ são a matriz e vetor de pesos e $\mathbf{E}_x$ e $\mathbf{e}_y$ são os resíduos de $\mathbf{X}$ e de $\mathbf{y}$.

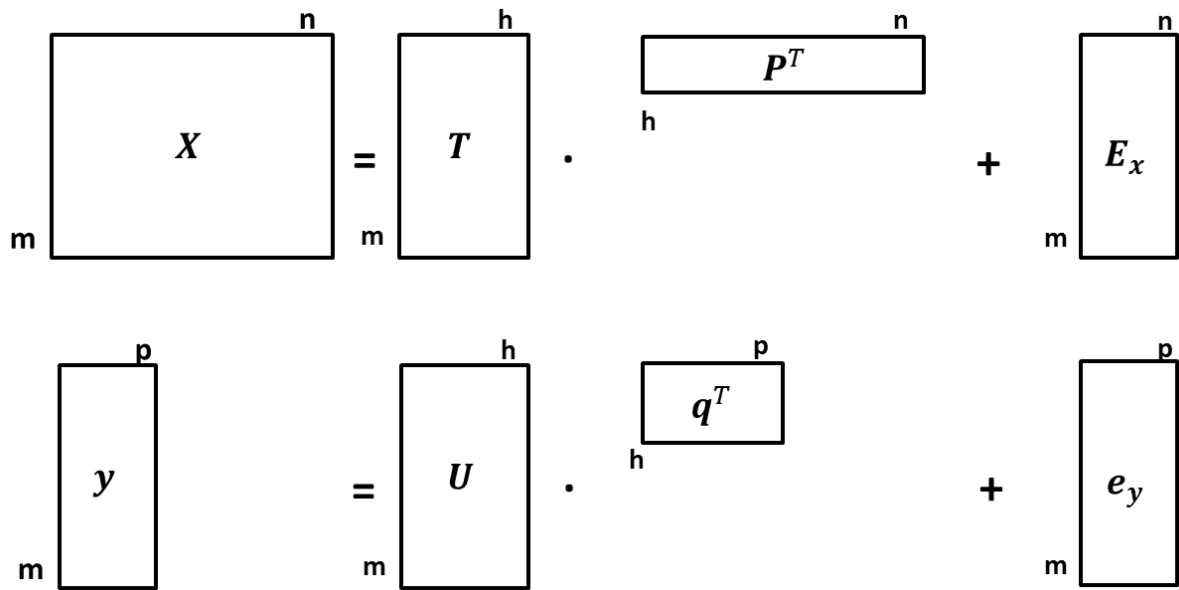


Figura 3.17 Demonstração gráfica do princípio do PLS.
 Fonte: adaptado de Geladi e Kowalski (1986)⁸¹

A matriz de escores é estimada como uma combinação linear das variáveis originais com coeficientes de fator-peso proporcional à covariância entre X e y (Equação 3.6). A matriz de resíduos, E_x contém a informação em X que não foi modelada, assim como o vetor de resíduos e_y para y .

$$T = X \cdot W \quad (3.6)$$

Onde W é a matriz de fatores-peso. O fator peso $\hat{w}_i$ é estimado para cada variável latente, como um coeficiente de regressão por quadrados mínimos das colunas de X em y no espaço das variáveis originais e normalizado (Equação 3.7 e Equação 3.8).

$$\hat{w}_i = X^T \cdot y (y^T \cdot y)^{-1} \quad (3.7)$$

$$w_i = \frac{X^T \cdot y}{\|X^T \cdot y\|} \quad (3.8)$$

A partir da matriz de fatores-peso W são obtidos os coeficientes de regressão b_{PLS} (Equação 3.9), que darão a relação linear entre os dados de X e y .

$$\mathbf{b}_{PLS} = \mathbf{W}(\mathbf{P}^T \cdot \mathbf{W})^{-1} \cdot \mathbf{q}^T \quad (3.9)$$

Dessa forma, o modelo de regressão linear PLS pode ser representado pela equação 3.10 e amostras futuras podem ter valores de propriedade de interesse previstos utilizando a combinação linear da nova matriz pelos coeficientes do modelo construído.

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{b}_{PLS} \quad (3.10)$$

Uma importante diferença do PLS para o PCR é que o PLS leva em consideração o erro nas estimativas do parâmetro de interesse e nos dados instrumentais.⁸¹ Assim como na PCA, no PLS o produto de $\mathbf{T} \cdot \mathbf{P}$ aproxima os dados instrumentais e o produto entre $\mathbf{U} \cdot \mathbf{q}$ aproxima os dados do parâmetro de interesse. A matriz de escores é gerada tanto para os dados instrumentais quanto para os dados de propriedade. Os autovalores geralmente são calculados como a soma dos quadrados dos escores $\mathbf{T}$ e $\mathbf{U}$, respectivamente. Pelo PLS os autovalores distinguem-se do PCA, pois dependem de ambos os conjuntos de dados.^{29,35,79-81}

Cada componente principal, que é denominado no PLS de variável latente, irá aproximar melhor os dados instrumentais à propriedade de interesse. Para cada variável latente haverá um vetor de escores $\mathbf{t}_m$, um vetor de escores $\mathbf{u}_m$ um vetor de *loading* $\mathbf{p}_h$ e um vetor de *loading* $\mathbf{q}_h$.

A validação do modelo é realizada para se encontrar o melhor número de fatores, ou seja, otimizar número de variáveis latentes. A validação do modelo pode ser um processo externo ou interno (validação cruzada, ou CV, do inglês *cross-validation*). O processo externo utiliza amostras que não tenham relação com o conjunto de calibração, mas que sejam representativas tanto da calibração quanto de amostras futuras. O processo interno, validação cruzada, tem o conjunto de calibração dividido em vários segmentos, dos quais um é designado para a validação do modelo e os demais para construção de modelos de calibração. São construídos modelos de acordo com o número de segmentos, excluindo cada segmento do modelo de

calibração e utilizando na validação, sendo selecionado um modelo com menor resíduos de resposta como modelo final. ^{29,35,58,79-81}

Dentre os métodos de validação interna destacam-se os dois mais comumente aplicados: *leave-one-out* e *k-fold*. No método *leave-one-out* uma amostra é removida do conjunto de calibração para utilizar na validação. É construído o modelo com as amostras restantes ($n-1$). O processo de construção do modelo é repetido até que todas as amostras de calibração tenham sido utilizadas para validação. No método *k-fold* as amostras de calibração são divididas em k subconjuntos de mesmo tamanho. Enquanto um subconjunto é utilizado para validação, os demais são utilizados na construção do modelo ($k-1$). O procedimento é repetido até que todos os subconjuntos tenham sido utilizados na validação. Para obtenção dos subconjuntos pode-se utilizar diferentes parâmetros, como: blocos (*contíguos block*); aleatórios (*random block*); ou ordenados, também denominado persianas (*venetian blinds*). ^{29,35,58,79-81}

Na validação cruzada diferentes modelos são gerados na previsão do parâmetro de interesse para as amostras de calibração, sem que sejam inseridas de forma simultânea em modelo de calibração e de validação. O número ótimo de variáveis latentes para construção do modelo é determinado pela avaliação da raiz quadrada do erro quadrático médio de validação cruzada (RMSECV, do inglês *root mean square error of cross validation*), que é calculado pelos valores previstos das amostras de calibração na validação cruzada (Equação 3.11). ^{29,35,58,79-81}

$$RMSECV = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{n}} \quad (3.11)$$

Onde y_i e $\hat{y}_i$ são os valores de referência e estimado da propriedade de interesse sendo modelada, pelo modelo de validação cruzada para n amostras do conjunto de calibração. Na etapa de validação cruzada o modelo gera um gráfico de RMSECV em função de número de variáveis latentes, representado na Figura 3.18. Pelo gráfico é determinado um número ótimo de VL que corresponda ao menor valor de RMSECV, ou menor número de VL em que não altere significativamente o RMSECV.

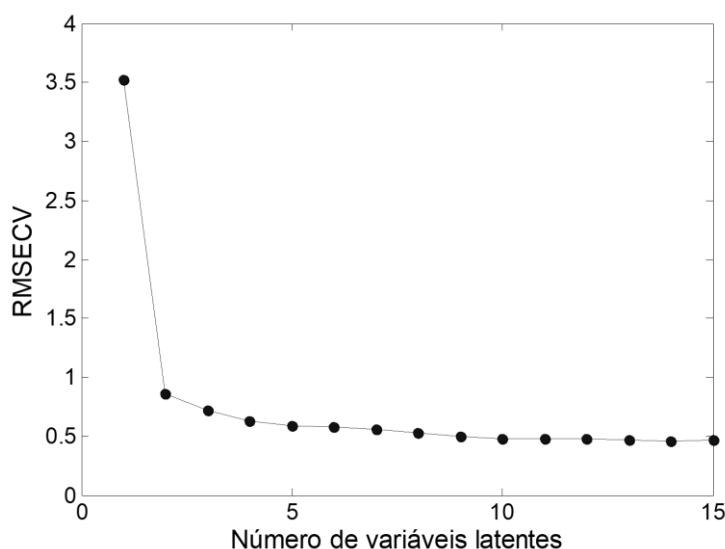


Figura 3.18 Gráfico demonstrativo de RMSECV em função de número de variáveis latentes na etapa de otimização de modelo de regressão por PLS, método de validação cruzada k-fold.
Fonte: dados de projeto deste trabalho.

O número de variáveis latentes é uma importante propriedade do modelo. Um número pequeno de fatores pode prejudicar a relação entre valores previstos e de referência da propriedade de interesse, porque faltará componente para descrever o modelo (dimensionalidade do modelo) ou fontes de variações sistemáticas. Isso é chamado de subajuste do modelo. Se o número de fatores estiver em excesso haverá um bom ajuste da curva com dados de calibração, porém no conjunto de previsão poderá haver ajuste insatisfatório, possivelmente pela entrada de ruído ou outro interferente instrumental ao modelo de calibração. Neste caso, ocorreu um sobreajuste do modelo. ^{29,35,58,79-81}

Uma vez realizada a otimização do modelo com sua validação, o modelo está pronto para previsão de um conjunto denominado de previsão ou de teste. Este conjunto compreende amostras não utilizadas no conjunto de calibração nem de validação do modelo, cuja separação do conjunto de dados original é realizada antecipadamente após devidos pré-processamentos dos dados.

Uma importante aplicação dos modelos de regressão é a predição das variáveis da propriedade de interesse y' a partir dos dados instrumentais X' de novas amostras. A predição é realizada pela decomposição do bloco X' e construção do bloco y' . Isso é possível com a utilização dos vetores de peso, fatores-peso e

coeficientes de regressão construídos no modelo de calibração, estando o modelo pronto para ser aplicado ao novo conjunto amostral.

3.3.5 Pré-tratamento e pré-processamento dos dados

Muitas vezes, antes de iniciar o desenvolvimento do modelo, realizam-se procedimentos nos dados após estes serem organizados em matrizes, para que sejam reduzidas as variações indesejáveis. As técnicas instrumentais que geram sinais como resposta, apresentam em seus espectros, ou cromatogramas, por exemplo, três componentes: sinal, ruído e fundo, que se diferem pela frequência. O ruído é caracterizado pela alta frequência em comparação ao sinal verdadeiro, o fundo é o de menor frequência e o sinal verdadeiro, usualmente, uma frequência intermediária. Consideram-se duas contribuições fornecidas pelos sinais registrados no instrumento: o sinal verdadeiro, com as informações sistemáticas, denominada contribuição determinística do sinal; e contribuição estocástica, que corresponde às variações indesejadas, denominada de ruído. Estas variações indesejadas, quando não eliminadas pelos sistemas de aquisição de dados, muitas vezes acoplados ao instrumento, se não reduzidas antes da análise quimiométrica, podem influenciar os resultados.^{27,32}

O ruído aleatório é constituído de erros experimentais, de medida, de amostragem, ou outra fonte. Outras fontes de variação de baixa frequência indesejáveis podem ser informações sistemáticas, mas, que não estejam relacionadas com o processo avaliado. Essas variações indesejadas apresentadas juntamente com o sinal verdadeiro, sejam de origem determinística ou aleatória, dependerão do tipo de instrumento e das condições da análise.³² Dessa forma, é primordial a realização de uma etapa de preparação dos dados para a análise com pré-tratamentos adequados. O pré-tratamento consiste em transformações necessárias dos dados instrumentais originais, minimizando variações aleatórias e também sistemáticas como ruído, linha de base, entre outros. A escolha adequada do método é essencial para o sucesso da análise multivariada, devendo-se, geralmente, testar vários métodos para se alcançar a preparação mais adequada.³²

Os procedimentos podem ser aplicados aos dados referentes às amostras (linhas da matriz de dados), ou aos dados instrumentais, às variáveis (colunas da matriz). Neste trabalho, são descritos pré-tratamentos os métodos aplicados às amostras e pré-processamentos, os aplicados às variáveis.³² Seguem abaixo a

descrição de alguns métodos comumente aplicados, para pré-tratamento e pré-processamento dos dados, aplicados ou testados neste trabalho.^{27,32,79-80,83-87}

3.3.5.1 Correção de deslocamentos ao longo da linha de base

Deslocamentos de picos ao longo da linha de base são comuns em cromatogramas e espectro de RMN, porém podem ainda ser encontrados em espectros de infravermelho, por variações em temperatura na análise.⁸³⁻⁸⁴ Na cromatografia os deslocamentos de pico ocorrem devido às variações de temperatura, tempo de uso da coluna, possíveis alterações na fase móvel e diferenças na injeção da amostra. A correção posterior desse deslocamento é realizada por meio de algoritmos de alinhamento dos picos. O alinhamento consiste em alinhar um cromatograma desalinhado a um cromatograma de referência, que pode ser um cromatograma médio ou que apresente perfil de eluição com picos bem resolvidos.⁸³⁻

84

Existem diversos algoritmos para esse fim, sendo mais comumente utilizado para alinhamento cromatográfico o método de correlação otimizada (COW, do inglês *correlation optimized warping*).^{27,29,83-84} Este método consiste em corrigir mudanças de pico em um cromatograma por meio de alongamento linear e compressão, até que a correlação entre o cromatograma e a referência, denominado cromatograma alvo, seja máxima. Dois parâmetros de configuração para ajuste da qualidade do alinhamento são o número de seções em que o cromatograma será dividido e o parâmetro de distorção que define a flexibilidade das distorções para o alinhamento.

Neste trabalho utilizou-se o método de correlação por intervalos (*Icoshift*, do inglês *interval-correlation-shifting* proposto por Savorani *et al.* (2010)).⁸³ Este método, desenvolvido inicialmente para alinhamento de espectros de RMN, mas também aplicável a outros espectros e cromatogramas,^{51,83-84} é baseado no deslocamento de intervalos em um cromatograma, de forma simultânea em todos os cromatogramas ou espectros, de um conjunto de dados. O algoritmo utiliza um mecanismo de otimização da correlação baseado na Transformada de Fourier, denominado FFT (do inglês, *Fast Fourier Trasnformation*) para promover velocidade de cálculos e automação do alinhamento. Os deslocamentos resultam no alinhamento de cada pico de um cromatograma para um cromatograma alvo, maximizando a correlação cruzada entre os intervalos definidos pelo usuário.^{51,83-84}

Os parâmetros de configuração para realização do alinhamento pelo método do *Icoshift* são divididos em três partes: definição de intervalos; maximização da correlação cruzada, pelo mecanismo de FFT, entre cada intervalo definido; e reconstrução do sinal.⁸³⁻⁸⁴

3.3.5.2 Correção de linha de base

A correção de desvios de linha de base pode ser realizada pela suavização com derivada.^{29,85} Outra forma é a utilização do método iterativamente adaptativo por mínimos quadrados ponderados e penalizados, denominado airPLS.⁸⁶ O algoritmo promove o ajuste da linha de base rapidamente, por meio da soma dos erros dos quadrados entre uma linha de base montada e sinais originais, cálculo de pesos e nova iteração. O gráfico resultante mantém o aspecto visual dos dados originais, facilitando a interpretação dos resultados.⁸⁶

3.3.5.3 Derivada

Pré-tratamento com derivadas é comum quando se analisa espectros em que ocorram problemas de deslocamento e inclinação da linha de base. Assume-se que os ruídos têm alta frequência em relação ao sinal de interesse. Esta aplicação pode ser desenvolvida com vários tipos de filtros: filtros de média móvel; filtros polinomiais (Savitzki-Golay); filtros baseados na Transformada de Fourier. Vale ressaltar que ao se utilizar a derivada, a relação entre a área abaixo de um pico do espectro e a concentração é preservada, sendo que permanece a correspondência quantitativa entre o espectro e a concentração, sendo inconveniente seu uso em relação à razão sinal/ruído, que pode ocorrer um decréscimo e por isso se a razão sinal/ruído for baixa pode não ser aconselhável.²⁹

3.3.5.4 Normalização

Os valores de cada uma das variáveis de uma amostra i são divididos por um fator de normalização, como, por exemplo, pela norma dessa amostra, que está representada na equação 3.12 $\|x_i\|$. Como resultado as amostras ficam localizadas em uma escala pré-determinada.^{29,85, 87-88}

$$x_{ij(norm)} = \frac{x_{ij}}{\|x_i\|} \quad (3.12)$$

3.3.5.5 Correção multiplicativa de espalhamento (MSC)

Também denominada Correção multiplicativa de sinal, a MSC (do inglês *Multiplicative Scatter Correction*), esse pré-tratamento é utilizado para corrigir efeitos de espalhamento aditivos e multiplicativos, comumente causados por meio de fenômenos físicos, como sensibilidade do detector, variações de temperatura e pressão, entre outros. Removem-se os efeitos de espalhamento com deslocamento e escalamento do espectro para que ele se ajuste ao espectro “ideal”, que pode ser o espectro de uma das amostras do conjunto de dados, ou o espectro médio. Os efeitos aditivos são corrigidos pelos parâmetros de deslocamento e os efeitos multiplicativos pelos de escalamento. Sendo ambos estimados simultaneamente por meio de regressão linear (Equação 3.13).^{29,85, 87-88}

$$\mathbf{x}_i = a_i \mathbf{1} + b_i \cdot \mathbf{x}_m + e_i \quad (3.13)$$

Onde $\mathbf{x}_i$ é um vetor coluna de cada cromatograma ou espectro, a_i e b_i são coeficientes de regressão estimados por quadrados mínimos de cada espectro ao espectro médio. O espectro corrigido é obtido pela subtração do termo constante a_i em cada variável dos dados originais $\mathbf{x}_i$ como correção dos efeitos aditivos e divisão desse resultado pelo coeficiente angular b_i como correção dos efeitos multiplicativos.²⁹

$$\mathbf{x}_{ij(msc)} = \frac{x_{ij} - a_i}{b_i} \quad (3.14)$$

3.3.5.6 Padronização normal de sinal (SNV)

SNV (do inglês, *Standard Normal Variate*) é apropriado para remover interferências de espalhamento e de tamanho de partícula sólida, além de corrigir efeitos aditivos e multiplicativos, assim como o MSC. Corresponde a um autoescalamento de cada linha da matriz original dos dados. A base do método corresponde então a padronizar a variação por espectro, obtendo resultado proporcional ao método MSC, com a subtração de cada variável na matriz de dados

pela média da respectiva linha, seguido de divisão desse resultado pelo desvio-padrão dessa linha (Equação 3.15).^{29,85, 87-88}

$$x_{ij(snv)} = \frac{x_{ij} - \bar{x}_i}{\sigma_i} \quad (3.15)$$

3.3.5.7 Centragem na média

Centragem na média é o pré-processamento mais frequente quando se trabalha com dados espectrais, tendo como resultado uma translação de eixos para o valor médio de cada um deles, preservando a estrutura dos dados. Centragem dos dados na média: calcula-se a média de cada coluna da matriz de dados (variáveis) e, em seguida, subtrai-se essa média de cada uma das variáveis da respectiva coluna, como na equação 3.16.²⁹

$$x_{ij(cm)} = x_{ij} - \bar{x}_j \quad (3.16)$$

3.3.5.8 Autoescalamento

Os métodos de escalamento consistem em dividir os dados por algum fator de escala de medida da dispersão dos dados. O autoescalamento é utilizado para comparação de variáveis de diferentes dimensões, devendo ser utilizado quando os dados foram adquiridos em uma mesma faixa de amplitude de sinal. No autoescalamento cada coluna da matriz de dados é centrada na média e, em seguida, é escalada pelo desvio-padrão dessa coluna (Equação 3.17).²⁹

$$x_{ij(a)} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (3.17)$$

É importante ressaltar, que o método de pré-tratamento e/ou pré-processamento devem ser definidos de acordo com o tipo de dados e com o objetivo, de modo a facilitar a análise quimiométrica e a interpretação das informações químicas relevantes.²⁹

3.3.6 Avaliação dos modelos

Uma etapa importante na construção do modelo PLS é a avaliação da qualidade preditiva do modelo construído. Além do número de variáveis latentes definido na etapa de validação do modelo, outros fatores importantes são a presença de amostras anômalas (*outliers*) e a presença erros sistemáticos e de tendência nos resíduos.^{29,47,80} Amostras anômalas identificadas pelas ferramentas da PCA e confirmadas na etapa de avaliação do modelo devem ser avaliadas criteriosamente quanto à sua exclusão, para não prejudicar a capacidade preditiva do modelo. Entretanto, algumas amostras podem apresentar comportamento atípico por representar um tipo de variabilidade importante dos dados instrumentais e sua exclusão pode não ser indicada.²⁹ A avaliação estatística do modelo inclui a avaliação dos resíduos gerados E_x e e_y (Equação 3.18). A soma dos quadrados dos resíduos de y é indicativo de qualidade preditiva do modelo. A soma dos quadrados dos resíduos de X é indicativo de quantidade de dados de X não usados no modelo, o que pode significar propriedades inesperadas nos dados ou presença de erros experimentais, quando grande parte de X não foi usado no modelo.⁸⁰

$$e_i = y_i - \hat{y}_i \quad (3.18)$$

Onde y_i é o valor de referência, $\hat{y}_i$ o valor previsto e e_i o valor residual da amostra i . Um modelo bem ajustado geralmente apresenta resíduos distribuídos normalmente acerca de zero. Por outro lado, modelos fracamente ajustados refletem tendências nas distribuições dos resíduos, indicando que informações relevantes podem não ter sido incorporadas no modelo.¹⁷

A exatidão (acurácia) do modelo para o conjunto de amostras de calibração é avaliada pela métrica da raiz quadrada do erro quadrático médio do conjunto de calibração (RMSEC, do inglês *Root Mean Square Error of Calibration*), pela Equação 3.19.⁸⁹⁻⁹⁰

$$RMSEC = \sqrt{\frac{1}{v} \sum_{i=1}^{ncal} (y_i - \hat{y}_i)^2} \quad (3.19)$$

Onde $ncal$ é o número de amostras do conjunto de calibração, y_i e $\hat{y}_i$ são os valores da propriedade de interesse, de referência e previsto, respectivamente e ν o número de graus de liberdade. Este último, ν , é estimado no modelo de calibração pelo PLS em: $\nu = ncal - LV - 1$, sendo LV o número de variáveis latentes.⁸⁹⁻⁹⁰

A capacidade preditiva para amostras desconhecidas ou acurácia de previsão do modelo é avaliada pela métrica de raiz quadrada do erro quadrático médio do conjunto de previsão (RMSEP, do inglês *Root Mean Square Error of Prediction*), Equação 3.20.^{17, 89-91}

$$RMSEP = \sqrt{\frac{1}{\nu} \sum_{i=1}^{npre} (y_i - \hat{y}_i)^2} \quad (3.20)$$

Onde $npre$ é o número de amostras do conjunto de previsão e ν neste caso será $\nu = npre$. Um modelo com maior acurácia requer que o erro de previsão seja minimizado.^{17, 89-91}

A linearidade do modelo descreve a qualidade do ajuste linear entre os valores de referência e os previstos pelo modelo de regressão. O coeficiente de determinação R^2 (Equação 3.21) é o parâmetro de avaliação do ajuste linear dos modelos PLS.^{17, 89-91}

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (3.21)$$

Onde $\bar{y}$ corresponde ao valor médio de referência. O valor do coeficiente de determinação varia de 0 a 1, sendo que o modelo de melhor ajuste resultará em um R^2 mais próximo de 1. A Equação 3.21 é utilizada para avaliação dos modelos de validação cruzada, R^2_{cv} , de calibração, R^2_c e de previsão, R^2_p .

O modelo deve ainda ser avaliado quanto à presença de erros aleatórios e sistemáticos, que podem ser originados de dados experimentais ou de calibração não adequada. Modelos com presença de erros sistemáticos e de tendência nos resíduos podem ser considerados insatisfatórios.^{17, 89-91}

3.3.6.1 Avaliação de erros sistemáticos.

Se houver erros sistemáticos significativos no modelo, este apresenta viés, ou seja, uma tendência e os erros não serão centrados numa distribuição gaussiana. Essa tendência afeta a estimativa sempre para um sentido, fornecendo resultados abaixo ou acima do esperado. O método de avaliação do modelo para identificar a presença de erros de viés geralmente é pelo teste para bias, que representam o desvio médio de n medições (Equação 3.22).^{17, 89-91}

$$Bias = \frac{\sum_{i=1}^I (y_i - \hat{y}_i)}{n} \quad (3.22)$$

Onde n é o número de amostras. O teste para bias geralmente é realizado para avaliação do modelo construído para o conjunto de calibração e para amostras de previsão.^{17,90}

A performance do teste de bias é avaliada pelo teste- t bicaudal (Equação 3.23), cujas hipóteses testadas são:

$$H_0: bias = 0;$$

$$H_1: bias \neq 0.$$

$$t_{calc} = \frac{|bias| \cdot \sqrt{n}}{SDE} \quad (3.23)$$

Onde SDE é o desvio padrão dos erros (SDE , do inglês *Standard Deviation of the Errors*), que é calculado pela Equação 3.24.^{17, 89-91}

$$SDE = \frac{\sum_{i=1}^I (y_i - \hat{y}_i - bias)^2}{n-1} \quad (3.24)$$

O valor do teste estatístico calculado t_{calc} é comparado com valor de t tabelado t_{tab} considerando distribuição t -Student com $n - 1$ graus de liberdade e nível de significância $\alpha = 0,05$. Se o valor de t -calculado for menor que t -tabelado nessa configuração, não há evidências estatísticas para rejeição da hipótese nula H_0 , indicando que não há evidências de erros sistemáticos nos resíduos, ao nível de

significância adotado. No sentido oposto, t -calculado maior que t -tabelado, há evidências para rejeição de H_0 , implicando admissão de presença de erros sistemáticos nos resíduos. ^{17, 89-91}

3.3.6.2 Avaliação de erros de tendência pelo teste não paramétrico

O método para detectar presença de tendências nos resíduos é o teste não-paramétrico de permutação. Em testes paramétricos de permutação há o princípio de distribuição conhecida dos dados do modelo. Nos testes não-paramétricos não há suposição prévia da distribuição dos dados. ^{17,89-92}

O teste de permutação não-paramétrico é baseado na aleatoriedade nos valores de referência, y , sem alterar a ordem nos dados instrumentais, X .¹⁷ Este teste consiste em uma reordenação repetitiva de cada variável com valores de referência (y_i). Novos valores de resposta são obtidos, reordenando e mudando as posições de todos os n valores de y . Esses novos valores de resposta, y permutado, não devem ter associação com o resíduo previsto da calibração, ou se houver, que seja muito limitada.¹⁷ O número possível de permutações aumenta com o número de elementos, com número de combinações dado por $N!$, sendo N o número de elementos. ^{17,89-92}

Inicialmente são definidas as hipóteses testadas para os resíduos não relacionados aos valores de referência, ou relacionados conforme a Equação 3.25.

H_0 - Hipótese nula: resíduos de previsão e_i são independentes de y_i ;¹⁷

H_1 - Hipótese alternativa: resíduos de previsão e_i são relacionados (dependentes) de y_i ;¹⁷

$$e_i = g(y_i) + \varepsilon_i \quad (3.25)$$

Onde ε_i é um erro aleatório independente e $g(y_i)$ é a função polinomial que poderá modelar a relação entre os resíduos e os valores de referência. Por sua vez, $g(y_i)$ segue a forma abaixo (Equação 3,26).¹⁷

$$g(y_i) = b_n y_i^n + b_{n-1} y_i^{n-1} + \dots + b_i y_i + b_0 \quad (3.26)$$

Onde b_n é o coeficiente polinomial de ordem n . Na hipótese alternativa em que se propõe a dependência dos resíduos com os valores de referência, admite-se que o efeito aleatório é decorrente do erro aleatório independente ε_i . Se o coeficiente polinomial de maior ordem b_n (Equação 3.26) for estatisticamente significativo ao nível de significância adotado existe evidências estatísticas significativas de tendência nos resíduos.¹⁷

O resultado do teste é avaliado pelo p-valor obtido. O p-valor é determinado pela proporção de vezes em que o coeficiente polinomial de maior ordem calculado para os dados originais b_n for maior que o coeficiente polinomial de maior ordem calculado para os dados permutados b'_n . Se o p-valor for menor que o nível de significância adotado, há evidências estatísticas significativas para rejeitar a hipótese nula H_0 , admitindo-se presença de tendência significativa nos resíduos. Em contrapartida, se p-valor for maior que nível de significância, H_0 é aceita, admitindo-se aleatoriedade nos resíduos, ou seja, ausência de erros tendenciosos.¹⁷

3.3.7 Fusão de dados

A combinação de dados ou informações oriundas de várias fontes é denominada fusão de dados. Essa integração é realizada geralmente com a finalidade de produção de modelos de classificação ou de regressão para estimativas de parâmetros.⁹⁵⁻⁹⁶

A fusão de dados pode fornecer a possibilidade de inferências estatísticas com maior precisão em relação aos modelos desenvolvidos com os conjuntos individuais. Geralmente a fusão de dados gera resultados satisfatórios quando os modelos construídos com dados fundidos atendem ao critério de maior complexidade e especificidade, quando comparados com os modelos obtidos com os dados individuais.⁹⁵⁻⁹⁹

A fusão de dados é uma área de ampla multidisciplinaridade. Os métodos geralmente seguem o preceito de relações possíveis entre diferentes fontes de dados. Estas relações podem ter caráter de complementaridade, redundância ou cooperatividade. Existem várias formas de classificação das técnicas de fusão de dados, entretanto adota-se neste trabalho uma classificação baseada no nível de abstração dos dados, em que podem ser: fusão de nível baixo; fusão de nível médio; e fusão de nível alto.⁹⁵⁻⁹⁹

A fusão de dados de nível baixo consiste na concatenação das matrizes de dados, previamente pré-processadas ou não. A fusão de nível médio consiste na extração de vetores informativos das matrizes de dados individuais antes da fusão, para a realização de outras tarefas após a fusão. A fusão de nível alto é comumente chamada de fusão de informações ou de decisão. Este tipo de fusão combina resultados obtidos em modelos de regressão ou de classificação, para obter uma decisão final mais precisa.⁹⁵⁻⁹⁹

Os modelos originais devem representar da melhor forma possível os dados individuais para garantir melhora na exatidão do modelo após a fusão dos dados. Uma outra classificação engloba a combinação de dados com diferentes níveis de abstração, denominada fusão de múltiplos níveis.⁹⁵⁻⁹⁹

Embora seja discreto o número de trabalhos na literatura com a fusão de dados em química, algumas aplicações são encontradas com finalidade de melhoria de modelos de classificação em diversas áreas, sendo que em análise de petróleo, a fusão de dados tem sido aplicada para fornecimento de modelos mais robustos e de melhor exatidão, com dados de diferentes fontes instrumentais, ou de mesma fonte.⁹⁵⁻

4 METODOLOGIA

O presente trabalho foi desenvolvido com dados cromatográficos e medições de propriedades físico-químicas de um total de 128 amostras de petróleo bruto da bacia sedimentar brasileira de densidade API variando de óleo leve (54,0 °API) a óleo pesado (17,0 °API).⁵⁹

A partir dos cromatogramas por cromatografia gasosa de alta temperatura (HTGC) e análise detalhada de hidrocarbonetos (DHA), diversos modelos de calibração multivariada por PLS foram desenvolvidos para determinação das propriedades ponto de fulgor e pressão de vapor Reid. Os modelos de maior eficiência preditiva foram selecionados e estão descritos neste trabalho. Os parâmetros de °API, PF e PVR das amostras de petróleo utilizadas neste trabalho, bem como a indicação dos ensaios realizados podem ser vistos na Tabela A1 (Anexos).

A fusão de dados foi utilizada na construção de modelos de regressão pelo PLS e para fornecimento de curva de destilação simulada com maior informação de componentes voláteis, compensando uma limitação da destilação simulada pela HTGC na região inicial da curva PEV.

4.1 Análise por cromatografia gasosa

Das 128 amostras iniciais, 103 foram analisadas pelo método HTGC e 67 pelo método DHA. Para a fusão de dados utilizou-se as informações de 42 amostras que apresentavam ambos os ensaios cromatográficos.

As análises cromatográficas DHA foram realizadas pelo Centro de Pesquisa e Desenvolvimento da Petrobras (CENPES), localizado no Rio de Janeiro, Brasil. A descrição da análise é apresentada neste trabalho para melhor compreensão dos dados obtidos.

O ensaio de DHA seguiu o método padrão ASTM D7900⁷⁴. Neste método o cromatógrafo é equipado com pré-coluna DB-5, 1m; coluna analítica HP-1, 100 m x 0,25 mm, revestida de filme polidimetilsiloxano 0,5 µm; válvula para reversão de fluxo, responsável pelo descarte da fração *n*-C15⁺; e detector por ionização em chama (FID). As condições cromatográficas em que se realizou o ensaio foram: gás carreador Hélio; temperatura da pré-coluna de 200 °C; temperatura do injetor e do detector de 300 °C; programação da temperatura do forno: 35 °C a 60 °C, 1 °C.min⁻¹, isoterma de 20 min., 60 °C a 200 °C, 2 °C.min⁻¹, isoterma de 20 min.

Foi adicionado padrão interno, 3,3-dimetilbuteno, no preparo da amostra, com concentração de 3% em massa. O padrão interno faz parte de um método de alinhamento do sistema acoplado ao equipamento, como a referência para alinhamento dos cromatogramas de diferentes amostras, adequando a reprodutibilidade do método.⁶⁸ O ensaio não utiliza réplicas. A alíquota da mistura foi injetada na entrada da pré-coluna, que está em série com a coluna analítica. No momento determinado previamente, pela calibração com padrão de *n*-parafinas para separação das frações, a válvula foi alternada para reversão do fluxo e a fração de interesse havia passado para coluna analítica. O tempo total da análise geralmente é estimado em 140 minutos.

As análises cromatográficas pela técnica HTGC foram realizadas no Núcleo de Competências em Química do Petróleo (NCQP-UFES), em cromatógrafo do fabricante Agilent Technologies, modelo 6890N.

O ensaio pela HTGC seguiu a ASTM D7169⁷⁶ com extensão de calibração (C5-C120) de parafinas normais. As amostras de petróleo foram analisadas em quadruplicata em um cromatógrafo com coluna capilar metálica revestida internamente com polidimetilsiloxano, 5 m x 0,53 mm de diâmetro interno, 0,09 - 0,15 µm de espessura de fase estacionária e detector de ionização por chama. As condições cromatográficas em que se realizou o ensaio foram: temperatura inicial do forno a -20 °C, programação de 10 °C.min⁻¹ a 430 °C, isoterma por 2 min; temperatura do injetor a 430 °C e do detector a 435 °C; gás de arraste He, com fluxo de 20 mL.min⁻¹; ar sintético e H₂ para alimentar a chama no detector.

O tempo de análise foi de aproximadamente 47 minutos, com utilização de 1 mL de amostra para cada réplica. O *software* utilizado para o tratamento do cromatograma foi o *GC ChemStation* Ver.A.10.02 (1757) *Agilent Technologies*. Para processamento de dados utilizou-se a média das quadruplicatas dos cromatogramas para compor o conjunto de dados.

4.2 Medições das propriedades físico-químicas

Os ensaios dos parâmetros físico-químicos ponto de fulgor e pressão de vapor Reid foram realizados pelo Centro de Pesquisa e Desenvolvimento da Petrobras (CENPES), localizado no Rio de Janeiro, Brasil. Para melhor compreensão dos resultados obtidos, os ensaios para obtenção dos parâmetros são brevemente descritos a seguir.

4.2.1 Ponto de fulgor

O ponto de fulgor foi medido seguindo a norma ISO 13736⁶³. A faixa de abrangência para ponto de fulgor para as amostras de petróleo bruto variou de menor que -30,0 °C a 59,5 °C. Segundo a norma padrão, a repetitividade e a reprodutibilidade são 1,0 °C e 1,5 °C, respectivamente.⁶³

Para o ensaio de ponto de fulgor a amostra foi previamente resfriada e em seguida adicionada ao copo teste, acoplado ao equipamento (*Abel closed-cup method*), e aquecida sob agitação constante. Em pequenos intervalos de temperatura, entre 1,0 °C e 2,0 °C, uma centelha foi disparada através de uma abertura na tampa do copo teste. A menor temperatura em que a superfície do petróleo inflamou momentaneamente com a centelha de teste, foi registrada como ponto de fulgor, equivalente à pressão atmosférica ambiente.^{6,63}

4.2.2 Pressão de vapor Reid

A pressão de vapor Reid foi determinada seguindo a norma ASTM D323⁶⁵. Os valores medidos de pressão de vapor Reid pelo método padronizado estavam distribuídos em 10,1 kPa a 73,0 kPa. A repetitividade e a reprodutibilidade do método padrão ASTM D323^{65portalfindes} é de 3,2 kPa e 5,2 kPa, respectivamente, para amplitude de 35,0 a 100,0 kPa de pressão de vapor Reid.

Para o ensaio de pressão de vapor Reid a amostra foi previamente resfriada e mantida entre 0,0 e 1,0 °C e aquecida em uma câmara de vapor equipada com um manômetro, imersa em um banho termostático a 37,8 °C (60 °F). A pressão de vapor Reid foi registrada após consecutivas medidas no manômetro no qual o sistema apresentou estar em equilíbrio.

4.3 Modelagem dos dados

Todos os procedimentos de processamento de dados foram realizados no software MATLAB® (The Mathworks, Natick, USA). Inicialmente os cromatogramas foram alinhados pelo algoritmo Icoshift.⁸³ Para a realização do alinhamento, os cromatogramas foram configurados com 50 intervalos de espaços iguais pelo tempo de retenção e nesses intervalos os picos foram alinhados tendo como alvo um cromatograma médio. Em seguida, as variações na linha de base foram removidas utilizando-se o algoritmo airPLS.⁸⁶

4.3.1 Determinação das propriedades físico-químicas

Os dados foram divididos em diferentes conjuntos para construção de modelos para determinação das propriedades físico-químicas a partir dos cromatogramas. Um fluxograma da elaboração dos conjuntos de dados com a quantidade de amostras disponíveis para cada técnica pode ser visto na Figura 4.1.

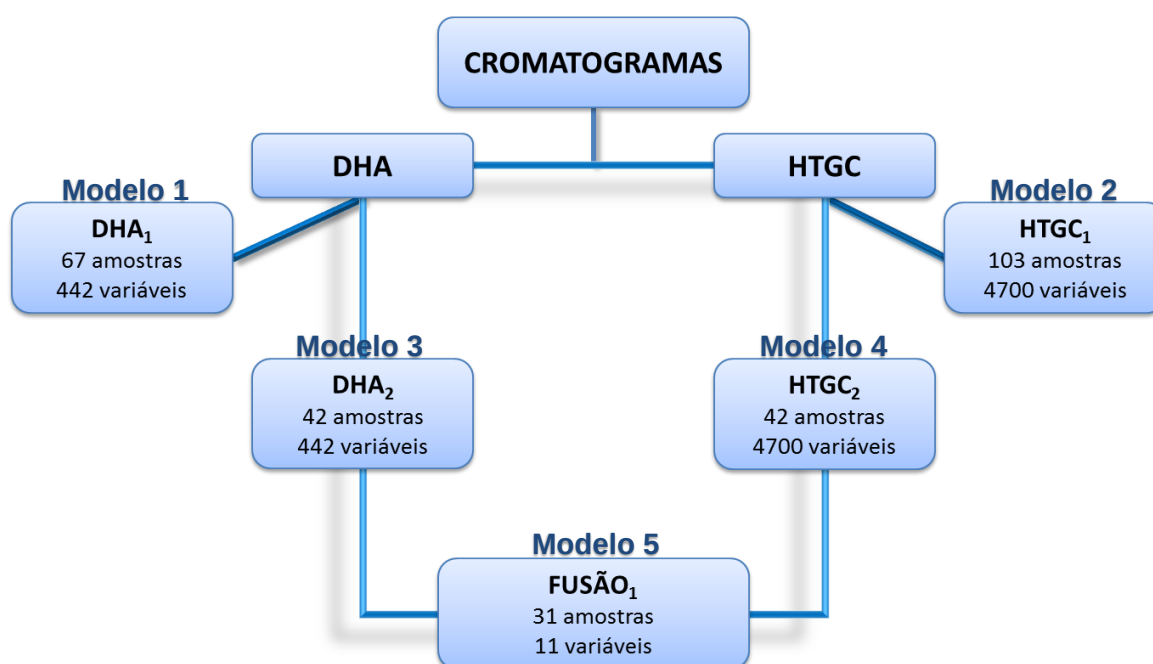


Figura 4.1 Fluxograma de elaboração dos conjuntos de dados instrumentais a partir dos cromatogramas DHA e HTGC para construção dos modelos pelo PLS.

O número de amostras variou para cada propriedade estimada devido adaptação ao número de amostras com as medições realizadas da referida propriedade.

Todos os modelos foram construídos seguindo-se os mesmos procedimentos. Antes da construção dos modelos de calibração multivariada, os conjuntos de dados passaram pelo pré-tratamento para alinhamento e remoção de linha de base. Entre as técnicas testadas as que se apresentaram adequadas aos dados foram: pré-tratamento pelo método SNV⁸⁸ e pré-processamento pela centragem dos dados na média. Em seguida, o conjunto de dados pré-processado foi separado em conjunto de calibração e conjunto de teste, na proporção de 70% para conjunto de calibração pelo método Kennard Stone¹⁰⁰ modificado, denominado "rank kS",¹⁰¹ utilizando na

separação informações da propriedade de interesse proporcionando separação com maior representatividade da amplitude dos valores da propriedade de interesse dentro do conjunto de calibração.

Então o número de variáveis latentes foi selecionado na etapa de otimização do modelo, pelo método validação cruzada *venetian blind 5-fold*. A partir disso os modelos foram construídos e selecionados.

4.3.1.1 Ponto de fulgor

Para ponto de fulgor foram construídos cinco modelos: 1 - DHA₁ com 66 amostras, utilizando dados de DHA; 2 - HTGC₁ com 51 amostras utilizando dados de HTGC; 3 - DHA₂ e 4 - HTGC₂ com 42 amostras, para as amostras analisadas em ambas as técnicas cromatográficas, seguindo-se à seleção do melhor modelo e fusão de nível alto para desenvolvimento do modelo 5 - FUSÃO₁, com 31 amostras.

4.3.1.2 Pressão de vapor Reid

Para determinação de pressão de vapor Reid foram desenvolvidos os modelos: 1 - DHA₁, com 50 amostras com dados de DHA; 2 - HTGC₁ com 84 amostras, com dados de HTGC; 3 - DHA₂ e 4 - HTGC₂ com 42 amostras, analisadas em ambas as técnicas, seguindo-se à seleção dos modelos para construção do modelo 5 - FUSÃO₁, com 31 amostras.

Os modelos de fusão de dados, tanto para a predição de ponto de fulgor quanto para pressão de vapor Reid foram desenvolvidos pela combinação de informações de modelos individuais com dados de DHA e de HTGC para obtenção de um resultado final.

4.3.1.3 Avaliação dos modelos PLS

Todos os modelos foram avaliados quanto à exatidão utilizando as métricas previamente descritas, como erro quadrático médio dos conjuntos de calibração e de previsão, RMSEC e RMSEP, utilizando as equações 3.19 e 3.20, respectivamente.

A avaliação da qualidade de linearidade dos modelos foi realizada pelo coeficiente de determinação R^2 (Equação 3.21), tanto para calibração (R^2_c), quanto para previsão (R^2_p).

Todos os modelos foram avaliados por meio de testes estatísticos para verificação de presença de erros sistemáticos, pelo teste de bias e erros tendenciosos nos resíduos de previsão, pelo teste de permutação não-paramétrico.¹⁷

Os modelos foram comparados quanto à acurácia para avaliar se houve melhora na exatidão com a construção de modelos pela fusão de dados. O Teste F foi aplicado para verificar a igualdade entre as acurácias ao nível de significância de 0,05, de acordo com a Equação 4.1.^{5,17}

$$F = \frac{RMSEP_1^2}{RMSEP_2^2} \quad (4.1)$$

Onde $RMSEP_2$ deve ser equivalente ao modelo supostamente de maior exatidão, assim, $RMSEP_1 > RMSEP_2$. Quando o p-valor do Teste F for menor que o nível de confiança adotado, a hipótese de igualdade entre as acurácias é rejeitada, indicando que há uma diferença significativa entre as acurácias dos modelos avaliados.

Para avaliação da significância química dos modelos, as variáveis de maior peso em cada variável latente formada, foram identificadas por meio dos *loadings*. Dessa forma, foram detectadas e identificadas as variáveis de maior influência para a variação nos modelos, uma vez que cada variável latente aproxima melhor os dados instrumentais à propriedade de interesse.

Adicionalmente, os resultados dos modelos desenvolvidos foram analisados quanto à significância estatística, pelo Teste F , de acordo com ASTM E1655⁹⁰, e pela possibilidade de aleatoriedade de seus resultados, pelo teste de permutação da propriedade de interesse.^{17,93} O teste de permutação é baseado na aleatoriedade nos valores de referência, y , sem alterar a ordem nos dados instrumentais, X .¹⁷ Novas avaliações da acurácia do modelo pela métrica RMSEP (Equação 3.20), são realizadas para cada reordenação de y . A significância estatística do modelo é avaliada considerando distribuição *t-Student* com $n - 1$ graus de liberdade e nível de significância $\alpha = 0,05$. Se o RMSEP, obtido pelo modelo sendo avaliado, estiver dentro da distribuição dos RMSEP dos modelos permutados, a hipótese de aleatoriedade do modelo é aceita. No sentido inverso, se o valor de RMSEP do modelo

avaliado, posicionar-se fora da distribuição dos RMSEP permutados, a hipótese de aleatoriedade é rejeitada. Para cada teste, foram realizadas 250 permutações.^{17, 89-93}

Em resumo, o quadro da Figura 4.2, mostra as etapas de processamento dos dados para construção dos modelos preditivos, assim como configurações e ferramentas utilizadas para cada finalidade.

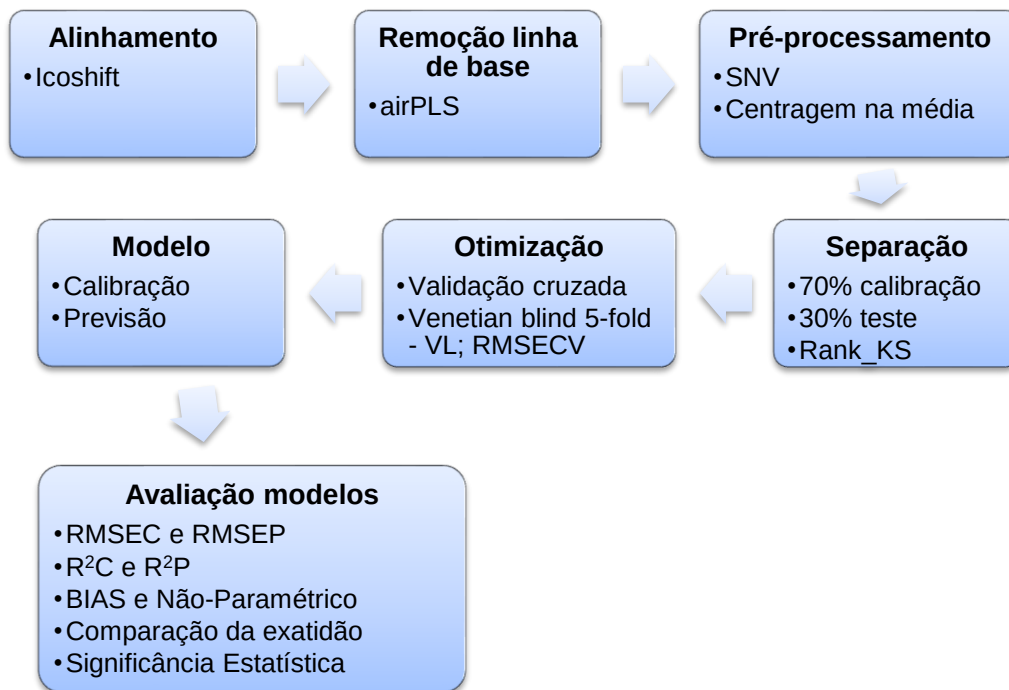


Figura 4.2 Diagrama de etapas para construção de modelos de calibração multivariada.

4.4 Fusão de dados para obtenção de curva de destilação simulada

A fusão dos dados foi realizada para combinação das curvas de distribuição de pontos de ebulição e recuperação em massa, obtida pela técnica DHA e destilação simulada (SIMDIS) pela técnica HTGC, seguindo-se a norma ASTM D7169⁷⁶.

A mesclagem das curvas foi realizada em cada amostra. Para isso cada curva obtida por ambas as técnicas foi avaliada graficamente para identificação do ponto de junção, também chamado temperatura de junção (TJ) (Figura 4.3)

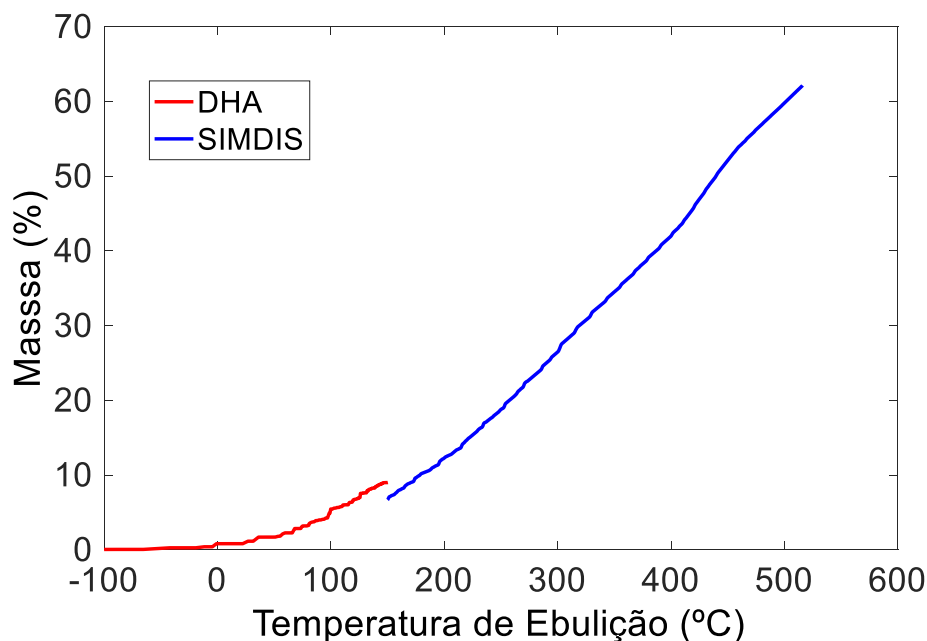


Figura 4.3 Aspecto visual de curvas de destilação simulada DHA e SIMDIS para identificação de ponto de junção, correções e mesclagem de dados

O ponto de junção entre as curvas foi definido em T10%, que compreende a temperatura que ocorre eluição de aproximadamente 10% de rendimento em massa e eluição de *n*-parafina até C9 (nonano). Nesta TJ foram calculadas as correções necessárias, uma vez que a fusão requer que os dados estejam em uma mesma magnitude. Dessa forma, as equações abaixo descrevem a correção realizada para a curva SIMDIS (Equação 4.2 e 4.3).

Para temperaturas (T) até ponto de junção (TJ):

$$\%SIMDIS'_T = \%DHA_T \quad (4.2)$$

Para temperaturas acima ponto de junção (TJ):

$$\%SIMDIS'_T = (\%DHA_{TJ} - \%SIMDIS_{TJ}) + \%SIMDIS_T \quad (4.3)$$

Para construção da curva completa, foi utilizado interpolação para estimar os valores da função no ponto do intervalo gerado entre as extremidades das curvas DHA e SIMDIS. Os valores calculados das extremidades da curva de DHA e da curva SIMDIS no ponto de junção foram utilizados como termos de entrada da função polinomial de interpolação. A curva final obtida foi suavizada por meio de equações polinomiais, cuja ordem foi selecionada de acordo com o ajuste necessário de cada

curva, variando entre ordem 6 a 7. Uma representação gráfica dessas etapas de obtenção de curva de destilação simulada pela fusão das duas técnicas é apresentada na Figura 4.4.

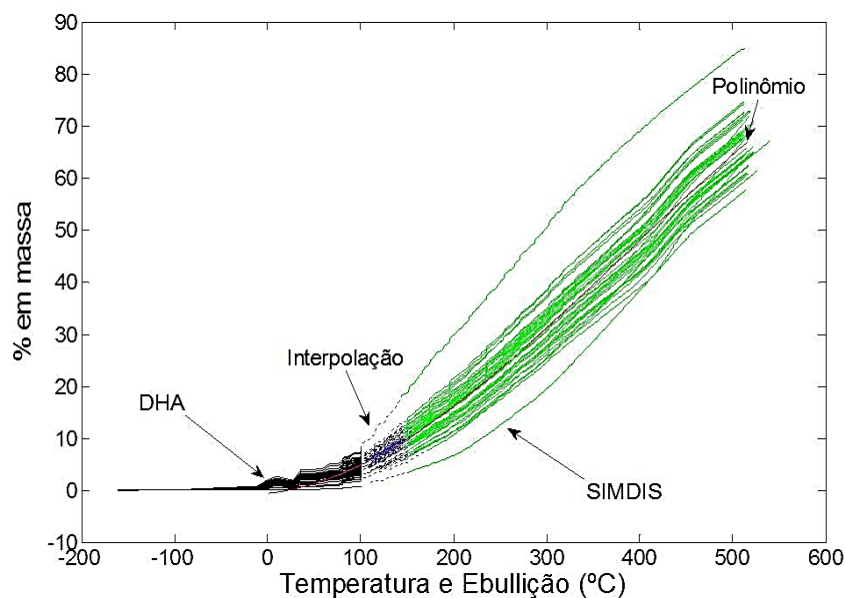


Figura 4.4 Aspecto gráfico de componentes para mesclagem de curvas DHA e SIMDIS.

As curvas estimadas pela SIMDIS e pela fusão (Y_p) de 5 amostras foram comparadas às curvas PEV(Y_R) para cálculo de Erro Relativo $ER(\%)$ de cada amostra (Equação 5).⁵²⁻⁵³

$$ER(\%) = \frac{|Y_p - Y_r|}{Y_R} * 100 \quad (4.4)$$

5 RESULTADOS E DISCUSSÕES

5.1 Perfil cromatográfico

O perfil cromatográfico de um óleo leve, um óleo médio e um óleo pesado estão representados na Figura 5.1 com os respectivos valores de densidade API, ponto de fulgor (PF) e pressão de vapor Reid (PVR).

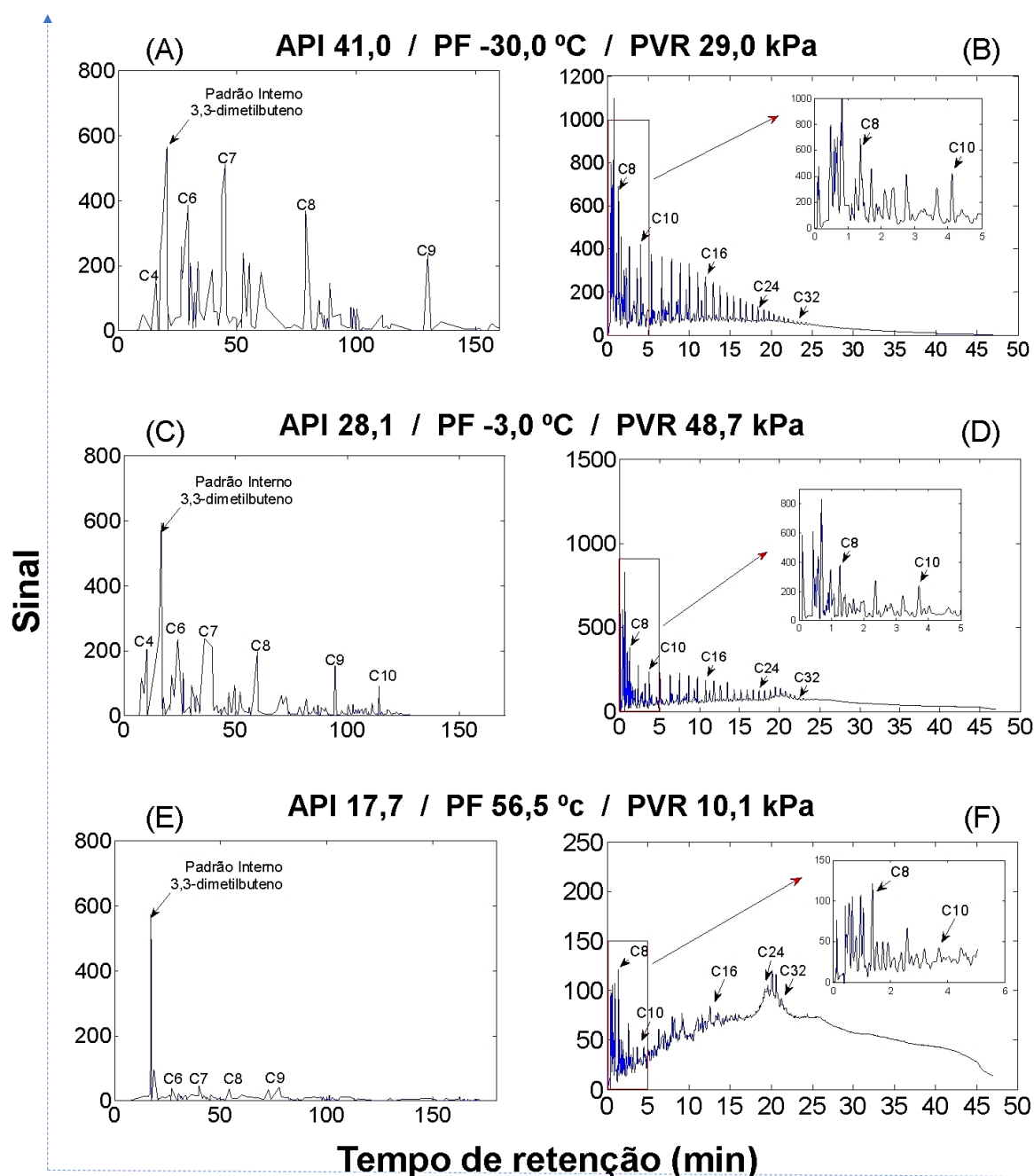


Figura 5.1 Cromatogramas de três amostras de petróleo pelas técnicas DHA e HTGC. (A) Cromatograma DHA, óleo leve; (B) Cromatograma HTGC, óleo leve; (C) Cromatograma DHA, óleo médio; (D) Cromatograma HTGC, óleo médio; (E) Cromatograma DHA, óleo pesado; (F) Cromatograma HTGC, óleo pesado.

Picos correspondendo aos *n*-alcanos são observados nos cromatogramas DHA (Figura 5.1 (A), (C) e (E)), os quais são descritos na literatura como mal resolvidos, nos primeiros cinco minutos, nos cromatogramas pelo HTGC, destacados na Figura 5.1 (B), (D) e (F).^{25-26,49}

O pico observado nos cromatogramas DHA (Figura 5.1 (A), (C) e (E)), em aproximadamente 17 minutos, de mesma intensidade, corresponde ao padrão interno 3,3-dimetilbuteno, que foi utilizado para alinhamento dos cromatogramas entre as amostras. Os componentes *n*-alcanos apresentam-se predominantes no óleo leve (Figura 5.1 (A), e (B)) uma vez que este tipo de óleo apresenta maior quantidade de compostos voláteis. Consequentemente neste tipo de óleo há mais substâncias para evaporar, logo, havendo uma relação inversa com o valor de ponto de fulgor. Os cromatogramas de óleo médio (Figura 5.1 (C) e (D)) apresentam os mesmos picos, porém com considerável redução de intensidade, como esperado para este tipo de óleo. Observa-se no cromatograma DHA de óleo médio (Figura 5.1 (C)) alargamento de pico na região de 36 a 39 min, correspondendo aos isômeros *n*-heptano e metilciclohexano. Os cromatogramas de óleo pesado (Figura 5.1 (E) e (F)) apresentam ausência desses picos.

Pelos cromatogramas HTGC pode-se observar um efeito na linha de base a partir do óleo médio (Figura 5.1 (C) e (D)), devido ao aumento na resposta do detector de fluxo de massa dos componentes mais pesados.^{42,49} Este efeito é mais visível no cromatograma mais ampliado de óleo pesado (Figura 5.1 (F)).

Nos cromatogramas do óleo pesado (Figura 5.1 (E) e (F)) é observado um perfil de biodegradação do óleo pela diminuição da intensidade dos *n*-alcanos. Entretanto, na região de 20 minutos (Figura 5.1 (F)), há presença de picos entre *n*-C24 a *n*-C32, possivelmente correspondentes aos biomarcadores hopanos, que são compostos pentacíclicos (C30 e C31) que são mais resistentes à biodegradação do que *n*-alcanos.¹⁰²

5.2 Modelagem dos dados

Como etapa inicial de processamento dos dados para a construção de modelos de calibração multivariada, o alinhamento dos picos dos cromatogramas pelo icoshift⁸³ corrigiu o desalinhamento dos tempos de retenção. Um recorte com o efeito visual do alinhamento de um dos intervalos dos cromatogramas HTGC é apresentado na Figura 5.2.

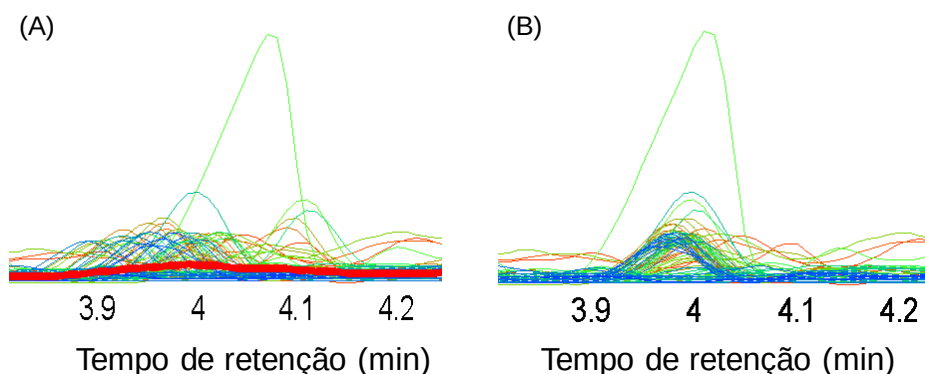


Figura 5.2 Efeito visual do alinhamento dos cromatogramas HTGC pelo Icosift. (A) Antes do alinhamento; (B) Após alinhamento.

Todos os conjuntos de dados foram testados em diversos métodos de pré-processamento, no entanto os que se apresentaram mais adequados aos dados foram SNV⁸⁸ e Centragem na média. Para cada grupo de dados foram separadas as amostras para o conjunto de calibração (70%) e de teste (30%) utilizando-se *rank_Ks*.¹⁰¹ Assim, as propriedades de ponto de fulgor e pressão de vapor Reid foram estimadas por meio da construção de modelos de regressão por PLS.

5.2.1 Estimativa de ponto de fulgor

A distribuição dos valores de referência de ponto de fulgor (°C) pela quantidade de amostras disponíveis nas duas técnicas cromatográficas pode ser vista pelos histogramas na Figura 5.3. Essa distribuição é uma representação da faixa de valores de ponto de fulgor do petróleo localizado na bacia sedimentar brasileira, uma vez que as amostras foram selecionadas para abranger essa representatividade.

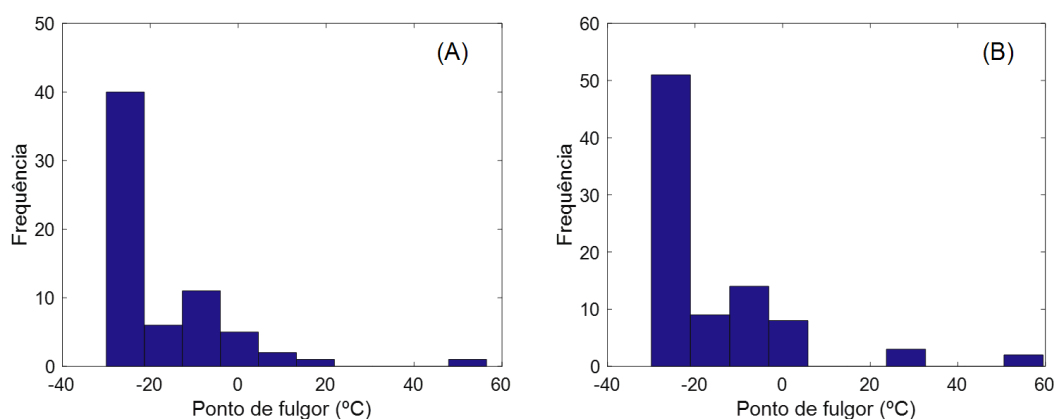


Figura 5.3 Histogramas com distribuição da propriedade ponto de fulgor medida pelo método padrão. (A) 66 amostras analisadas pelo DHA; (B) 87 amostras analisadas pelo HTGC.

Os modelos que apresentaram maior capacidade preditiva para estimar ponto de fulgor com os dados cromatográficos são apresentados na Tabela 5.1. Em todos os modelos descritos não há evidências de erros sistemáticos (teste bias) e tendenciosos nos resíduos (teste não paramétrico de permutação) ao nível de significância de 5%.

Tabela 5.1 Principais parâmetros estatísticos dos modelos PLS para predição de ponto de fulgor

Parâmetros/Dados	1 DHA ₁	2 HTGC ₁	3 DHA ₂	4 HTGC ₂	5 FUSÃO ₁
Pré-processamento	SNV	SNV	SNV	Cent.	Cent.
Amostras	66	51	42	42	31
Conjunto de calibração	47	36	31	31	23
Conjunto de previsão	19	15	11	11	8
Min. Cal (°C)	<-30	<-30	<-30	<-30	<-30
Max. Cal (°C)	56.5	59.5	56.5	56.5	56.5
VL	2	4	4	7	7
RMSEC (°C)	11.3	9.0	7.3	4.1	3.5
RMSECV (°C)	12.6	17.6	14.4	16.1	6.1
RMSEP (°C)	8.0	12.4	11.6	17.2	5.3
R ² c	0.5820	0.8518	0.8684	0.9634	0.9797
R ² cv	0.4503	0.3998	0.4282	0.3665	0.9092
R ² p	0.5373	0.7393	0.8211	0.6950	0.8963
Bias Calibração ^a	1,0000	0,7934	1,0000	1,0000	0,8124
Bias Previsão ^b	0,2660	0,3372	0,0639	0,0532	0,0770
Permut. Calibração ^c	0.07	0.06	0.25	0.14	0.15
Permut. Previsão ^d	0.25	0.49	0.50	0.42	0.23

^ap-valor de teste de bias de conjunto de calibração.

^bp-valor de teste de bias conjunto de conjunto de previsão.

^cp-valor de teste de permutação não paramétrico para avaliar tendências nos resíduos do conjunto de calibração.

^dp-valor de teste de permutação não paramétrico para avaliar tendências nos resíduos do conjunto de previsão.

Entre os modelos desenvolvidos para dados de cada técnica isolada na predição do ponto de fulgor, o Modelo 3 (DHA₂) apresentou RMSEC de 7,3 °C e RMSEP de 11,6 °C, R²c de 0.8684 e R²p de 0.8211. Apesar de o Modelo 1 (DHA₁) ter apresentado menor RMSEP (8.0 °C) que o Modelo 3 (DHA₂), pelo Teste *F*, não se indicou diferença estatisticamente significativa entre RMSEP desses dois modelos (p-valor igual a 0,1488).

Os gráficos com as correlações entre valores preditos e de referência para ponto de fulgor são apresentados na Figura 5.4.

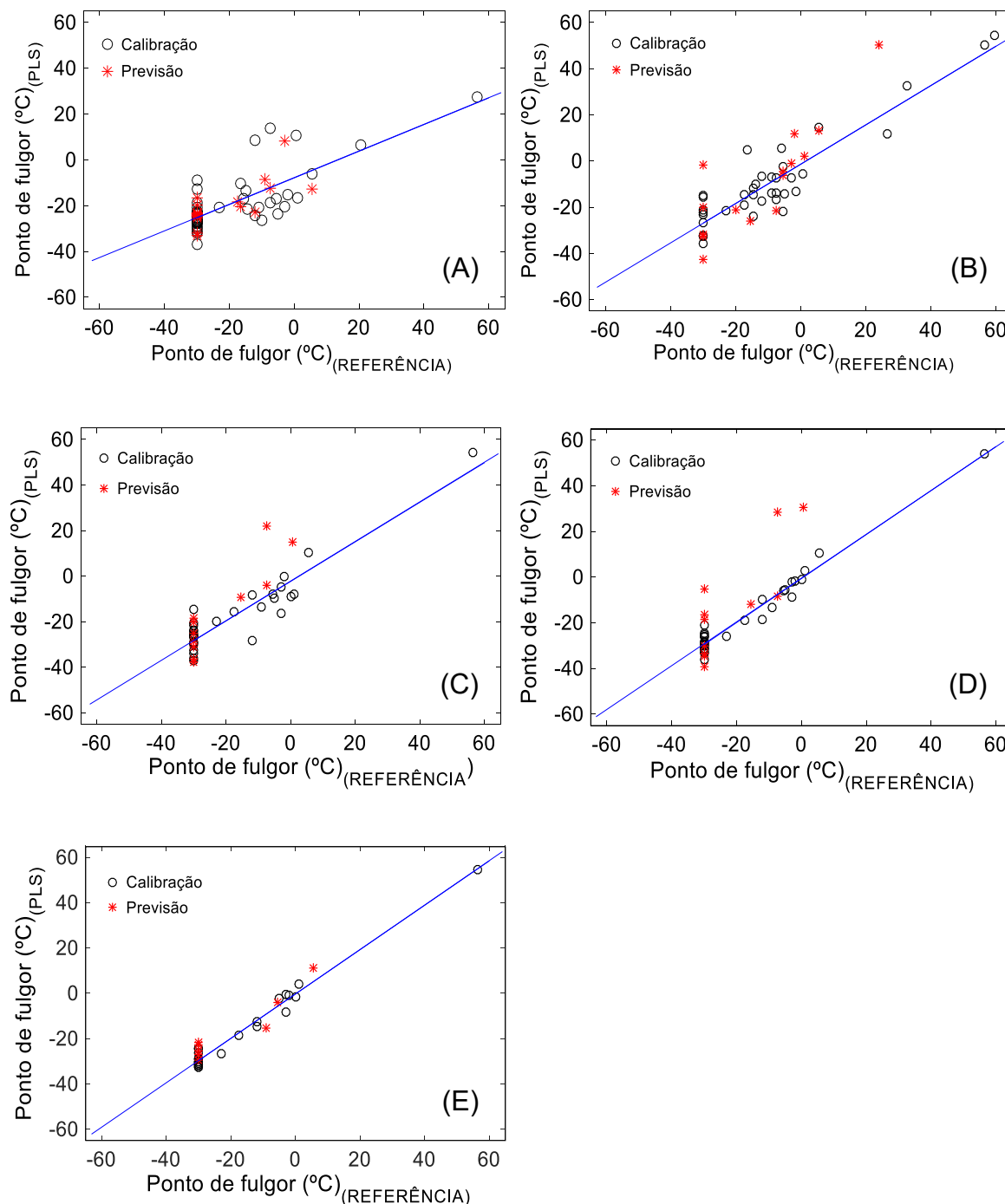


Figura 5.4 Gráficos de correlação entre valores de ponto de fulgor (°C) de referência versus previsto pelo modelo PLS. (A), Modelo 1, DHA₁; (B) Modelo 2, HTGC₁; (C) Modelo 3, DHA₂; (D) Modelo 4, HTGC₂; (E) Modelo 5, FUSÃO₁.

O modelo desenvolvido com fusão de dados (Modelo 5) demonstrou que houve ganho de informação, com RMSEC de 3.5 °C; RMSEP de 5.3; R^2c de 0,9797 e R^2p de 0,8963. O Teste F de comparação entre os modelos indicou uma melhora na acurácia do Modelo 5.

Na análise de significância da regressão, todos os modelos apresentaram valores altos no Teste F , que comparados aos valores de $F_{crítico, GL, \alpha=0,05}$ para os graus de liberdade e nível de confiança, indicaram regressão significativa.⁹⁰⁻⁹¹

Considerando que o modelo 5 apresentou melhora na exatidão, podemos afirmar que as informações contidas nos cromatogramas DHA e HTGC são complementares. Compreender as variáveis de maior influência nos modelos das duas técnicas pode esclarecer as contribuições de cada cromatograma na construção do modelo com fusão de dados. Por meio de visualização gráfica do cromatograma DHA, em função dos *loadings* do modelo PLS 3 (DHA₂), utilizando o índice de Kóvats⁷⁵ foram identificadas as variáveis de maiores *loadings* (Figura 5.5).

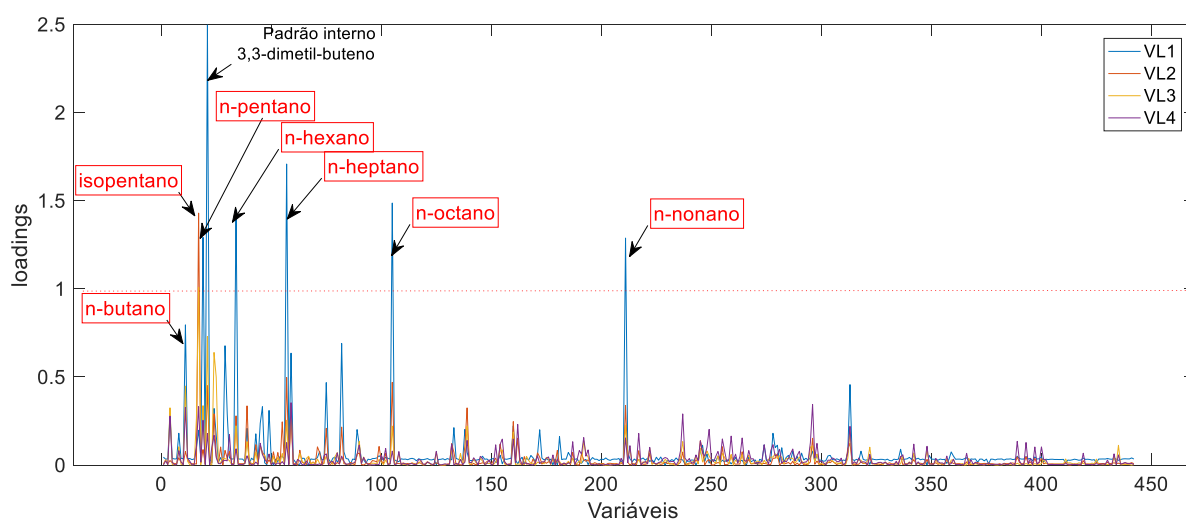


Figura 5.5 Representação do cromatograma pelo DHA, das variáveis latentes (VL), em função dos *loadings* do modelo 3 – DHA₂ na previsão de ponto de fulgor.

Podemos observar que as variáveis de maior contribuição para as VL com dados de DHA são os *n*-alcanos de menor peso molecular. Essas variáveis indicam que possuem maior influência na construção das variáveis latentes. Para avaliação das variáveis de maior influência pelo cromatograma HTGC (Figura 5.6), utilizou-se o tempo de retenção em função dos *loadings*, para identificação dos picos pelo tempo de retenção.

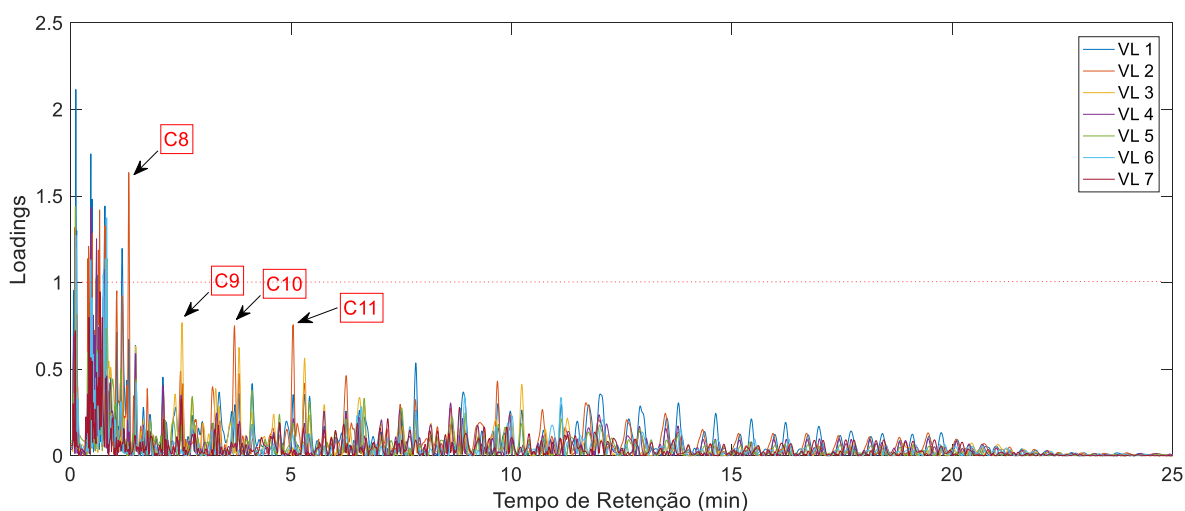


Figura 5.6 Representação do cromatograma pelo HTGC, das variáveis latentes (VL), em função dos *loadings* do modelo 4 – HTGC₂ na previsão de ponto de fulgor.

Considerando os *loadings* de maior valor, pelo cromatograma HTGC pode-se observar que as variáveis de maior influência estão localizadas antes dos 5 minutos iniciais do cromatograma, região correspondente aos *n*-alcanos de nC6 a nC10, indicando que estas variáveis possuem maiores influências na construção das variáveis latentes, para a propriedade ponto de fulgor. Havendo uma cooperatividade entre as informações contidas no cromatograma DHA e no HTGC, a soma de suas informações de maior relevância, para esta propriedade, formam as variáveis do modelo de fusão de dados (Modelo 5).

Os testes permutação para avaliação da significância estatística dos modelos desenvolvidos para ponto de fulgor demonstraram que o Modelo 5, pela fusão de dados, destacado como de melhor desempenho, juntamente com o Modelo 1, pelos dados de DHA, não são resultados aleatórios, como pode ser observado na Figura 5.7. Esta figura consiste nos histogramas com a distribuição do RMSEP gerado na avaliação dos modelos a partir da permutação de valor de Ponto de Fulgor (°C). Uma linha contínua vermelha indica a localização do valor de RMSEP do modelo sendo avaliado quanto à significância estatística. Se esta linha estiver localizada dentro do histograma, o RMSEP do modelo avaliado pertence à distribuição de RMSEPs aleatórios, no sentido inverso, linha vermelha fora da distribuição, o RMSEP do modelo sendo avaliado não pertence à distribuição aleatória, indicando haver significância estatística.

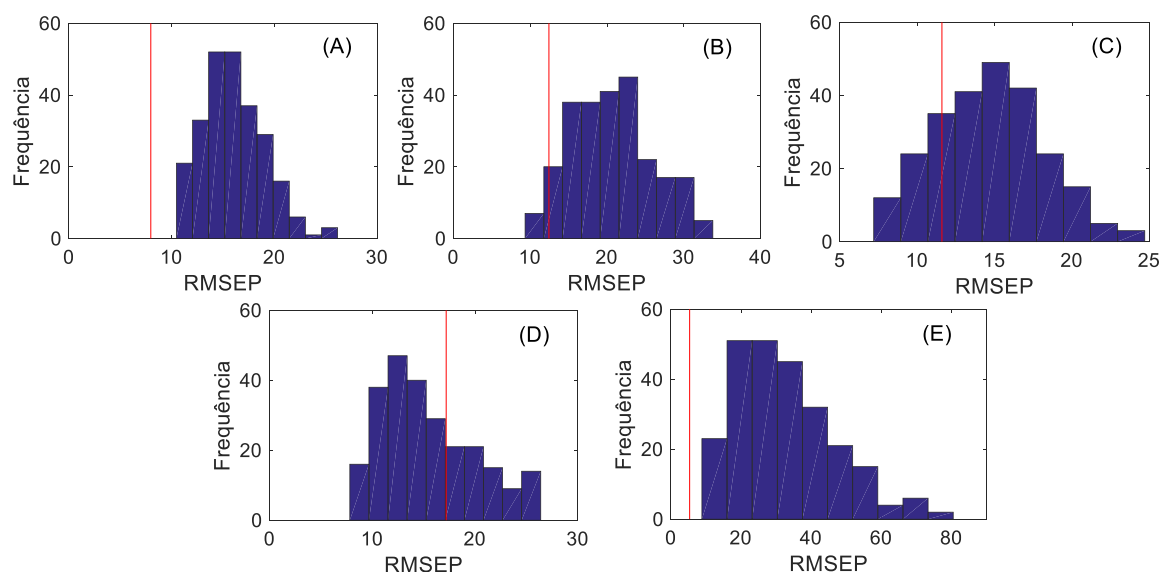


Figura 5.7 Resultados avaliação teste significância estatística Modelos preditivos PF. (A) Modelo 1 - DHA₁; (B) Modelo 2 - HTGC₁; (C) Modelo 3 - DHA₂; (D) Modelo 4 - HTGC₂; (E) Modelo 5 - FUSÃO₁

Dentre os modelos desenvolvidos para determinação de ponto de fulgor, apresentaram resultados aleatórios, indicando não haver significância estatística, os modelos: 2 - HTGC₁ (Figura 5.7(B)); 3 - DHA₂ (Figura 5.7(C)) e 4 - HTGC₂ (Figura 5.7(D)). Como pode ser observado, todos estes modelos que apresentaram resultado indicativo de aleatoriedade, apresentaram RMSEP acima de 10,0 °C, enquanto que os dois modelos que não foram indicados como aleatórios (Modelo 1 e Modelo 5), apresentaram RMSEP de 8,0 °C e 5,3 °C, respectivamente.

Os resultados obtidos para ponto de fulgor em petróleo bruto brasileiro, que geralmente estão entre os valores de -30,0 °C a 59,5 °C, apresentaram valores aceitáveis quando comparados com resultados reportados na literatura para derivados de petróleo. Vários trabalhos são reportados com determinação de diesel, cuja faixa de amplitude é menor em comparação ao petróleo bruto. O diesel geralmente apresenta valores de ponto de fulgor entre 24,5 °C a 76,5 °C. Dai et al (2015)¹⁰ utilizaram modelos de contribuição de grupos moleculares e regressão linear múltipla para previsão de ponto de fulgor em misturas de hidrocarboneto, obtendo RMSEP de 10,5 K e R²p de 0,8730. Alves et al. (2012)¹⁵ reportaram resultados de estimativa de ponto de fulgor em diesel com RMSEP igual a 3,8 °C e R²p de 0,698 pelo PLS e RMSEP de 2,0 °C com R²p de 0,936 pelo modelo de regressão por vetor de suporte. Figueiras et al (2014)¹⁷ descreveram modelos preditivos para óleo diesel com RMSEP de 3,7 °C e R²p de 0.844.

5.2.2 Estimativa pressão de vapor Reid

A distribuição dos valores de referência de pressão de vapor Reid (kPa) pela quantidade de amostras disponíveis das duas técnicas cromatográficas pode ser vista pelos histogramas na Figura 5.8. Pela representatividade dessas amostras, esta distribuição expressa uma faixa, geralmente encontrada, de pressão de vapor do petróleo brasileiro.

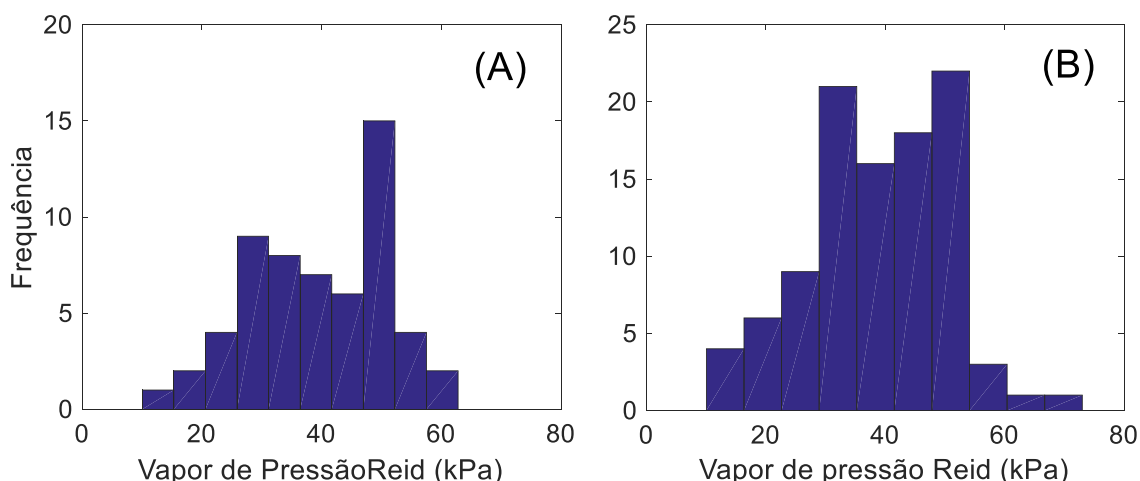


Figura 5.8 Histogramas com distribuição da propriedade pressão de vapor Reid medida pelo método padrão. (A) 58 amostras analisadas por DHA; (B) 101 amostras analisadas pelo HTGC.

Os resultados obtidos na estimativa de pressão de vapor Reid mostraram uma ótima relação com os valores de referência (Tabela 5.2 e Figura 5.6). O Modelo 2 (HTGC₁) apresentou R^2_c de 0,9954 e R^2_p de 0,9976. Este modelo apresentou melhor acurácia com RMSEC de 0,6 kPa e RMSEP de 0,4 kPa, abaixo dos valores de repetitividade e reprodutibilidade do método padrão⁶⁵ (3,2 kPa e 5,2 kPa, respectivamente). O Teste F indicou que há diferença estatisticamente significativa na acurácia do Modelo 2 (HTGC₁), uma vez que apresentou diferenças significativas entre acurácia deste com todos os demais modelos. Entretanto, o Teste F indicou não haver diferença estatisticamente significativa entre Modelo 1 (DHA₁), Modelo 3 (DHA₂), Modelo 4 (HTGC₂) e Modelo 5 (FUSÃO₁) ao nível de significância de 5%.

Tabela 5.2 Principais parâmetros dos modelos PLS para previsão de pressão de vapor Reid.

Parâmetros/Dados	1	2	3	4	5
	DHA ₁	HTGC ₁	DHA ₂	HTGC ₂	FUSÃO ₁
Pré-processamento	Cent.	Cent.	SNV	Cent.	Cent.
Amostras	50	84	39	39	28
Conjunto de calibração	36	60	28	28	20
Conjunto de previsão	14	24	11	11	8
Mín. Calibração (kPa)	10.0	23.2	20.2	20.2	20.2
Max. Calibração (kPa)	62.8	51.0	53.3	53.3	53.3
VL	8	4	6	5	2
RMSEC (kPa)	3.7	0.6	3.3	5.5	3.5
RMSECV (kPa)	11.0	0.6	9.1	9.7	4.7
RMSEP (kPa)	6.8	0.4	8.1	7.4	3.6
R ² c	0.9285	0.9954	0.9057	0.7303	0.8851
R ² cv	0.2783	0.9945	0.2154	0.1615	0.7588
R ² p	0.7581	0.9976	0.6989	0.4616	0.8933
Bias Calibração ^a	0,9998	0,9737	0,8121	1,0000	0,9487
Bias Previsão ^b	0,0538	0,9564	0,0348	0,8047	0,6459
Permutação Calibração ^c	0.4157	0.3455	0.0530	0.4500	0.4441
Permutação Previsão ^d	0.3923	0.0835	0.5059	0.2151	0.0874

^ap-valor de teste de bias de conjunto de calibração.

^bp-valor de teste de bias conjunto de conjunto de previsão.

^cp-valor de teste de permutação não paramétrico para avaliar tendências nos resíduos do conjunto de calibração.

^dp-valor de teste de permutação não paramétrico para avaliar tendências nos resíduos do conjunto de previsão.

A redução no número de amostras (modelos DHA₂ e HTGC₂) implicou redução da capacidade preditiva dos modelos com dados individuais (DHA₁ e HTGC₁), sendo mais visível no modelo com dados de HTGC. Porém, o Teste *F* indicou diferença significativa entre acurácia dos modelos 2 (HTGC₁) e 4 (HTGC₂), mas indicou não haver diferença significativa entre acurácias dos modelos 1 (DHA₁) e 3 (DHA₂).

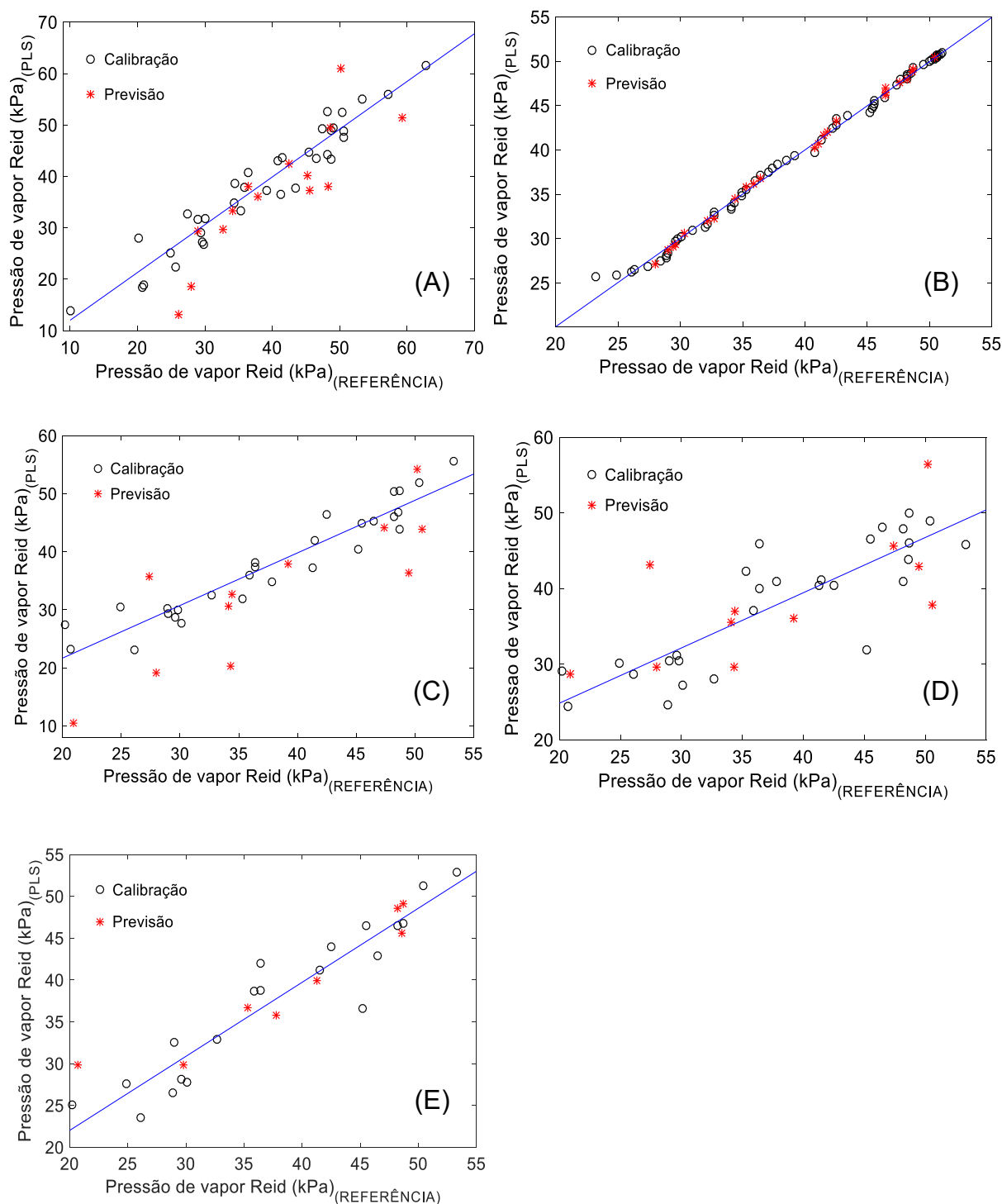


Figura 5.9 Gráficos de correlação entre valores de pressão de vapor Reid (kPa) de referência versus previsto pelo modelo PLS. (A), Modelo 1, DHA₁; (B) Modelo 2, HTGC₁; (C) Modelo 3, DHA₂; (D) Modelo 4, HTGC₂; (E) Modelo 5, FUSÃO₁.

A fusão de dados não resultou ganho de informação na capacidade preditiva, mas reduziu os erros dos modelos com menor número de amostras. Neste modelo com fusão de dados foi obtido RMSEC de 3,5 kPa e RMSEP de 3,6 kPa. O Teste *F* indicou diferença significativa entre a acurácia do modelo FUSÃO₁ e modelo HTGC₂

(p-valor de 0.034). Entretanto, também houve diferença significativa quando comparados os RMSEPs dos modelos FUSÃO₁ e HTGC₁, indicando que o modelo HTGC₁ possui maior acurácia. Portanto, com a fusão de dados não houve melhora na acurácia, mas uma redução de 44% e 48% no RMSEP dos modelos construídos, modelos 3 (DHA₂) e 4 (HTGC₂), respectivamente.

Os resultados indicam que os dados de HTGC apresentam melhor capacidade preditiva para a determinação de pressão de vapor Reid em petróleo. A limitação da técnica na região de componentes mais voláteis não interferiu nos resultados com abrangência de vapor de pressão Reid de 23,2 kPa a 51 kPa. Isto é possivelmente explicado pela relação direta da pressão de vapor como princípio da cromatografia gasosa. Estes resultados encontrados estão de acordo com resultados relatados na literatura para estimativa de PVR de gasolina utilizando dados de destilação pelo PLS. Mendes, Aleme e Barbeira (2017)¹⁹ utilizaram amostras de gasolina entre 49,1 kPa a 67,4 kPa e obtiveram resultados com R^2_p de cerca de 0,97. Devendo-se destacar que pela maior complexidade da matriz, petróleo bruto, os resultados deste trabalho demonstram serem mais expressivos frente a derivados e frações.

Todos os modelos foram avaliados estatisticamente e não há evidências de erros sistemáticos (teste bias) e tendência nos resíduos (teste de permutação não paramétrico) ao nível de significância de 5%. Quando avaliados quanto à significância de sua regressão, todos os modelos apresentaram valores altos no Teste F , que comparados aos valores de $F_{crítico, GL, \alpha=0,05}$ para os graus de liberdade e nível de confiança, indicaram regressão significativa.⁹⁰⁻⁹¹

Quanto à significância química, o modelo que apresentou melhor desempenho, Modelo 2 (HTGC₁), foi avaliado pelos *loadings* para identificar as variáveis de maior influência na construção das 4 variáveis latentes do modelo. De acordo com a Figura 5.10, a série homóloga de *n*-alcanos foi a principal fonte de variação do modelo para previsão de pressão de vapor Reid.

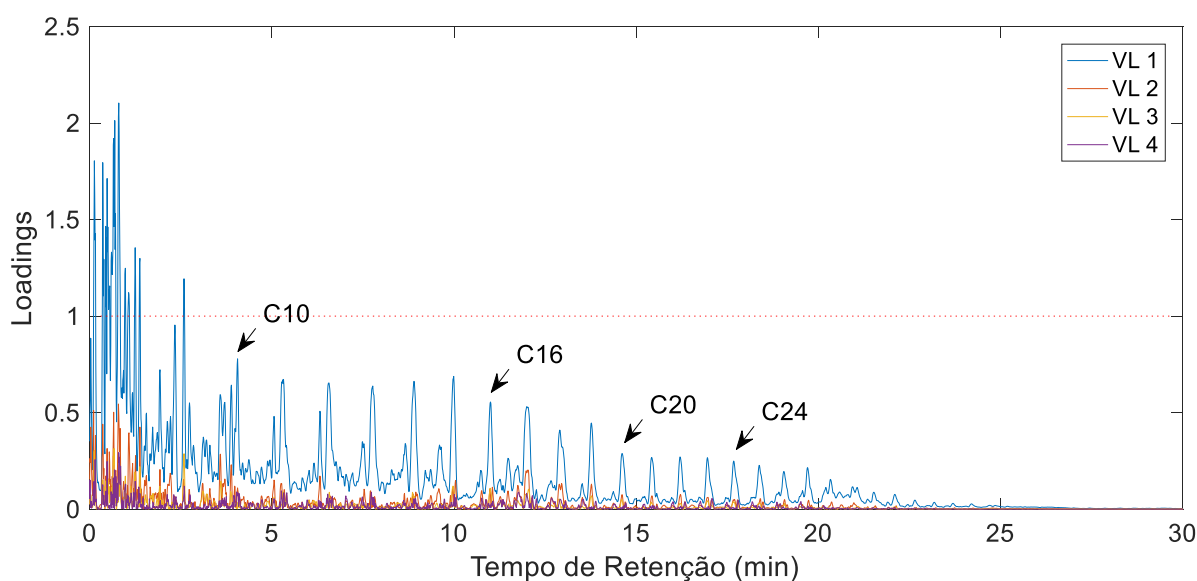


Figura 5.10 Representação do cromatograma pelo HTGC, das variáveis latentes (VL), em função dos *loadings* do modelo 2 - HTGC₁ na previsão de pressão de vapor Reid.

Outra observação importante a partir da Figura 5.10, é a região utilizada como fonte de variação para a previsão da pressão de vapor. O cromatograma pelo HTGC abrange até 47 minutos, porém, a fonte de variação utilizada na previsão da pressão de vapor Reid compreende até aproximadamente 25 minutos, considerando-se que a partir de C24 não há fonte de variação aproveitada na previsão desta propriedade. Este comportamento corrobora com o que seria esperado na construção do modelo, uma vez que a partir de C32, aproximadamente 25 minutos no cromatograma HTGC, ocorre a eluição da fração equivalente ao resíduo na destilação do petróleo bruto.⁷⁶

A avaliação da significância estatística dos modelos, que testou a hipótese de aleatoriedade de RMSEP de cada modelo desenvolvido para determinação de pressão de vapor Reid, indicou que o Modelo 2 (HTGC₁) não apresenta aleatoriedade de seu RMSEP. Os histogramas do teste de permutação podem ser observados na Figura 5.11.

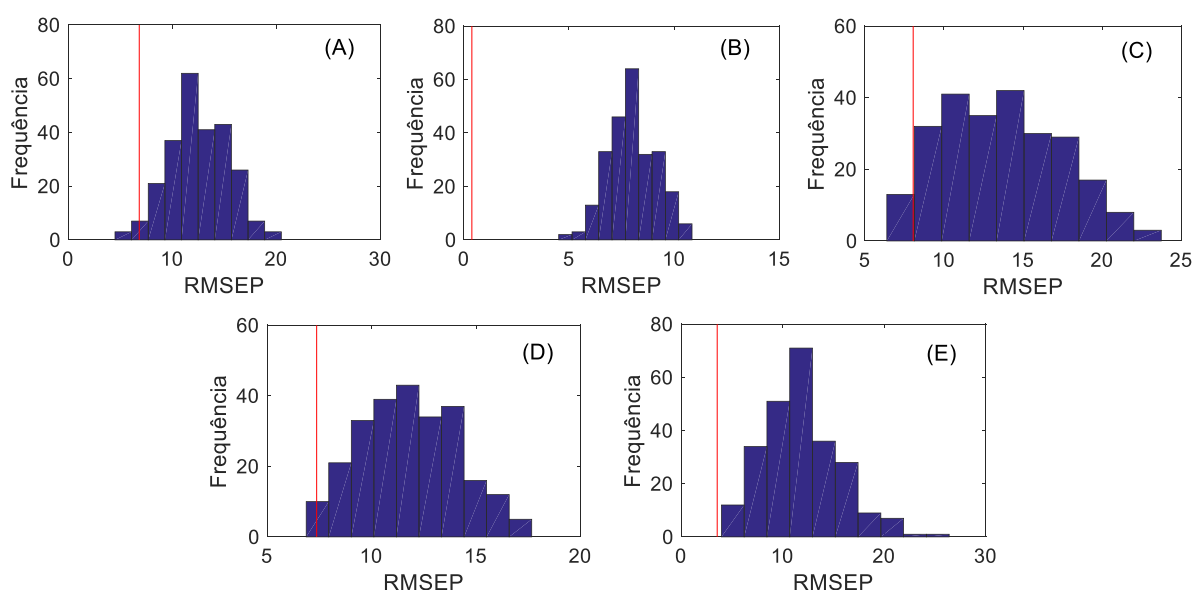


Figura 5.11 Resultados avaliação teste significância estatística Modelos preditivos PVR. (A) Modelo 1 - DHA₁; (B) Modelo 2 - HTGC₁; (C) Modelo 3 - DHA₂; (D) Modelo 4 - HTGC₂; (E) Modelo 5 - FUSÃO₁

Como pode ser observado, todos os modelos cujo RMSEP está dentro da distribuição aleatória, apresentaram RMSEP acima de 3,6 kPa, enquanto que os modelos que estão fora da distribuição de RMSEP aleatórios (Modelo 2 (HTGC₁) e Modelo 5 (FUSÃO₁)), apresentaram valores de 0,6 kPa e 3,6 kPa, respectivamente.

5.3 Fusão de dados para estimativa da curva PEV

A fusão de dados descrita acima compreendeu uma fusão de alto nível,⁹⁵ onde os resultados de um modelo de calibração foram unidos para promover um resultado final com maior precisão na estimativa das propriedades PF e PVR. Em contrapartida, para a estimativa da curva PEV pela destilação simulada, a fusão de dados está mais compreendida na classificação de fusão de médio nível,⁹⁵ uma vez que foram unidas informações que resultam da obtenção de curva de distribuição de pontos de ebulição pelo método cromatográfico, utilizando-as em outras modelagens.⁹⁵⁻⁹⁹

Com a finalidade de obtenção da curva PEV pelas técnicas cromatográficas DHA e SIMDIS, foi necessário efetuar correções na curva de destilação simulada SIMDIS para fusão de dados e interpolação linear, como previamente descrito (Metodologia). A curva final obtida apresentou maior contribuição no conhecimento composicional do óleo analisado.

A curva PEV simulada por meio da fusão dos dados apresenta dados pela técnica DHA na região inicial até T10%, região comumente descrita como *light-end*. A partir deste ponto apresenta dados de destilação simulada SIMDIS. A Figura 5.12. Apresenta graficamente os principais procedimentos na fusão dos dados para obtenção da curva de destilação simulada.

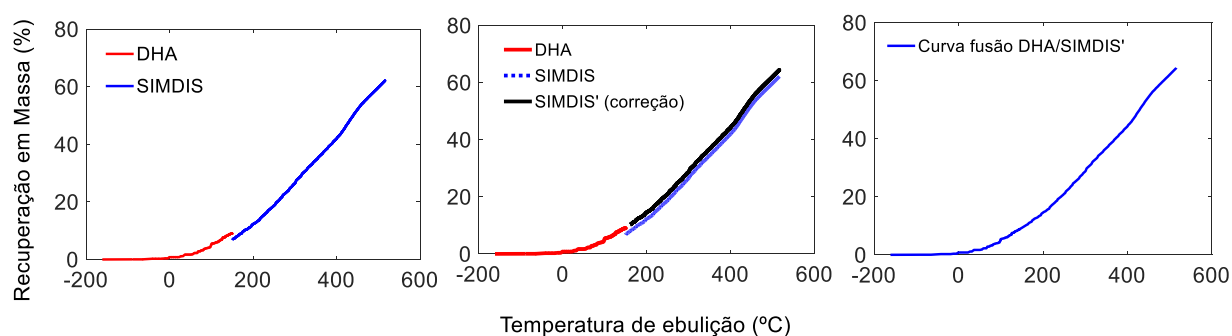


Figura 5.12 Efeito visual da correção das curvas para interpolação e fusão. (A) Curva DHA e SIMDIS, sem correção; (B) Curvas DHA, SIMDIS e SIMDIS' (com correção); (C) Curva final obtida DHA/SIMDIS'.

A curva final obtida para uma amostra de óleo médio está representada na Figura 5.8. O ajuste polinomial foi decisivo na suavização da curva final. A ordem do polinômio foi definida de acordo com o melhor ajuste da curva, variando entre 5 a 6 graus para as 42 amostras. A amostra representada na Figura 5.8 foi ajustada com polinômio de grau 6.

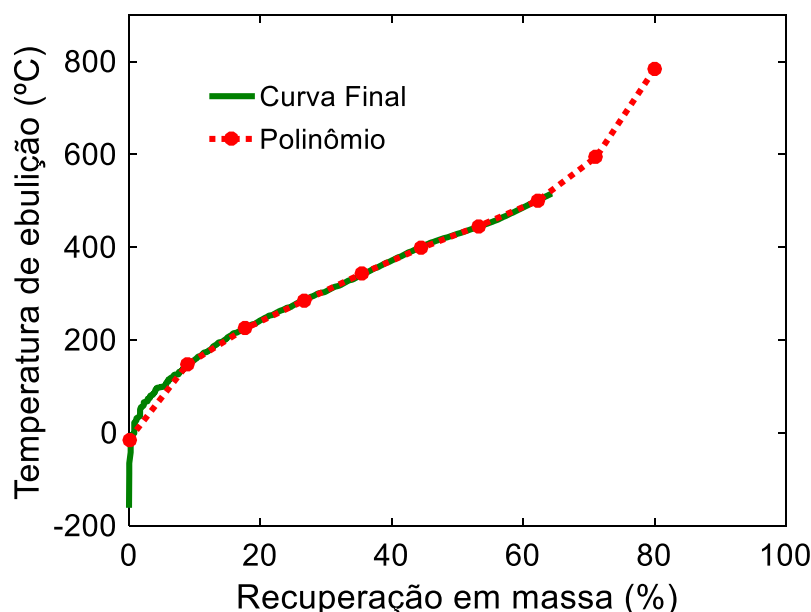


Figura 5.13 Curva simulada desenvolvida pela fusão de dados de DHA com SIMDIS para um óleo médio.

Uma comparação visual das curvas utilizadas para fusão e a curva final, para uma amostra de óleo médio são apresentados na figura 5.9.

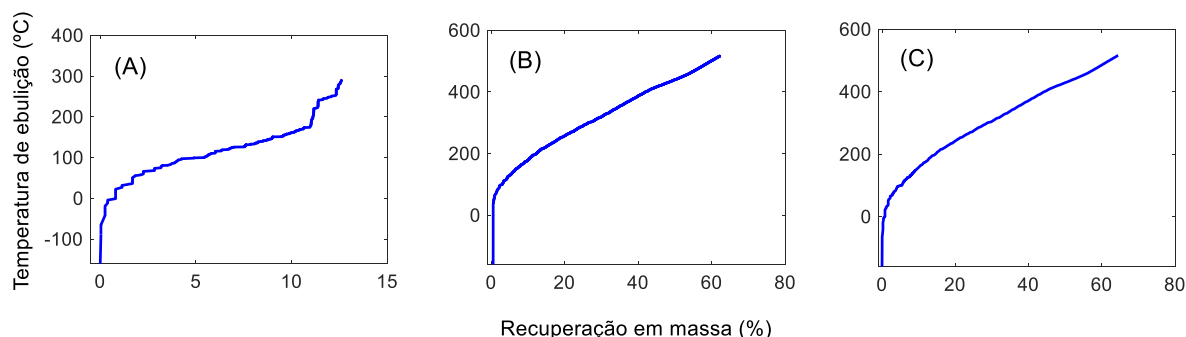


Figura 5.14 Comparação entre as curvas de amostra de petróleo médio para as técnicas de DHA, SIMDIS e fusão dos dados DHA/SIMDIS. (A) Curva DHA; (B) Curva SIMDIS; (C) Curva fusão DHA/SIMDIS.

Como pode ser observado na Figura 5.9, a curva apresentou-se mais suave na região inicial após a fusão dos dados (Figura 5.9 (C)), devido à eluição de diversos componentes com temperatura de ebulição muito próxima no DHA. A curva PEV pela destilação simulada apresenta maior contribuição de composição em relação a curva PEV pela destilação física, o que pode estar relacionado ao maior alcance de temperatura que a técnica SIMDIS proporciona.⁴²⁻⁴³

A partir de uma comparação entre curva PEV pela destilação física e a curva estimada pelo SIMDIS sem e com fusão de dados de cinco amostras deste conjunto

de dados foi possível identificar que houve uma redução do erro relativo correspondente ao SIMDIS quando correlacionado com a PEV (Figura 5.10). Os valores de erro relativo das temperaturas de destilação T10%, T20%, T30%, T40%, T50% e T60%, para as cinco amostras foram calculados com relação às suas respectivas PEV (Ver Anexo II).^{52,102} Os valores obtidos demonstram erros relativos menores para a fusão, principalmente na região inicial da curva. Porém, uma das amostras (Amostra 100) demonstrou resultado diferentemente do esperado, com valor de erro relativo discrepante na região da fusão de dados. Esse resultado pode compreender um erro experimental ou na coleta amostras, por exemplo.

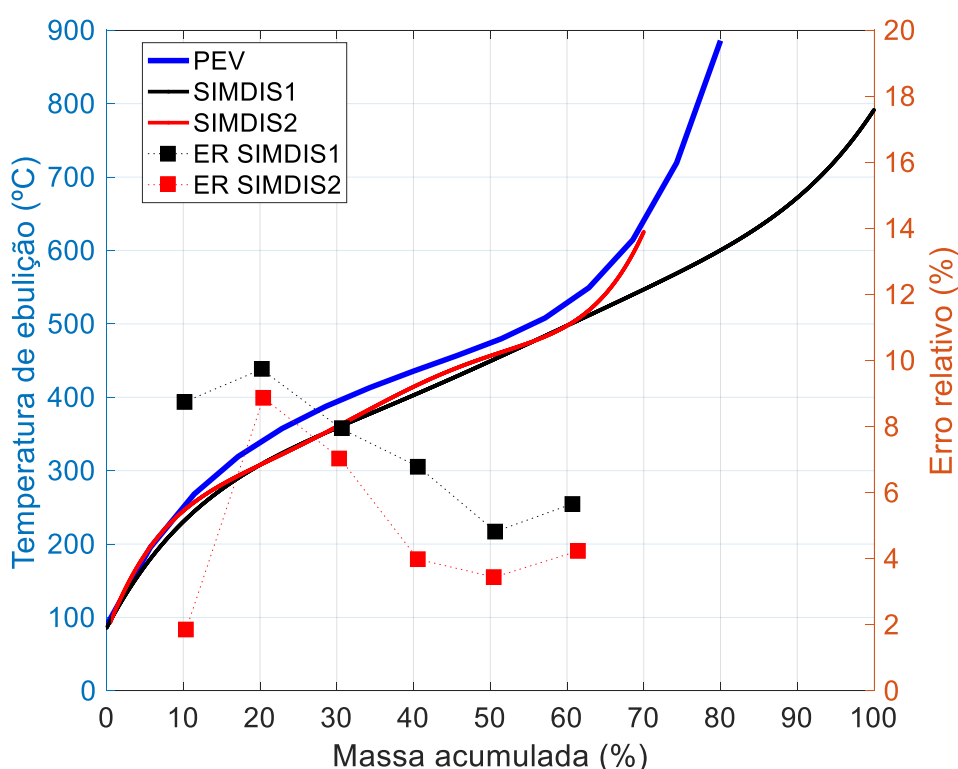


Figura 5.15 Gráfico de comparação entre erros relativos das curvas PEV, SIMDIS1 e SIMDIS2 de amostra de óleo pesado (API 17,7; PF 57,0 °C; PVR 10,1 kPa)..
 PEV: Destilação física; SIMDIS1: Destilação simulada por cromatografia gasosa SIMDIS;
 ER: Erro Relativo; SIMDIS2: Destilação simulada pela fusão de dados DHA/SIMDIS..

6 CONCLUSÕES

Neste trabalho foi proposta uma metodologia para estimativa da curva de destilação simulada de amostras de petróleo e suas propriedades físico-químicas, ponto de fulgor e pressão de vapor Reid, por meio de análise detalhada de hidrocarbonetos (DHA) e cromatografia gasosa de alta temperatura (HTGC) associados com calibração multivariada.

Utilizando a fusão de informações das duas técnicas cromatográficas (DHA e HTGC) é possível estimar o ponto de fulgor com um erro médio de 5,3 °C e coeficiente de determinação aproximadamente 0,9. A fusão de dados representa ganho de informações na determinação do ponto de fulgor, entretanto para determinar a pressão de vapor Reid somente os dados cromatográficos de HTGC são suficientes. A metodologia com regressão por PLS expressou melhor resultado para pressão de vapor Reid pela HTGC com erro médio de 0,4 kPa e coeficiente de determinação aproximadamente 0,99.

A fusão de informações é importante para obtenção de curva PEV simulada pela técnica SIMDIS de forma completa. A compensação da limitação na região inicial da curva de destilação simulada foi realizada de forma satisfatória por meio de quimiometria. Esse procedimento fornece uma curva de destilação simulada que pode trazer maior contribuição ao conhecimento composicional da amostra de petróleo.

Ressalta-se que a destilação simulada (SIMDIS) por cromatografia gasosa de alta temperatura (HTGC) é amplamente utilizada para obtenção da curva PEV simulada laboratorial. As metodologias propostas neste trabalho agregam aos dados obtidos pela HTGC outras importantes atribuições, como a determinação das propriedades físico-químicas. Dessa forma, informações de elevada relevância podem ser inferidas à indústria do petróleo, de forma rápida e precisa. Assim, decisões e processos podem ser otimizados, além de contribuições de importância para segurança do processo industrial.

7 REFERÊNCIAS

1. THOMAS, J. E. **Fundamentos de engenharia de petróleo**, Editora Interciência Ltda, 2004.
2. SPEIGHT, J. G. **Handbook of petroleum product analysis**, John Wiley & Sons, 2002.
3. RIAZI, M. R. **Characterization and properties of petroleum fractions**. ASTM Stock Number: MNL50, 1.ed., USA, Philadelphia, PA, 2005.
4. FAHIM, M.; AL-SAHHAH, T.; ELKILANI, A. **Fundamentals of Petroleum Refining**. 1. ed. Oxford: Elsevier, 2010.
5. FILGUEIRAS, P.R.; SAD, C. M. S.; LOUREIRO A. R.; SANTOS, M. F. P.; CASTRO, E. V.R.; DIAS, J. C. M. Determination of API gravity, kinematic viscosity and water content in petroleum by ATR-FTIR spectroscopy and multivariate calibration. *Fuel*, **v.116**, p.123-130, 2014.
6. FARAH, M. A. **Petróleo e seus derivados: definição, constituição, aplicação, especificações, características de qualidade**. Rio de Janeiro: LTC, 2012, 226p.
7. RÉ-POPPI, N.; ALMEIDA, F.F.P.; CARDOSO, C.AL.; RAPOSO-JR, VIANA, L.H.; SILVA,T.Q.; SOUZA,J.L.C. Screening analysis of type C Brazilian gasoline by gas chromatography – Flame ionization detector. *Fuel*. **v.88**, p.418-423, 2009.
8. FLUMIGNAN, D.L.; TININIS, A.G.; FERREIRA, F.O.; OLIVEIRA, J.E. Screening Brazilian C gasoline quality: Application of the SIMCA chemometric method to gas chromatographic data. *Anal. Chim. Acta*, **v.595**, p.128-135, 2007.
9. BUTLER, R. M.; COOKE, G. M.; LUKK, G. G.; JAMESON, B. G. Prediction of flash points of middle distillates. *Ind. Eng. Chem.*, **v.48**, p.808-8012, 1956.
10. DAI, Y.; LIU, H.; CHEN, X.; LIU, Y.; LI, X.; ZHU, Z.; ZHANG, Y., CAO Z. A new group contribution-based method for estimation of flash point temperature of alkanes. *J. Cent. South. Univ.*, **v.22**, p.30-36, 2015.
11. ALQAHEEM, S. S.; RIAZI, M. R. Flash points of hydrocarbons and petroleum products: prediction and evaluation of methods. *Energy Fuels*, **v.31**, p.3578-3584, 2017.

12. MONTIN, T. Online flash point calculation improves crude unit control. *Hydrocarbon Processing*, **v.77**, 12, 81, Dec. 1998. ISSN: 00188190.
13. SHUVALOV, G. V. An experimental investigation of a method of determining the flash point of petroleum products. *Measurement Techniques*, **v.52**, n.10, p.1119-1121 2009.
14. COOPLER, J. B.; WISE, K. L.; GROVES, J.; WELCH, W. T. Determination of octane numbers and Reid vapor-pressure of commercial petroleum fuels using FT-Raman spectroscopy and partial least squares regression-analyses. *Anal. Chem.*, **v.22**, n.67, p.4096-4100, 1995.
15. ALVES, J., C., L.; HENRIQUES, C., B.; POPPI, R., J. Determination of diesel quality parameters using support vector regression and near infrared spectroscopy for an in-line blending optimizer system. *Fuel*, **v.97**, p.710-717, 2012.
16. ROCHA, W., F., C.; NOGUEIRA, R.; VAZ, B., G. Validation of model of multivariate calibration: an application to the determination of biodiesel blend levels in diesel by near-infrared spectroscopy. *Journal of chemometrics* , **v.26**, p.456-461, 2012.
17. FILGUEIRAS, P. R.; ALVES, J. C. L.; SAD, C. M. S.; CASTRO, E. V. R.; DIAS, J. C. M.; POPPI, R. J. Evaluation of trends in residuals of multivariate calibration models by permutation test. *Chemometrics intel lab syst*, **v.133**, p.33-41, 2014.
18. CÔCO, L. C.; YAMAMOTO, I.; MEIEN, O. F. Study of correlations for physicochemical properties of Brazilian gasoline. *Chemometrics intel lab syst*, **v.76**, p.55-63, 2005.
19. MENDES, G.; ALEME, H. G.; BARBEIRA, P. J.S. Reid vapor prediction of automotive gasoline using distillation curves and multivariate calibration. *Fuel*, **v.187**, 167-172, 2017.
20. MALACARNE, M. M. **Implantação e otimização de uma unidade manual de destilação de petróleos a vácuo no LabPetro/UFES** . Dissertação de Mestrado. Departamento de Química da Universidade Federal do Espírito Santo, Vitória, 2011, 129p.
21. EGGERTSEN, F. T.; GROENNINGS, S.; HOLST, J.J. Analytical Distillation by Gas Chromatography Programmed Temperature Operation. *Anal. Chem.*, **v.32**, p.904-909, 1960.

22. GREEN, L.E.; SCHMAUCH, L.J.; WORMAN, J.C. Simulated Distillation by Gas Chromatography. *Anal. Chem.*, **v.36**, 1512-1516, 1964.
23. PETROFF, N.; HSCHEITT, A.; DURAND, J.P. Automated simulated distillation by gas chromatography. *J. Chromatogr.*, **v.395**, p.241-254, 1987.
24. FERREIRA, A. A.; NETO, F. R. A. A Destilação simulada na Indústria do Petróleo. *Quim. Nova*, **v.28**, p.478-482, 2005.
25. GRECSEK, H.; OSENBACH, T.; TIPLER, A. Integrating Detailed Hydrocarbon Analysis Data with Simulated Distillation to Improve the characterization of crude oils by Gas Chromatography. *Anal. Instrum., GC Focus*, p.20-22, 2010
26. BLOMBERG. J.; SCHOENMAKERS. P. J.; BRINKMAN. U. A. Th. Gas chromatographic methods for oil analysis. *J. Chromatogr.*, **v.972**, 137–173, 2002.
27. DASZYKOWSKI, M.; WALCZAK, B. Use and abuse of chemometrics in chromatography. *Trends in Analytical Chemistry*, **v.25**, n.11, p.1091-1096, 2006.
28. SENDRA-BOSQUE, J. M.; RODRIGUEZ, L. C.; SAMBIÁS, C. R.; LA-MATA, A. P. Combining chromatography and chemometrics for the characterization and authentication of fats and oils from triacylglycerol compositional data – A review. *Anal. Chim. Acta*, **v.724**, p.1-11, 2012.
29. BEENS J., BRINKMAN U., A., T. The role of gas chromatography in compositional analyses in the petroleum industry. *Trends in Analytical Chemistry*, **v.19**, p.260-275, 2000.
30. MÜHLEN, C. V.; ZINI, C. A.; CARAMÃO, E. B. Nomenclatura na língua portuguesa em cromatografia multidimensional abrangente. *Quím, Nova*, **v.30**, n.3, p.682-687, 2007.
31. NIZIO, K. D.; MCGINITIE, T. M.; HARYNUK, J., J. Comprehensive multidimensional separations for the analysis of petroleum. *J. Chromatogr.*, **v.1255**, p.12-23, 2012.
32. FERREIRA, M. M. C. **Quimiometria - Conceitos, métodos e aplicações**. Campinas: Editora da Unicamp, 2015, 493p.
33. NIELSEN, N. J.; BALLABIO, D.; TOMASI, G.; TODESCHINI, R.; CHRISTENSEN, J. H. Chemometric analysis of gas chromatography with flame ionization detection

- chromatograms: A novel method for classification of petroleum products. *J. Chromatogr.*, **v.1238**, p.121-127, 2012.
34. NETO, B. B.; SCARMINIO, I. S.; BRUNS, R. E. 25 anos de quimiometria no Brasil. *Quím. Nova*, **v.29**, p.1401-1406, 2006.
 35. FORINA, M.; LANTERI, S.; CASALE, M. Multivariate calibration - Review, *J. Chromatogr.*, **v.1158**, n.1–2, p.61–93, 2007.
 36. COLLINS, C.H.; BRAGA, G.L.; BONATO, P.S. **Fundamentos de cromatografia**. 1.ed. Campinas: Editora da UNICAMP, 2006, 453p.
 37. BONATO, P. S. **Cromatografia gasosa**. In: COLLINS, C.H.; BRAGA, G.L.; BONATO, P.S. *Fundamentos de cromatografia*. 1.ed. Campinas: Editora da UNICAMP, 2006. P.203-272.
 38. SCHNEIDER, W.; BRUDERRECK, H. Gas Chromatographic Separation of Hydrocarbons (C1 to C8) by Carbon Number Using Packed Capillary Columns. *Anal. Chem.*, **v.36**, n.8, p.1533-1540, 1964.
 39. PESEK, J.; BLAIR, B. termination of C8 and Heavier Molecular Weight Alklbenzenes in Petroleum Naphthas by Gas Chromatography. *Anal. Chem.*, **v.51**, p.12, 1979.
 40. ASTM International. ASTM D2892-05a. Standard Test Method for Distillation of Crude Petroleum: (15-Theoretical Plate Column). West Conshohocken, PA: American Society for Testing and Materials, 2005.
 41. ASTM International. ASTM D5236-07a. Standard Test Method for Distillation of Heavy Hydrocarbon Mixtures (Vacuum Potstill Method). West Conshohocken, PA: American Society for Testing and Materials, 2007.ASTM International.
 42. ESPINOSA-PEÑA, M.; FIGUEROA-GÓMEZ, Y.; JIMÉNEZ-CRUZ, F. Simulated Distillation Yield Curves in Heavy Crude Oils: A Comparison of Precision between ASTM D-5307 and ASTM D-2892 Physical Distillation. *Energy Fuels*, **v.18**, p.1832-1840, 2004.
 43. AUSTRICH, A. J.; BUENETRO-GONZALEZ, E.; LIRA-GALEANA, C. ASTM D-5307 and ASTM D-7169 SIMDIS Standards: A Comparison and Correlation of Methods. *Petroleum Science and Technology*, **v.33**, p.657–663, 2015.

44. ROUSSIS, S. G.; FITZGERALD, W. P. Gas Chromatographic Simulated Distillation-Mass Spectrometry for the Determination of the Boiling Point Distributions of Crude Oils. *Anal. Chem.*, **v.72**, p.1400-1409, 2000.
45. COTO, B.; COUTINHO, J.A.P.; MARTOS,C.; RBUSTILLO, M.D.; ESPADA, J.J.; PEÑA,J.L. Assessment and Improvement of n-Paraffin Distribution Obtained by HTGC To Predict Accurately Crude Oil Cold Properties. *Energy Fuels*, **v.25**, p.1153–1160, 2011.
46. ESPADA, J.J.; FERNANDEZ, S.; VELASCO, L.; COTO,B. Evaluation of different methodologies to determine the n-paraffin distribution of petroleum fractions. *Fuel*, **v.109**, p.470-475, 2013.
47. BRERETON, R. G. The evolution of chemometrics. *Anal. Methods*, **v.5**, p.3785-3789, 2013.
48. TOZZI, F. C.; SAD, C.M.S.; BASSANE,J.F.P.; SANTOS,F.D.; SILVA,M.; FILGUEIRAS,P.R.; DIAS,H.P.; ROMÃO,W.; CASTRO,E.V.R.; LACERDA JR,V. Improving the physicochemical properties of Brazilian onshore and offshore crude oils using the production of blends. *Fuel*, **v.159**, p.607-613, 2015.
49. MORALES-MEDINA, G.; GUZMÁN, A. Prediction Of Density And Viscosity Of Colombian Crude Oils From Chromatographic Data. *CT&F – Ciencia, Tecnología y Futuro*, **v.5**, p.57-74, 2012.
50. TERRA, L. A.; FILGUEIRAS, P. R.; TOSE, L. V.; ROMÃO, W.; CASTRO, E. V. R. Laser desorption ionization FT-ICR mass spectrometry and CARSPLS for predicting basic nitrogen and aromatics contents in crude oils. *Fuel*, **v.160**, p.274–281, 2015.
51. DUARTE, L. M.; FILGUEIRAS, P.R.; SILVA, S.R.C. et al. Determination of some physicochemical properties in Brazilian crude oil by ¹H NMR spectroscopy associated to chemometric approach. *Fuel*, **v.181**, p.660–669, 2016.
52. DUARTE, L. M.; FILGUEIRAS,P. R.; DIAS, J. C. M.; OLIVEIRA, L. M. S. L.;CASTRO, E. V. R.; OLIVEIRA, M. A. L. Study of Distillation Temperatura Curves from Brazilian Crude Oil by ¹H Nuclear Magnetic Resonance Spectroscopy in Association with Partial Least Squares Regression. *Energy Fuels*, **v. 31**, p.3892-3897, 2017,

53. FILGUEIRAS, P. R.; TERRA, L. A.; CASTRO, E. V. R.; OLIVEIRA, L. M. S. L.; DIAS, J.C. M.; RONNEY, J. L. Prediction of the distillation temperatures of crude oils using ^1H NMR and support vector regression with estimated confidence intervals. *Talanta*, **v.142**, p.197-205
54. BRASIL. Agência Nacional de Petróleo, gás natural e biocombustíveis. Portaria n. 47, de 04/09/2014. Disponível em: <<http://nxt.anp.gov.br/NXT/gateway.dll?f=templates&fn=default.htm&vid=anp:10.1048/enu>> Acesso em 23 out. 2016.
55. TISSOT, B. P.; WELTE, D. H. **Petroleum formation and occurrence**. New York: Springer-Verlag, 1984.
56. ULLER, V. C. **Oportunidades e Desafios da Colocação de Óleos Brasileiros no Mercado Internacional: O Refino e o Mercado Norte-Americano de Combustíveis**. Dissertação de Mestrado. Departamento de Ciências da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2007. 226p.
57. WINTER, A. **Caracterização de Frações Ultra-Pesadas de Petróleo Nacional por meio do Processo de Destilação Molecular**. Dissertação de Mestrado. Departamento de Engenharia Química da Universidade Estadual de Campinas, São Paulo, 2007. 137p.
58. FILGUEIRAS, P. R. **Determinação da composição de blends de petróleos utilizando FFTIR-ATR e calibração multivariada**. Dissertação de Mestrado. Departamento de Química da Universidade Federal do Espírito Santo, Vitória, 2011. 112p.
59. NETO, D. M. C. **Estudo reológico de emulsões água em óleo de petróleos pesados**. Dissertação de Mestrado. Departamento de Química da Universidade Federal do Espírito Santo, Vitória, 2017. 80p.
60. BRASIL. BRASIL. Agência Nacional de Petróleo, gás natural e biocombustíveis.. Resolução n.º 703/2017. Disponível em: <<http://nxt.anp.gov.br/NXT/gateway.dll?f=templates&fn=default.htm&vid=anp:10.1048/enu>> Acesso em 30 set. 2017.
61. BRASIL-. Agência Nacional de Petróleo, gás natural e biocombustíveis. Portaria n. 206, de 29/08/2000. Disponível em <<http://nxt.anp.gov.br/NXT/>

gateway.dll?f=templates&fn=default.htm&vid=anp:10.1048/enu>. Acesso em 23 out. 2016.

62. BRASIL. Ministério do Trabalho. Norma Regulamentadora N.º 20. Segurança e saúde no trabalho com inflamáveis e combustíveis.
63. ISO 13736:2008, Determination of flash point — Abel closed-cup method. British Standards, 2008.
64. BRASIL. Agência Nacional de Petróleo, gás natural e biocombustíveis. Regulamento Técnico 3/2013. Resolução n. 40, de 25/10/2013.
65. ASTM International. ASTM D323-15a. Standard Test Method for Vapor Pressure of Petroleum Products (Reid Method). West Conshohocken, PA: American Society for Testing and Material, 2015.
66. AQUINO NETO, F. R.; NUNES, D. S. **Cromatografia: princípios básicos e técnicas afins**. 1.ed. Rio de Janeiro: Interciência, 2003.
67. SKOOG, A.; HOLLER, F. J.; CROUCH, S. R. **Princípios de Análise Instrumental**. 6.ed. Porto Alegre: Bookman, 2009, cap.6.
68. AC ANALYTICAL CONTROLS. AC DHA Operator Manual, 131p, 2004.
69. ASTM International. ASTM D6729-14. Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 100 Metre Capillary High Resolution Gas Chromatography. West Conshohocken, PA: American Society for Testing and Material, 2015.
70. BRASIL. Agência Nacional de Petróleo, gás natural e biocombustíveis. Resolução n.684 de 29/06/2017.
71. ASTM D5134-13. Standard Test Method for Detailed Analysis of Petroleum Naphthas through n-Nonane by Capillary Gas Chromatography1. West Conshohocken, PA: American Society for Testing and Material, 2014.
72. ASTM International. ASTM D6730 – 01. Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 100–Metre Capillary (with Precolumn) High-Resolution Gas Chromatography. West Conshohocken, PA: American Society for Testing and Material, 2016.

73. ASTM International. ASTM D6733 – 01. Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 50-Metre Capillary High Resolution Gas Chromatography¹. West Conshohocken, PA: American Society for Testing and Material, 2016.
74. ASTM International. ASTM D7900-13. Standard Test Method for Determination of Light Hydrocarbons in Stabilized Crude Oils by Gas Chromatography. West Conshohocken, PA: American Society for Testing and Material, 2014.
75. ETTRE, L. S. The Kováts retention index system. *Anal. Chem.*, **v.36**, n.08, p.31A – 40A, 1964.
76. ASTM International. ASTM D7169-11. Standard Test method for boiling point distribution of samples with residues such as crude oils and atmospheric and vacuum residues by high temperature gas chromatography. West Conshohocken, PA: American Society for Testing and Material, 2014.
77. ASTM International. ASTM Standard D2887-16a: Standard Test Method for Boiling Range Distribution of Petroleum Fractions by Gas Chromatography; ASTM International: West Conshohocken, PA, 2016.
78. ASTM International. ASTM Standard D5307: Standard Test Method for Boiling Range Distribution of Petroleum Fractions by Gas Chromatography; ASTM International: West Conshohocken, PA, 2002.
79. FERREIRA, M.C.; ANTUNES, A.M.; MELGO, M.S.; VOLPE, P.L.O. Quimiometria I: calibração multivariada, um tutorial. *Química Nova*, v.22, n.5, p. 724-731., 1999.
80. BRERETON, R. G. Introduction to multivariate calibration in analytical chemistry. *Analyst.*, **v.125**, p.2125-2154, 2000.
81. GELADI, P.; KOWALSKI, B. R. Partial least-squares regression: a tutorial. *Anal. Chim. Acta*, **v.185**, p.1-17, 1986.
82. MARTENS, H.; NAES, T. **Multivariate Calibration**. Guildford: John Wiley & Sons Ltda, 1991, 419 p.
83. SAVORANI, F.; TOMASI, G.; ENGELSEN, S.B. icoshift: A versatile tool for the rapid alignment of 1D NM spectra. *J.Magn.Resson.*, **v.202**, 190-202, 2010.

84. TOMASI, G.; SAVORANI, F.; ENGELSEN, S. Icoshifit: an effective tool for the alignment of chromatographic data. *J. Chromatogr.*, **v.1218**, p.7832-7840, 2011.
85. BLANCO, M.; COELHO, J.; ITURRIAGA, H.; MASPOCH, S.; de la PEZUELA, C. Effect of data preprocessing methods in near-infrared diffuse reflectance spectroscopy for the determination of the active compound in a pharmaceutical preparation. *Appl. Spectrosc.*, **v.51**, n.2, p.240-246, 1997.
86. ELLERS, P. H. C. A Perfect smoother. *Anal. Chem.*, **v.75**, p.3631-3636, 2003.
87. BARNES, R. J.; DHANOA, M. S.; LISTER, S. J.; Standard Normal Variate transformation and De-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.*, **v.43**, n.5, p.772-777, 1989.
88. FEARN, T.; RICCIOLI, C.; GARRIDO-VARO, A.; GUERRERO-GINEL, J. E. On the geometry of SNV and MSC. *Chemometrics intel. Lab. Syst.*, **v.96**, p.22-26, 2009.
89. BRAGA, J. W. B.; POPPI, R. J. validação de modelos de calibração multivariada: uma aplicação na determinação de pureza polimórfica de carbamazepina por espectroscopia no infravermelho próximo. *Quim. Nova*, **v.27**, n.6, p.1004-1011, 2004.
90. ASTM. Standard Practices for Infrared Multivariate Quantitative Analysis, E-165505. West Conshohocken, PA: ASTM International; 2012.
91. VAN DER VOET, H. Pseudo-degrees of freedom for complex predictive models: the example of partial least squares. *J. Chem.*, **v.13**, p.195-208, 1999.
92. THOMAS, E. V. Non-parametric statistical methods for multivariate calibration model selection and comparison. *J. Chem.*, **v.17**, n.12, p.653-652, 2003.
93. PÉREZ-GUAITA, D.; KULIGOWSKI, J.; GARRIGUES, S.; QUINTÁS, G.; WOOD, B. R. Assessment of the statistical significance of classifications in infrared spectroscopy based diagnostic models. *Analyst.*, **v.140**, p.2422-2427, 2015.
94. TOMREN, A. L.; BARTH, T. Comparison of partial least squares calibration models of viscosity, acid number and asphaltene content in petroleum, based on GC and IR data. *Fuel*, **v.120**, p.8-21, 2014.

95. CASTANEDO, F., A review of data fusion techniques. *The Scientific World Journal* **v.2013**, p.1-19, 2013.
96. BORRÀS, E.; FERRÉ, J.; BOQUÉ, R.; MESTRES, M.; ACEÑA, L.; BUSTO, O., Data fusion methodologies for food and beverage authentication and quality assessment–A review. *Anal. Chim. Acta*, **v.891**, p.1-14, 2015.
97. CHAKRABORTY, S. A.; WEINDORFB, D. C.; LIC, B.; ALDABAA, A. A. A.; GHOSH, R. K.; ALI, S. T.; ALI, N. Development of a hybrid proximal sensing method for rapid identification of petroleum contaminated soils. *Science of the Total Environment*, **v.514**, p.399-408, 2015.
98. MA, H. Formation drillability prediction based on multi-source information fusion. *Journal of Petroleum Science and Engineering*, **v.78**, p.438-446, 2011.
99. PWASONG, A.; SATHASIVAN, S. On the fusion of regression and neural network methods. AIP Conference Proceedings 1750, 020002, 2016. Disponível em: <<https://doi.org/10.1063/1.4954515>>. Acesso em Dez. 2017.
100. KENNARD, R.W.; STONE, L.A. Computer aided design of experiments, *Technometrics*, **v.11**, p.137-148, 1969.
101. LIANG, C. et al. A new multivariate calibration model transfer method of near-infrared. *Chemometrics intel. Lab. Syst.*, **v.153**, p.51-57, 2016.
102. SOARES, R.F.; PEREIRA, R.; SILVA, R.S.F.; MOGOLLON, L.; AZEVEDO, D.E. Comprehensive two-dimensional gas chromatography coupled to time of flight mass spectrometry: new biomarker parameter proposition. *J Braz Chem Soc*, **v.24**, p.1570-1581, 2013.

ANEXOS

ANEXO I – Informações de 128 amostras de petróleo utilizadas neste trabalho

Tabela A1. Dados parâmetros de 128 amostras de petróleo utilizadas no trabalho.

AMOSTRA	DENSIDADE (°API) ISO 12185	PF (°C) ISO 13736	PF (°C) PLS	PVR (kPa) ASTM D 323	PVR (kPa) PLS	Ensaio	
						DHA	HTGC
001	54,0	----	----	56,3	----	----	X
002	41,9	≤-30,0	----	62,8	----	----	X
003	17,0	----	----	11,3	----	----	X
004	25,5	----	----	28,9	27,8	----	X
005	23,6	60,0	54,59	18,0	----	----	X
006	29,9	≤-30,0	----	29,6	29,4	----	X
007	53,3	----	----	45,4	44,7	----	X
008	26,8	-6,0	-4,55	32,7	32,3	----	X
009	26,7	≤-30,0	----	43,4	43,9	----	X
010	39,5	≤-30,0	----	37,1	37,5	----	X
011	29,0	≤-30,0	----	34,1	33,3	----	X
012	28,4	-15,0	-11,90	40,8	39,7	----	X
013	29,4	≤-30,0	----	48,2	47,9	----	X
014	24,4	-6,0	5,46	26,3	26,5	----	X
015	29,6	≤-30,0	----	48,2	48,1	----	X
016	22,0	27,0	11,73	16,6	----	----	X
017	30,8	----	----	37,4	37,9	----	X
018	30,6	-8,0	-21,49	34,9	34,9	----	X
019	23,4	----	----	32,2	31,7	----	X
020	28,7	≤-30,0	----	41,1	40,7	----	X
021	29,2	≤-30,0	----	45,6	45,2	----	X
022	29,8	≤-30,0	----	46,5	46,2	----	X
023	32,1	-15,0	-24,08	51,0	51,0	----	X
024	25,5	----	----	47,7	48,0	----	X
025	37,8	≤-30,0	----	50,9	50,8	----	X
026	30,2	≤-30,0	----	50,4	50,3	----	X
027	41,0	≤-30,0	----	29,0	28,4	----	X
028	25,9	≤-30,0	----	50,6	50,5	----	X
029	35,9	----	----	58,8	----	----	X

continua

Tabela A1. Dados parâmetros de 128 amostras de petróleo utilizadas no trabalho.

AMOSTRA	DENSIDADE (°API) ISO 12185	PF (°C) ISO 13736	PF (°C) PLS	PVR (kPa) ASTM D 323	PVR (kPa) PLS	Ensaio	
						DHA	HTGC
Continuação							
030	28,0	-18,0	-14,49	32,0	31,3	----	X
031	27,5	≤-30,0	----	50,0	50,0	----	X
032	29,5	-20,0	-21,02	47,6	47,6	----	X
033	40,5	----	----	38,5	38,9	----	X
034	21,0	≤-30,0	----	42,2	42,5	----	X
035	23,6	6,0	13,29	42,5	42,9	----	X
036	52,0	----	----	46,4	45,9	----	X
037	36,7	-17,0	5,00	29,4	29,0	----	X
038	30,8	----	----	28,4	27,5	----	X
039	34,0	≤-30,0	----	50,4	50,4	----	X
040	29,3	----	----	30,3	30,6	----	X
041	22,7	----	----	23,2	25,7	----	X
042	28,4	-2,0	-13,07	----	----	----	X
043	29,8	≤-30,0	----	46,5	46,5	----	X
044	27,8	-9,0	-13,87	48,5	48,6	----	X
045	44,6	≤-30,0	----	53,1	----	----	X
046	33,0	-6,0	-21,91	41,8	42,1	----	X
047	30,6	-8,0	-16,48	34,9	35,2	----	X
048	52,4	----	----	14,3	----	----	X
049	28,4	-2,0	-0,13	35,9	36,0	X	X
050	29,6	≤-30,0	----	45,6	45,5	----	X
051	25,9	≤-30,0	-20,94	50,6	50,5	X	X
052	25,6	0,0	-1,05	28,9	28,1	X	X
053	30,1	≤-30,0	-28,69	29,6	29,7	X	X
054	29,6	≤-30,0	-30,16	48,2	48,3	X	X
055	29,4	≤-30,0	-27,54	48,2	48,5	X	X
056	29,0	≤-30,0	-29,36	34,1	33,6	X	X
057	28,7	≤-30,0	-27,37	34,4	34,5	X	X
058	31,7	----	----	73,0	----	----	X
059	29,3	≤-30,0	----	50,7	50,7	----	X
060	26,8	-6,0	-2,28	32,7	32,6	----	X
061	23,6	6,0	2,69	42,5	43,2	X	X
062	28,4	-15,0	-14,42	40,8	40,2	----	X

continua

Tabela A1. Dados parâmetros de 128 amostras de petróleo utilizadas no trabalho.

AMOSTRA	DENSIDADE (°API) ISO 12185	PF (°C) ISO 13736	PF (°C) PLS	PVR (kPa) ASTM D323	PVR (kPa) PLS	Ensaio	
						DHA	HTGC
Continuação							
063	41,0	≤-30,0	-14,49	29,0	28,7	X	X
064	17,6	33,0	----	11,8	----	----	X
065	29,0	≤-30,0	-21,02	50,6	50,7	----	X
066	30,8	≤-30,0	----	41,3	41,2	X	X
067	28,8	1,0	----	----	----	X	X
068	29,6	≤-30,0	13,29	59,3	----	X	X
069	28,3	-9,0	----	45,5	44,9	X	X
070	28,1	-3,0	5,00	48,7	49,1	X	X
071	26,2	-8,0	----	27,4	26,8	X	X
072	17,7	57,0	----	10,1	13,9	X	X
073	26,9	-18,0	----	39,2	39,3	X	X
074	47,6	≤-30,0	----	31,0	30,9	----	X
075	22,1	1,0	-13,07	28,0	27,2	X	X
076	30,4	≤-30,0	----	41,5	41,6	X	X
077	28,1	≤-30,0	-13,87	36,4	36,8	X	X
078	21,9	-8,0	----	26,1	26,2	X	X
079	28,8	-16,0	-21,91	29,8	29,8	X	X
080	29,7	≤-30,0	-16,48	35,3	35,5	X	X
081	36,1	-15,0	----	----	----	X	----
082	30,2	≤-30,0	-0,13	48,6	48,8	X	X
083	32,3	≤-30,0	----	45,2	44,2	X	X
084	22,2	-14,0	-20,94	32,2	32,0	----	X
085	31,4	≤-30,0	-1,05	24,9	25,1	X	X
086	30,5	≤-30,0	-28,69	48,7	49,0	X	X
087	30,9	≤-30,0	-30,16	47,4	47,3	X	X
088	28,1	≤-30,0	-27,54	53,3	52,9	X	X
089	22,6	-12,0	-29,36	34,3	34,1	X	X
090	30,2	≤-30,0	-27,37	37,8	38,4	X	X
091	17,7	24,0	----	22,4	----	----	X
092	32,4	-12,0	----	20,2	25,1	X	X
093	29,1	≤-30,0	-2,28	49,5	49,6	X	X
094	30,8	-23,0	2,69	36,4	37,2	X	X
095	22,2	-3,0	-14,42	20,9	18,8	X	X

continua

Tabela A1. Dados parâmetros de 128 amostras de petróleo utilizadas no trabalho.

AMOSTRA	DENSIDADE (°API) ISO 12185	PF (°C) ISO 13736	PF (°C) PLS	PVR (kPa) ASTM D323	PVR (kPa) PLS	Ensaio	
						DHA	HTGC
Continuação							
096	26,8	-6,0	-5,47	32,7	32,9	X	X
097	30,2	≤-30,0	-28,75	50,4	50,4	X	X
098	27,9	≤-30,0	-31,82	20,7	18,4	X	X
099	46,4	≤-30,0	-25,28	45,4	44,7	X	----
100	28,0	-5,0	-8,72	30,1	30,2	X	X
101	36,7	-17,0	-10,34	29,4	29,1	X	----
102	29,8	≤-30,0	-32,47	46,5	47,0	X	X
103	29,0	≤-30,0	-12,73	50,6	47,6	X	----
104	28,4	-15,0	-21,39	40,8	43,0	X	----
105	29,7	≤-30,0	-15,07	35,3	35,8	----	X
106	30,6	-8,0	-18,86	25,6	22,4	X	----
107	23,6	6,0	-6,05	-----	----	X	----
108	28,7	≤-30,0	-20,69	49,1	49,5	X	----
109	28,3	-11,0	-20,77	48,3	----	X	----
110	28,8	≤-30,0	-19,91	57,2	56,0	X	----
111	31,9	≤-30,0	-24,10	43,5	37,7	X	----
112	29,8	≤-30,0	-31,05	50,2	50,2	X	X
113	28,5	≤-30,0	-36,70	62,8	61,6	X	----
114	28,9	-10,0	-26,34	56,8	----	X	----
115	20,6	21,0	6,41	17,7	----	X	----
116	30,7	-12,0	-22,91	-	----	X	----
117	26,9	≤-30,0	-22,61	40,3	----	X	----
118	30,6	≤-30,0	-24,41	40,9	----	X	----
119	32,2	≤-30,0	-23,43	-	----	X	----
120	29,3	≤-30,0	-23,09	48,5	----	X	----
121	28,4	≤-30,0	-16,47	52,6	----	X	----
122	29,0	-17,0	-20,38	-	----	X	----
123	27,6	≤-30,0	-24,81	-	----	X	----
124	30,3	≤-30,0	-31,73	36,0	36,5	----	X
125	29,5	≤-30,0	-31,36	-	----	X	----
126	27,7	≤-30,0	-26,94	-	----	X	----
127	28,5	≤-30,0	-27,05	47,5	----	X	----
128	30,6	≤-30,0	-32,63	42,5	43,6	----	X

ANEXO II – Comparação curvas PEV e destilação simulada DHA/SIMDIS

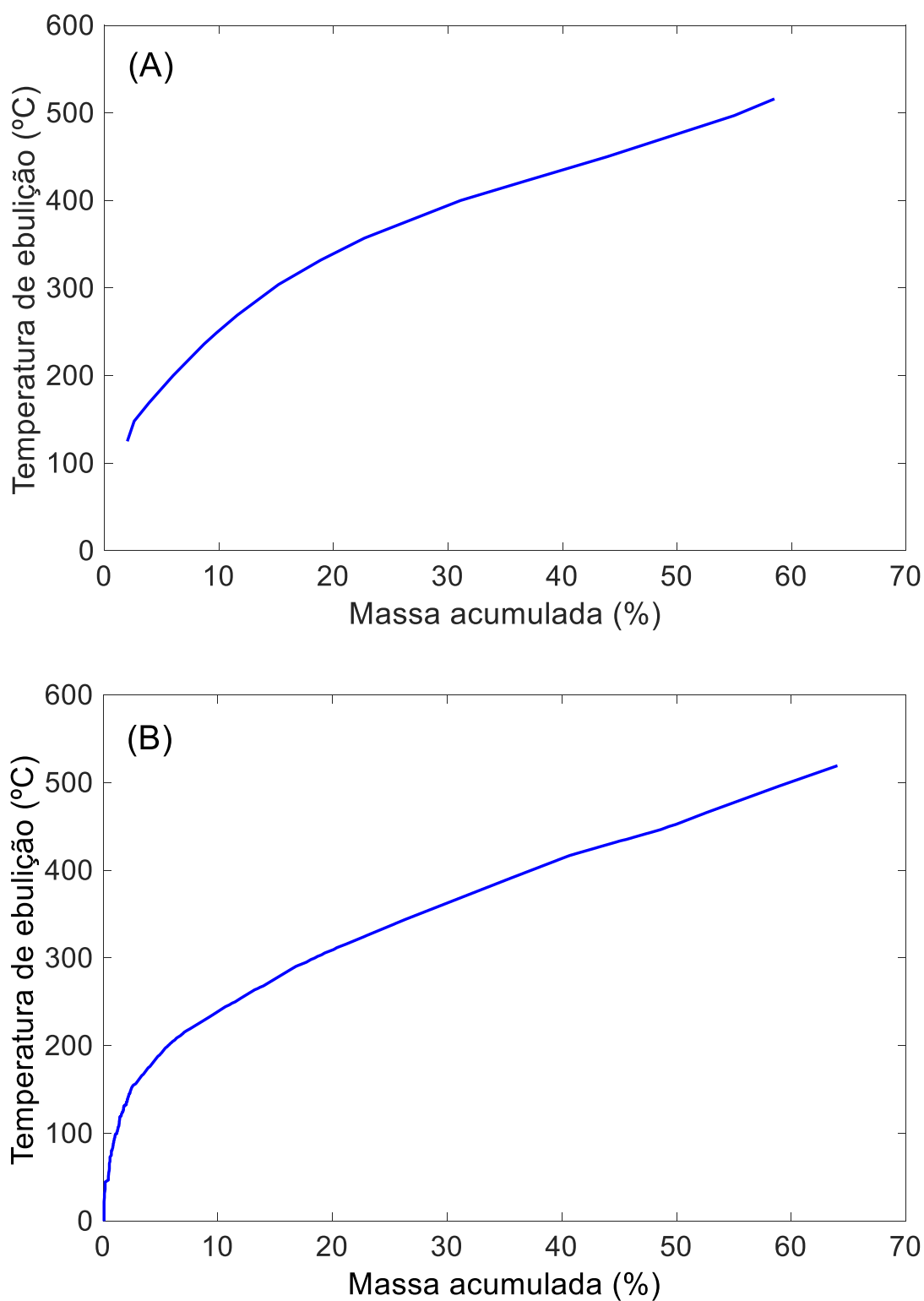


Figura A1 Curvas PEV de Amostra de óleo bruto. (A) Curva PEV Destilação física. (B) Curva de destilação simulada SIMDIS com fusão de dados DHA e SIMDIS. Amostra 072. Dados: API 17,7; PF 57,0 °C; PVR 10,1 kPa

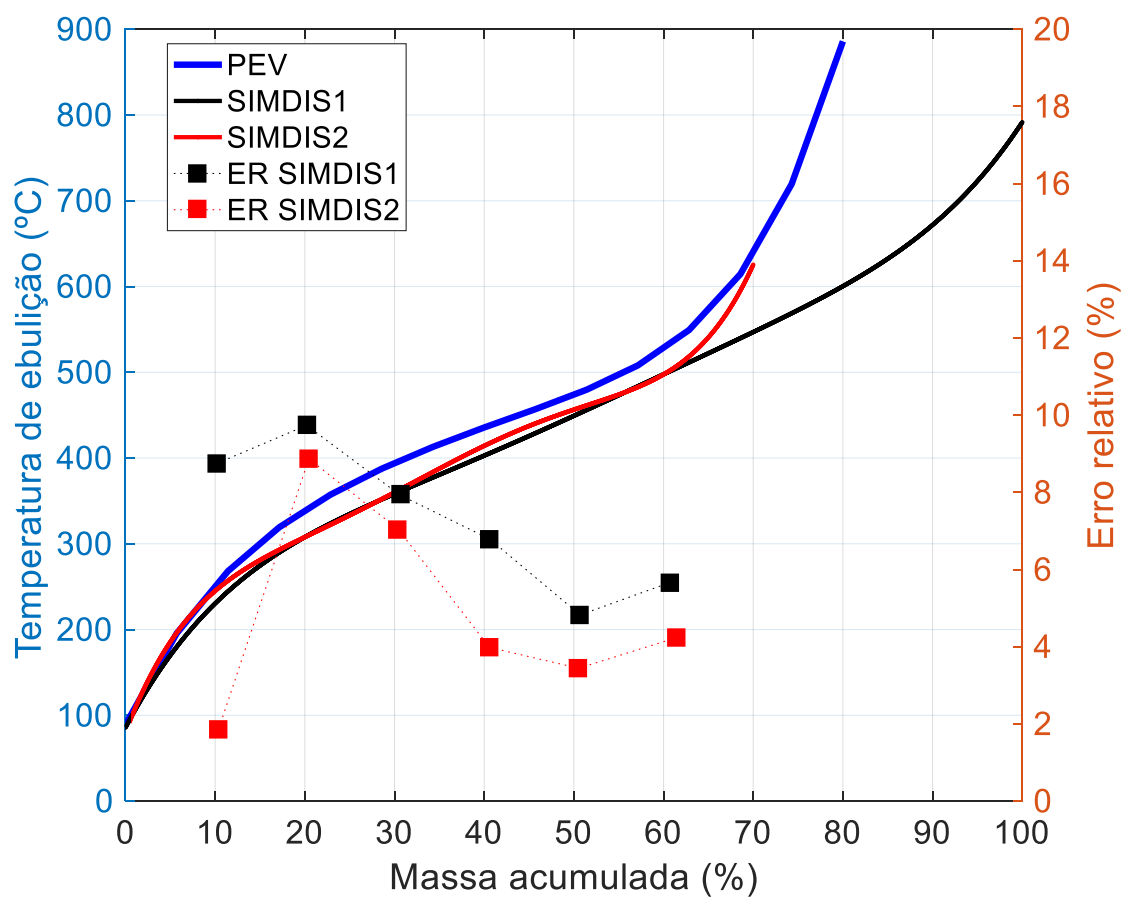


Figura A2 Gráfico de comparação entre erros relativos das curvas PEV, SIMDIS1 e SIMDIS2 de Amostra 072. PEV: Destilação física; SIMDIS1: Destilação simulada por cromatografia gasosa SIMDIS; ER: Erro Relativo; SIMDIS2 Destilação simulada pela fusão de dados DHA/SIMDIS.

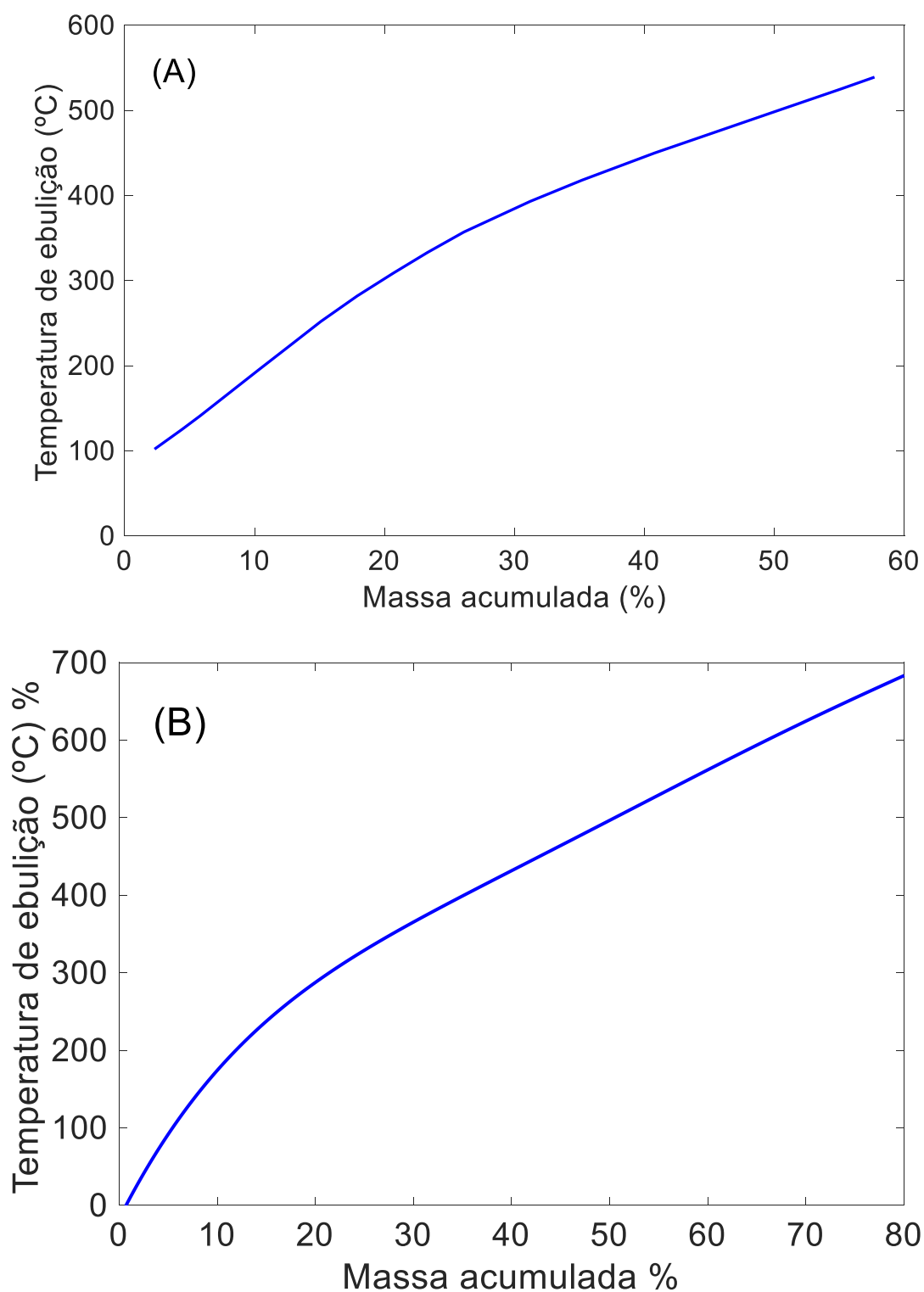


Figura A3 Curvas PEV de amostra de óleo bruto. (A) Curva PEV Destilação física. (B) Curva de destilação simulada SIMDIS com fusão de dados DHA e SIMDIS. Amostra 061. Dados: API 23,6; PF 6,0 °C; PVR 42,5 kPa.

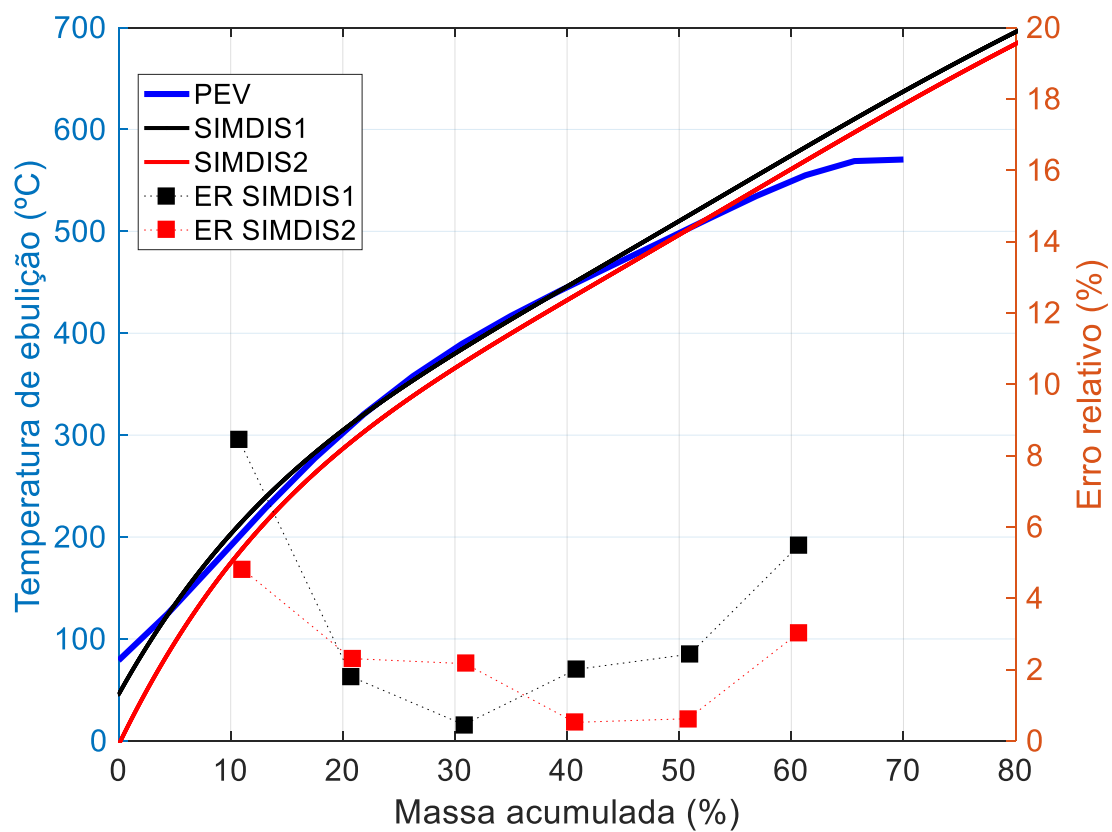


Figura A4 Gráfico de comparação entre erros relativos das curvas PEV, SIMDIS1 e SIMDIS2 de Amostra 061. PEV: Destilação física; SIMDIS1 : Destilação simulada por cromatografia gasosa SIMDIS; ER : Erro Relativo; SIMDIS2 : Destilação simulada pela fusão de dados DHA/SIMDIS.

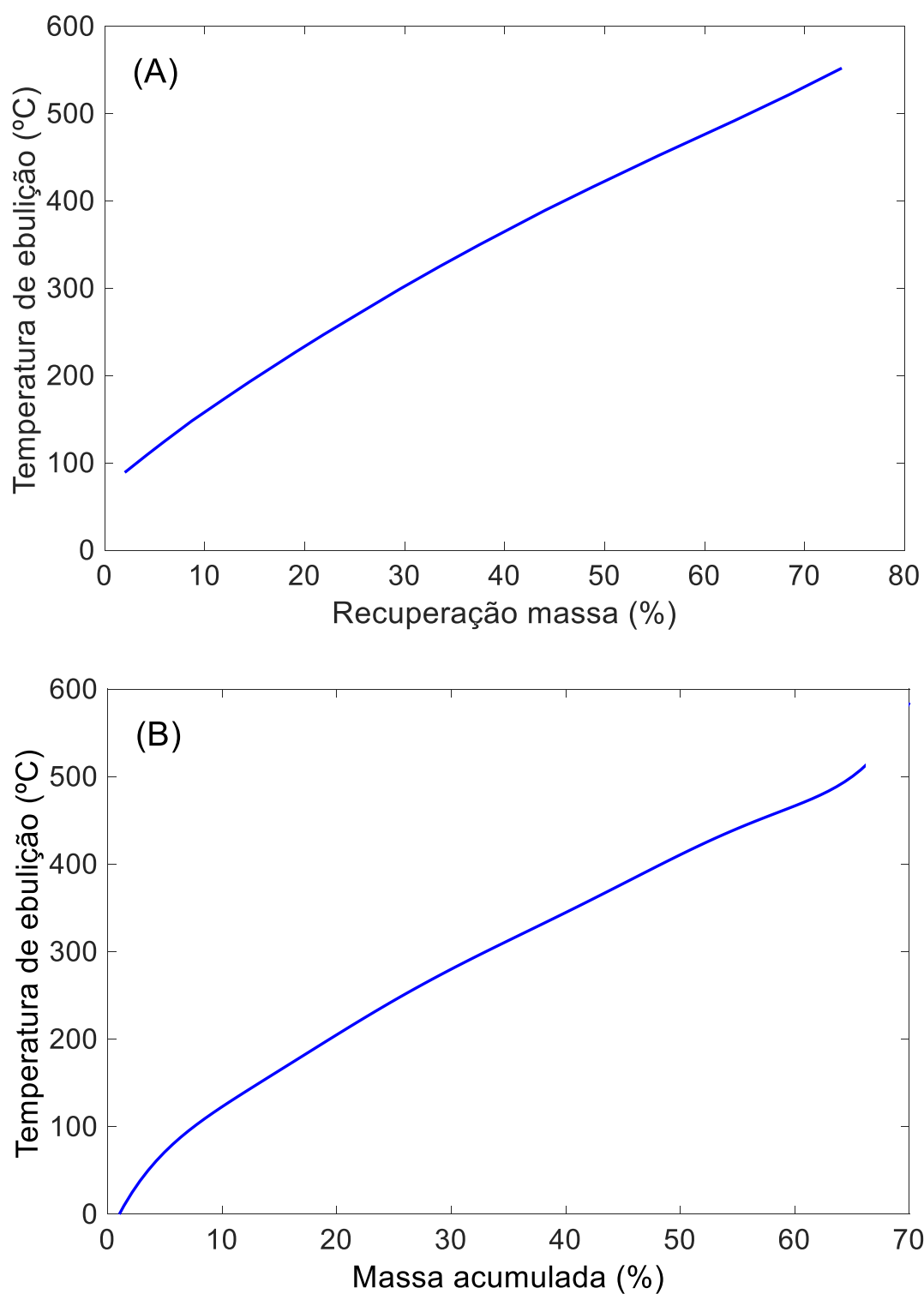


Figura A5 Curvas PEV de amostra e óleo bruto. (A) Curva PEV Destilação física. (B) Curva de destilação simulada SIMDIS com fusão de dados DHA e SIMDIS. Amostra 100. Dados: API 28,0; PF -5,0 °C; PVR 30,1 kPa.

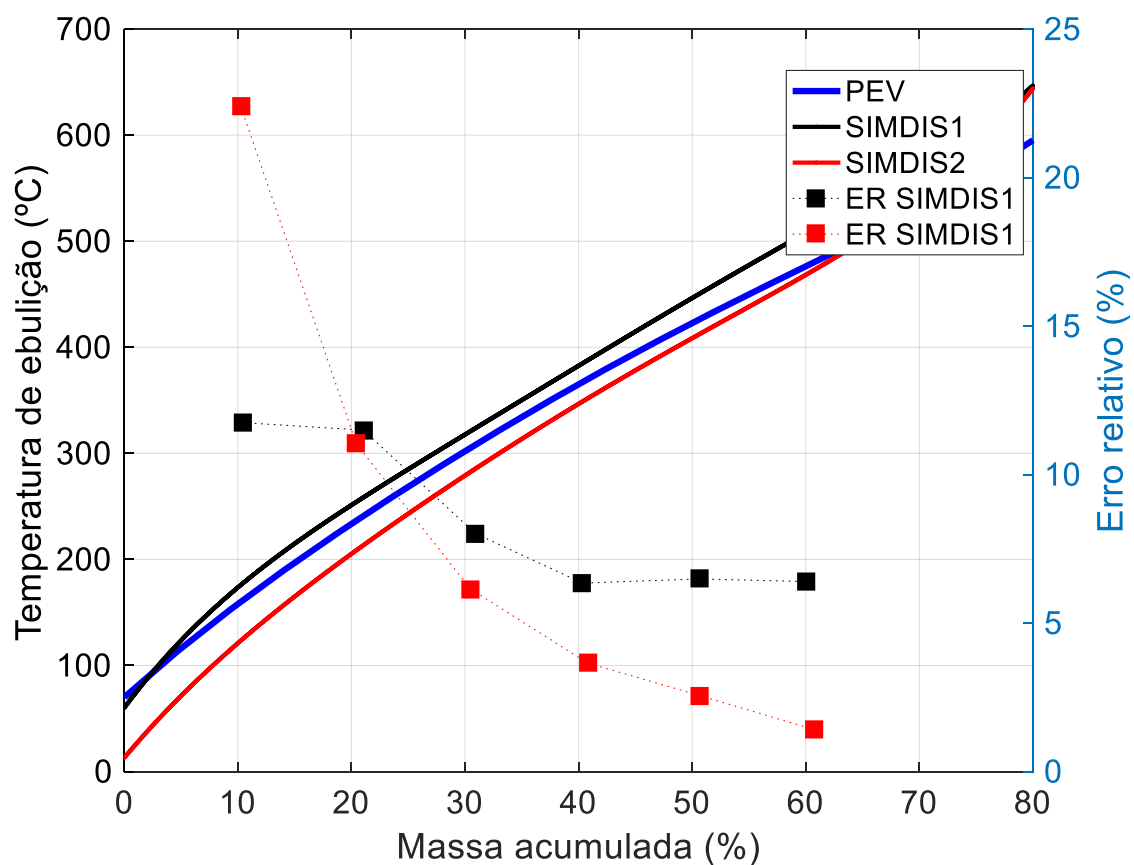


Figura A6 Gráfico de comparação entre erros relativos das curvas PEV, SIMDIS1 e SIMDIS2 de Amostra 100. PEV: Destilação física; SIMDIS1 : Destilação simulada por cromatografia gasosa SIMDIS; ER : Erro Relativo; SIMDIS2 : Destilação simulada pela fusão de dados DHA/SIMDIS.

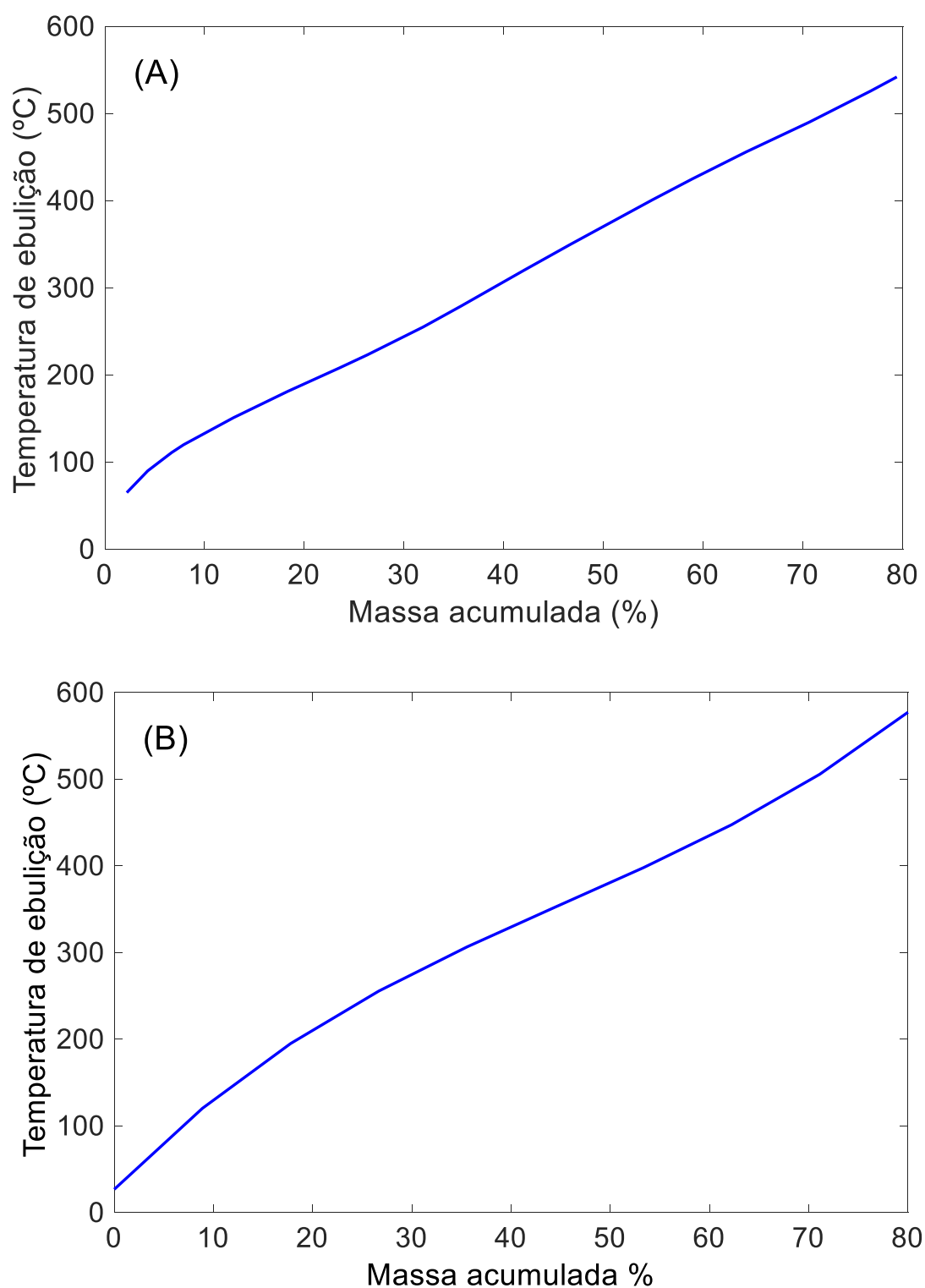


Figura A 7 Curvas PEV de amostra de petróleo. (A) Curva PEV Destilação física. (B) Curva de destilação simulada SIMDIS com fusão de dados DHA e SIMDIS. Amostra 053. Dados: API 30,1; PF $\leq$ -30,0 °C; PVR 29,6 kPa.

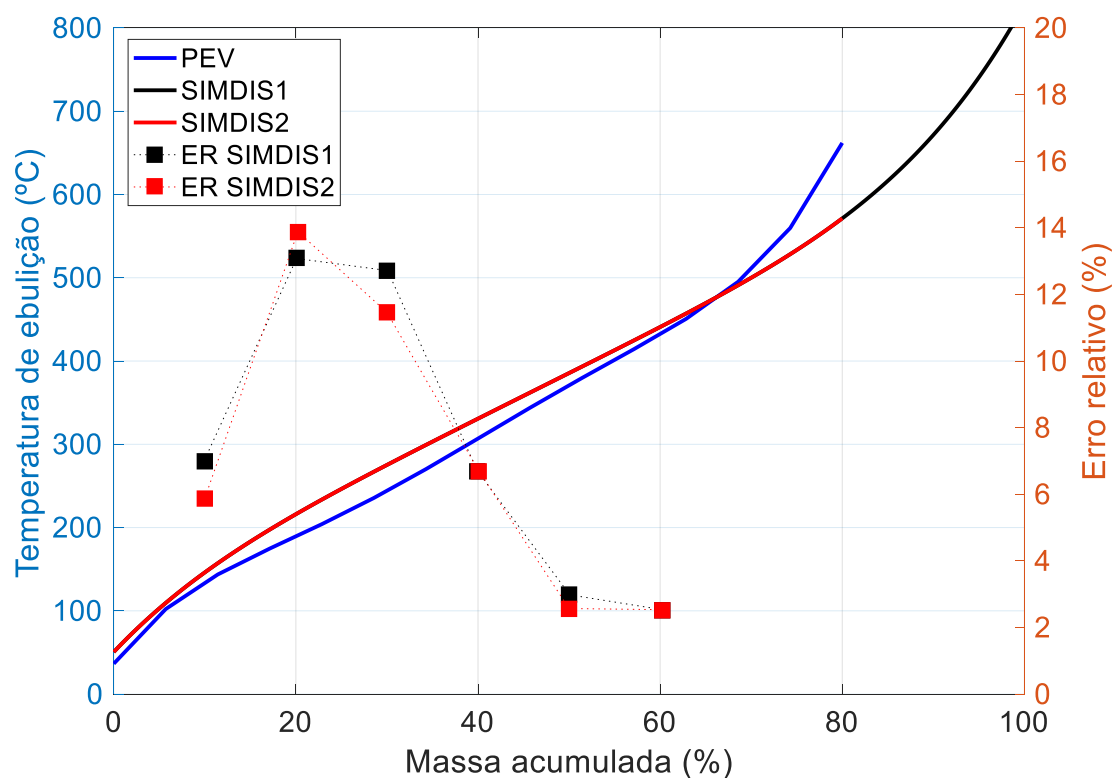


Figura A8 Gráfico de comparação entre erros relativos das curvas PEV, SIMDIS1 e SIMDIS2 de Amostra 053. PEV: Destilação física; SIMDIS1 : Destilação simulada por cromatografia gasosa SIMDIS; ER : Erro Relativo; SIMDIS2 : Destilação simulada pela fusão de dados DHA/SIMDIS.

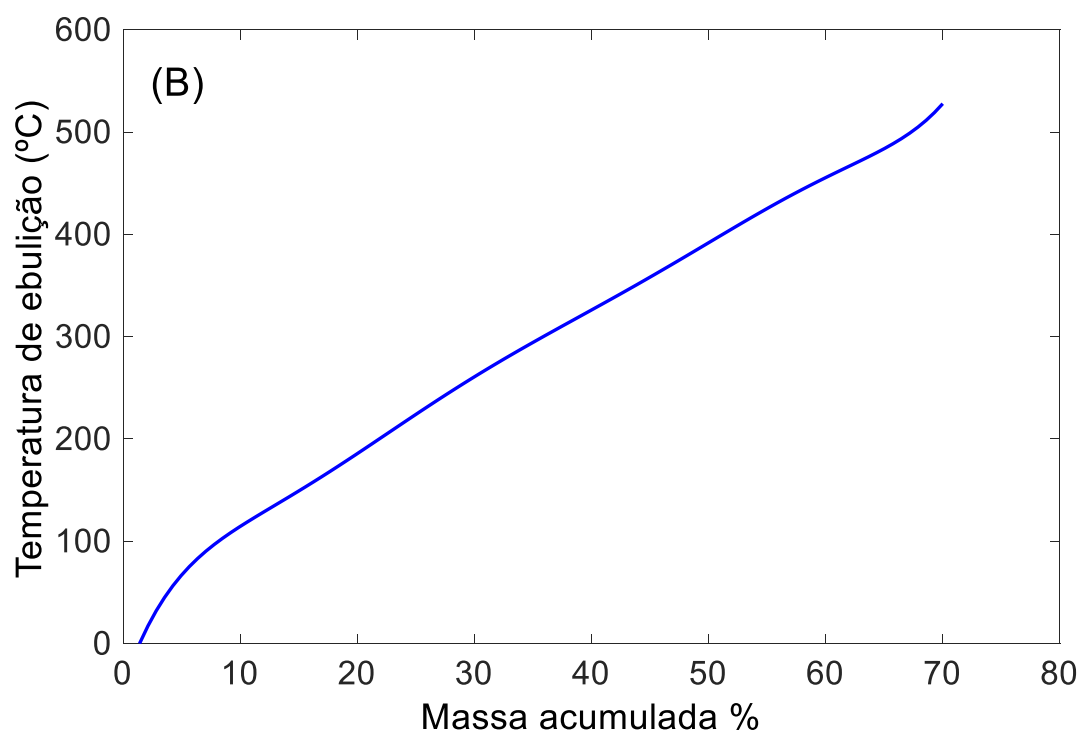
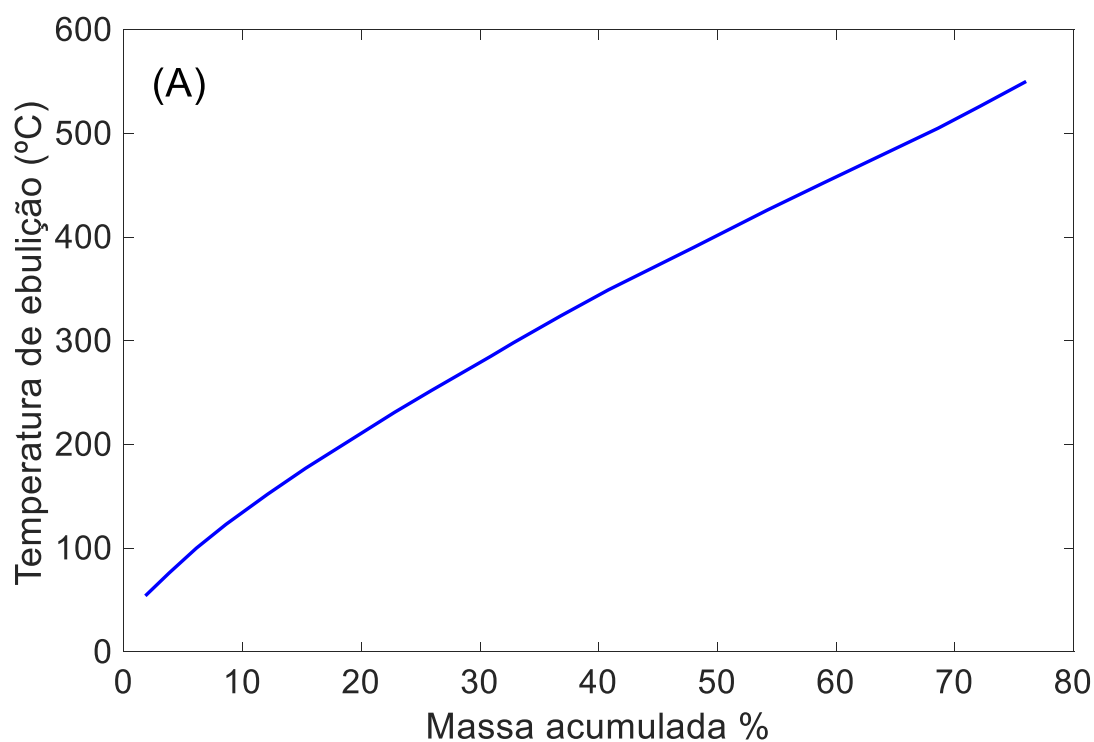


Figura A9 Curvas PEV de amostra de óleo bruto. (A) Curva PEV Destilação física. (B) Curva de destilação simulada SIMDIS com fusão de dados DHA e SIMDIS. Amostra 097. Dados: API 30,2; PF $\leq$ -30,0 °C; PVR 50,4 kPa.

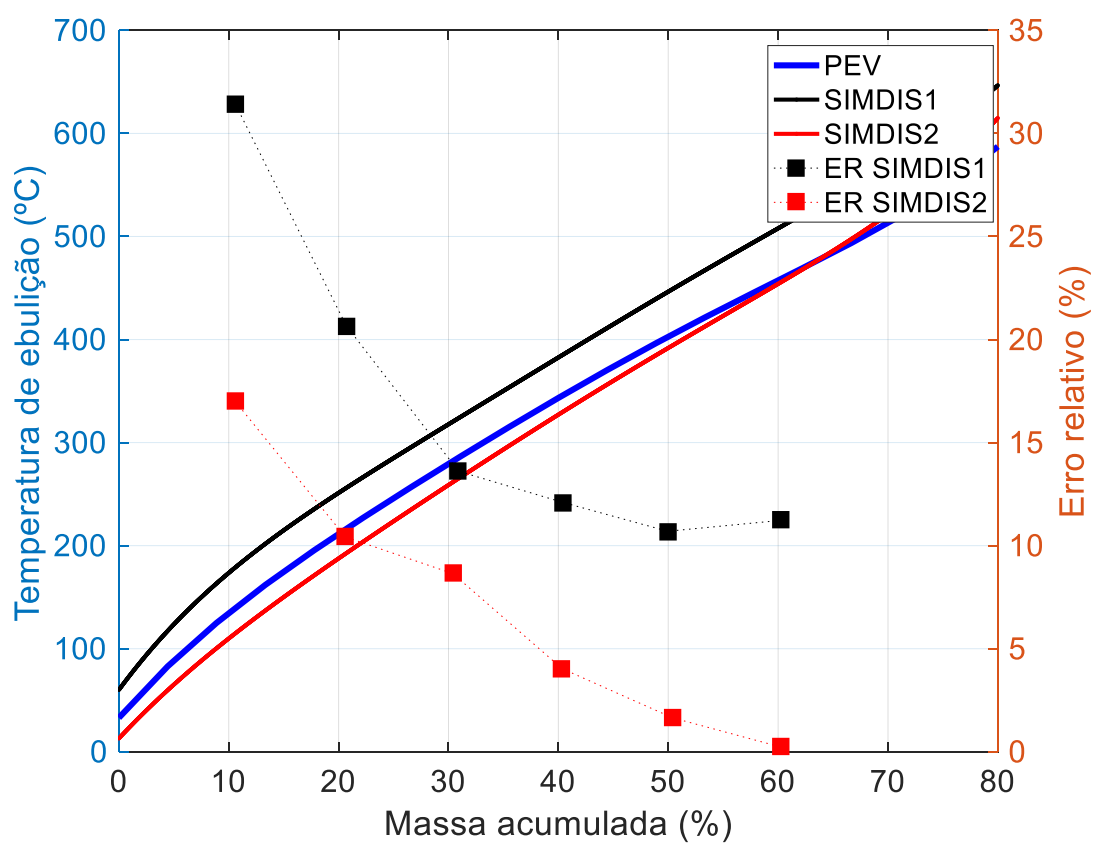


Figura A10 Gráfico de comparação entre erros relativos das curvas PEV, SIMDIS1 e SIMDIS2 de Amostra 097. PEV: Destilação física; SIMDIS1 : Destilação simulada por cromatografia gasosa SIMDIS; ER : Erro Relativo; SIMDIS2 : Destilação simulada pela fusão de dados DHA/SIMDIS.