



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA FUNDAMENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

dQF
Departamento
de Química Fundamental

NEIRIVALDO CAVALCANTE DA SILVA

**USO DA ESPECTROSCOPIA NIR NO DESENVOLVIMENTO DE UM
SIMULADOR PARA GASOLINA E NA TRANSFERÊNCIA DE CALIBRAÇÃO
ENTRE INSTRUMENTOS DE BANCADAS E PORTÁTIL**



RECIFE

2017

NEIRIVALDO CAVALCANTE DA SILVA

**USO DA ESPECTROSCOPIA NIR NO DESENVOLVIMENTO DE UM
SIMULADOR PARA GASOLINA E NA TRANSFERÊNCIA DE CALIBRAÇÃO
ENTRE INSTRUMENTOS DE BANCADAS E PORTÁTIL**

Tese apresentada ao Programa de Pós-Graduação em Química da Universidade Federal de Pernambuco como parte dos requisitos para a obtenção do título de Doutor em Química.

Orientadora:

Maria Fernanda Pimentel (Universidade Federal de Pernambuco, Brasil)

Coorientadores:

Fernanda Araújo Honorato (Universidade Federal de Pernambuco, Brasil)

Jose Manuel Amigo Rubio (Universidade de Copenhague, Dinamarca)

RECIFE

2017

Catálogo na fonte
Bibliotecária Joana D'Arc Leão Salvador CRB 4-572

S586u Silva, Neirivaldo Cavalcante da.
Uso da espectroscopia NIR no desenvolvimento de um simulador para gasolina e na transferência de calibração entre instrumentos de bancadas e portátil / Neirivaldo Cavalcante da Silva. – 2017.
123 f.: fig., tab.

Orientadora: Maria Fernanda Pimentel.
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN. Química Fundamental, Recife, 2017.
Inclui referências, apêndice e anexos.

1. Química analítica. 2. Infravermelho próximo. 3. Quimiometria. 4. Combustíveis. I. Pimentel, Maria Fernanda (Orientadora). II. Título.

543

CDD (22. ed.)

UFPE-FQ 2017-44

NEIRIVALDO CAVALCANTE DA SILVA

**USO DA ESPECTROSCOPIA NIR NO DESENVOLVIMENTO DE UM SIMULADOR
PARA GASOLINA E NA TRANSFERÊNCIA DE CAÇIBRAÇÃO ENTRE INSTRUMENTOS
DE BANCADAS E PORTÁTIL**

Tese apresentada ao Programa de Pós-Graduação no Departamento de Química Fundamental da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Química.

Aprovado em: 07 / 06 / 2017

BANCA EXAMINADORA

Profa. Maria Fernanda Pimentel Avelar (Orientador)

Departamento de Engenharia Química
Universidade Federal de Pernambuco

Profa. Claudete Fernandes Pereira
Departamento de Química Fundamental
Universidade Federal de Pernambuco

Prof. Luiz Stragevitch
Departamento de Engenharia Química
Universidade Federal de Pernambuco

Profa. Liliana Fátima Bezerra Lira
Departamento de Química
Universidade Federal da Paraíba

Profa. Simone da Silva Simões
Departamento de Química
Universidade Estadual da Paraíba

**Aos meus pais, os maiores
responsáveis pela minha formação, e
à minha esposa e melhor amiga.**

Com carinho dedico.

AGRADECIMENTOS

Tendo a consciência de que eu sozinho não poderia ter realizado tudo isso,

Eu agradeço a Deus, por todos os momentos de inspiração, e à minha família, meus pais, Erivaldo e Neide, e ao meu irmão Fagner, pela dedicação, cuidado e apoio que recebi durante todo o caminho percorrido até aqui, e pela compreensão do convívio algumas vezes sacrificado para a realização deste trabalho.

Agradeço de maneira muito especial à minha esposa Josiane por todo o companheirismo, compreensão, carinho, e constantes apoio e incentivo.

Agradeço à Profa. Dra. Maria Fernanda Pimentel, antes de tudo, por ter me dado a oportunidade de receber sua orientação, e por todas as discussões, conselhos e incentivo que foram essenciais para o meu desenvolvimento acadêmico e pessoal.

Agradeço à minha co-orientadora Profa. Dra. Fernanda Honorato por todas as contribuições, pelas discussões e apoio, que foram fundamentais.

Agradeço ao Prof. Dr. Jose Manuel Amigo pela valiosa co-orientação que recebi durante o período sanduíche na Universidade de Copenhague, Dinamarca.

Agradeço ao Dr. Márcio Rebouças que foi fundamental para a inicialização deste trabalho.

Agradeço à M.^a Ana Rosa, à Dra. Daniela Domingos e ao M.e Márcio Borges, por todo o suporte dado pelo laboratório de controle de qualidade da empresa parceira em um dos projetos.

Agradeço aos Professores Dr. Luiz Stragevitch, do Departamento de Engenharia Química da UFPE, Dra. Claudete Fernandes, do Departamento de Química Fundamental da UFPE, Dra. Simone Simões, do Departamento de Química da UEPB e Dra. Liliana Bezerra Lira, do Departamento de Química da UFPB, pela disponibilidade em fazer parte da banca examinadora.

Agradeço por todo o suporte dado pelo Laboratório de Combustíveis da UFPE (LAC), e a todos seus integrantes pelo agradável convívio diário.

Agradeço aos colegas de pós-graduação, especialmente, Alianda, Carolina, Eduardo, Leandro, Vitor e Francielson pelas frequentes colaborações e pelos conhecimentos compartilhados, conversas e cafezinhos vespertinos.

Agradeço aos professores do DQF-UFPE por todos os conhecimentos transmitidos.

Agradeço à CAPES, pela concessão da bolsa de doutorado, ao CNPQ, pela concessão da bolsa de doutorado sanduíche, e ao INCTAA, NUQAAPE e FACEPE por todos os auxílios concedidos.

RESUMO

Esta tese descreve dois trabalhos que abordam o desenvolvimento de metodologias analíticas, utilizando quimiometria e a espectroscopia NIR, para combustíveis. O primeiro aborda a transferência de calibração de um espectrômetro NIR de bancada para um NIR ultracompacto. Padrões virtuais foram criados por uma mistura matemática entre espectros de componentes puros da gasolina e misturas diesel/biodiesel e utilizados como amostras de transferência no método de Padronização Reversa (RS). Modelos de regressão por mínimos quadrados parciais (PLS) foram construídos para cinco parâmetros de qualidade da gasolina e um de misturas diesel/biodiesel. Os valores de RMSEP obtidos após os métodos de transferência foram equivalentes a reprodutibilidade dos métodos de referência para a maioria dos parâmetros. A metodologia demonstrou a viabilidade de transferir espectros de gasolina e misturas diesel/biodiesel adquiridos em um instrumento de bancada de alta resolução para um ultracompacto, utilizando padrões virtuais como amostras de transferência. O segundo trabalho envolve o processo de formulação da gasolina automotiva a partir das correntes de hidrocarbonetos obtidas no processo de craqueamento da nafta. O objetivo principal do trabalho foi desenvolver um simulador de octanagem de gasolina automotiva baseado na espectroscopia NIR e em ferramentas de quimiometria, com a finalidade de otimizar o processo de formulação. Os espectros foram adquiridos em dois equipamentos, um FT-NIR de bancada e um NIR ultracompacto. Modelos PLS foram desenvolvidos e validados para estimar a octanagem a partir de espectros simulados para potenciais formulações de gasolinas. A comparação entre as previsões dos valores de octanagem a partir dos espectros simulados, os espectros experimentais e os valores de referência, em diferentes tempos ao longo de 2 anos de produção, mostrou que as simulações foram capazes de fornecer valores preditos confiáveis. Ambos os espectrômetros forneceram resultados semelhantes, ratificando o potencial para utilização de instrumentos portáteis como mais uma ferramenta de otimização. O simulador de octanagem desenvolvido (NIRSIM) possui um alto potencial para implementação em processos de formulação de gasolina automotiva.

Palavras-chave: Infravermelho próximo. Combustível. Processo de formulação. Espectrômetro ultracompacto. Transferência de calibração. Padrões virtuais. Simulador off-line.

ABSTRACT

This thesis describes two works in which the NIR spectroscopy and chemometrics were employed in the development of methodologies for analyzing fuel. The first addresses the calibration transfer from a benchtop NIR spectrometer to a handheld NIR. Virtual standards were created by mathematically mixing spectra of pure components present in gasoline and diesel/biodiesel blends. The virtual standards were used as transfer samples for the reverse standardization (RS) method. Partial least squares regression (PLS) models were built for five quality parameters of gasoline and one of D/B blends. The RMSEP values obtained after the standardization approaches were equivalent to the reproducibility of the reference methods for most parameters. The methodology demonstrated the feasibility of transferring gasoline or D/B blend spectra acquired with a high-resolution benchtop instrument to a handheld NIR using virtual standards as transfer samples. The second work addresses the compounding process of automotive gasoline from hydrocarbon streams derived from the cracking process of naphtha. The main purpose of the work was to develop a NIR-based off-line simulator of octane number in automotive gasoline employing chemometrics in order to optimize the compounding process. Spectra were recorded using two instruments, a benchtop FT-NIR and a handheld NIR. PLS regression models were developed and validated to estimate the octane number from the simulated NIR spectra for potential formulations of gasolines. The comparison between the octane number predictions from the simulated spectra, the experimental spectra and, also, the reference values, at different times over 2 years of production, showed that the simulations were able to provide reliable predicted values. This work found out that the NIR spectra obtained both spectrometers provided similar results, confirming the potential for using portable instruments as a supplementary optimization tool. The developed octane simulator (NIRSIM) has an excellent potential for use in the automotive gasoline compounding processes.

Keywords: Near infrared. Fuel. Formulation process. Handheld spectrometer. Calibration transfer. Virtual standards. Off-line simulator.

LISTA DE FIGURAS

Figura 1. (A) Bandas sobrepostas, (B) primeira derivada e (C) segunda derivada. (Fonte: BRERETON, 2003 [56])	37
Figura 2. Esquema do procedimento matemático para criação dos padrões virtuais (acima) e transformação da matriz de componentes químicos puros por SVD (abaixo).....	54
Figura 3. Espectros brutos das amostras de gasolina adquiridos nos instrumentos primário e secundário para as amostras de gasolina (a.) e misturas diesel/biodiesel (b.).....	58
Figura 4. Espectros brutos dos componentes químicos puros presentes na composição da gasolina e misturas diesel/biodiesel, adquiridos nos instrumentos primário (a.) e secundário (b.).....	59
Figura 5. Espectros brutos para todas as amostras de gasolina (a.) e de misturas diesel/biodiesel (b.) adquiridos nos instrumentos primário e secundário na faixa de 11000 a 6000 cm^{-1}	60
Figura 6. Coeficientes de regressão dos modelos PLS do teor de biodiesel desenvolvidos no instrumento primário (a.) e secundário (b.) para a região espectral completa.....	61
Figura 7. Espectros brutos para os 12 componentes químicos puros adquiridos com os instrumentos primário (a.) e secundário (b.).....	62
Figura 8. Região de segundo sobretom de ligações C-H para os espectros pré-processados com SNV para todas as amostras de gasolina (a.) e diesel/biodiesel (b.). A região de segundo sobretom nos espectros dos componentes puros é mostrada em (c.) e (d.).....	63
Figura 9. Gráficos de valores preditos vs. referência para os modelos PLS construídos no instrumento primário e secundário na região de segundo sobretom de estiramentos de ligações C-H. As amostras de calibração estão em preto e de validação em vermelho.....	65
Figura 10. Padrões virtuais (verde) sobrepostos aos respectivos espectros experimentais de gasolina (a,b) e misturas diesel/biodiesel (c,d) para o instrumento primário (laranja) e secundário (azul).....	70
Figura 11. Comportamento dos valores de RMSEP em relação ao número de padrões após realizar o método de padronização reversa (RS) utilizando padrões experimentais. VL = Variável Latente; Repro: Reprodutibilidade do método de referência; RMSEP^P e RMSEP^S = Valores de RMSEP atingidos pelos modelos PLS construídos no instrumento primário (linha laranja) e secundário (linha azul), respectivamente.....	71
Figura 12. Valores preditos para os espectros de validação (vermelho) adquiridos no instrumento secundário, pelos modelos PLS construídos com os espectros do instrumento primário (preto).....	73
Figura 13. Gráfico de referência vs. valores preditos antes (esquerda) e após (direita) a correção SBC para os parâmetros T10, T50 e densidade. As 5 amostras utilizadas	

para realizar a correção estão circuladas nos gráficos. As amostras de calibração estão em preto e de valiração em vermelho.....	75
Figura 14. Espectros NIR das repetições em triplicata para a gasolina formulada em laboratório pelos analistas A e B. (a.) Espectros brutos e (b.) derivativos (1ª derivada com polinômio de 2º grau e janela de suavização de 3 pontos).....	77
Figura 15. Análise de componentes principais das repetições da gasolina M1 formulada em laboratório.....	78
Figura 16. Comparação entre os espectros das amostras de gasolina formuladas em laboratório (M1-M4) e os espectros simulados (S1-S4) a partir dos espectros de três correntes de hidrocarbonetos (T1-T3).....	80
Figura 17. Análise de componentes principais dos espectros experimentais (M) e simulados (S) de quatro gasolinas formuladas em laboratório.....	81
Figura 18. Espectros brutos das gasolinas e das correntes de hidrocarbonetos coletadas de 2015 a 2017 e analisadas no espectrometro de bancada e no ultracompacto.....	82
Figura 19. Espectros pré-processados das gasolinas coletados no espectrômetro de bancada (a.) e ultracompacto (b.). As cores azul e vermelho representam respectivamente as gasolinas do tipo 1 e 2.....	83
Figura 20. Análise de componentes principais dos espectros de gasolina automotiva adquiridos no espectrômetro de bancada. As cores azul e vermelho representam respectivamente as gasolinas do tipo 1 e 2.....	84
Figura 21. Análise de componentes principais dos espectros de gasolina automotiva adquiridos no espectrômetro ultracompacto. As cores azul e vermelho representam respectivamente as gasolinas do tipo 1 e 2.....	85
Figura 22. Comparação entre os espectros experimentais (vermelho) e as estimativas geradas pelos modelos CLS (sobrepostas em azul) no espectrômetro de bancada.....	87
Figura 23. Comparação entre os espectros experimentais (vermelho) e as estimativas geradas pelos modelos CLS (sobrepostas em azul) no espectrômetro ultracompacto....	88
Figura 24. Espectros do NIR de bancada de gasolinas Tipo 1 adquiridos do histórico de dados da companhia (preto) e coletados no período de 2015 a 2017 (azul).....	90
Figura 25. Modelo de regressão PLS para o parâmetro de octanagem MON, após atualização do modelo. (a.) Conjunto de calibração atual (azul) e do histórico (preto). (b.) Conjunto de calibração (preto) e previsão (vermelho) após atualização do modelo. (c.) Gráfico dos resíduos para amostras de calibração (preto) e previsão (vermelho). (d.) Gráfico dos escores para a importância das variáveis na projeção.....	92
Figura 26. Modelo de regressão PLS para o parâmetro de octanagem RON, após atualização do modelo. (a.) Conjunto de calibração atual (azul) e do histórico (preto). (b.) Conjunto de calibração (preto) e previsão (vermelho) após atualização do modelo. (c.) Gráfico dos resíduos para amostras de calibração (preto) e previsão (vermelho). (d.) Gráfico dos escores para a importância das variáveis na projeção.....	92
Figura 27. Modelo de regressão PLS para o parâmetro de octanagem IAD, após atualização do modelo. (a.) Conjunto de calibração atual (azul) e do histórico (preto).	

(b.) Conjunto de calibração (preto) e previsão (vermelho) após atualização do modelo. (c.) Gráfico dos resíduos para amostras de calibração (preto) e previsão (vermelho). (d.) Gráfico dos escores para a importância das variáveis na projeção.....	93
Figura 28. Modelos de regressão PLS desenvolvidos no NIR de bancada para os parâmetros de octanagem MON, RON e IAD. (a., b. e c.) Gráfico dos resíduos; (d., e. e f.) Gráfico dos valores preditos vs. referência; (g., h. e i.) Gráfico dos escores para a importância das variáveis na projeção; para MON, RON e IAD, respectivamente. Amostras de calibração (preto) e previsão (vermelho).....	95
Figura 29. Modelos de regressão PLS desenvolvidos no NIR ultracompacto para parâmetros de octanagem MON (esquerda) e RON (direita). (a., b. e c.) Gráfico dos resíduos; (d., e. e f.) Gráfico dos valores preditos vs. referência; (g., h. e i.) Gráfico dos escores para a importância das variáveis na projeção; para MON, RON e IAD, respectivamente. Amostras de calibração (preto) e previsão (vermelho).....	96
Figura 30. Diagrama de fluxo para o funcionamento do simulador de octanagem NIRSIM.....	102
Figura 31. Interface gráfica do simulador de octanagem NIRSIM para gasolina automotiva.....	103
Figura 32. Simulação de octanagem para gasolinas do Tipo 1 utilizando espectros de correntes de hidrocarbonetos adquiridos no NIR de bancada.....	104
Figura 33. Simulação de octanagem para gasolinas do Tipo 2 utilizando espectros de correntes de hidrocarbonetos adquiridos no NIR de bancada. (--- Limite máximo para a faixa de calibração do modelo PLS para o parâmetro de octanagem MON).....	105
Figura 34. Simulação de octanagem para gasolinas do Tipo 2 utilizando espectros de correntes de hidrocarbonetos adquiridos no NIR ultracompacto. (--- Limite máximo para a faixa de calibração do modelo PLS para o parâmetro de octanagem RON).....	106

LISTA DE TABELAS

Tabela 1. Figuras de mérito para os modelos PLS construídos com a faixa espectral completa e com a região de segundo sobretom.....	63
Tabela 2. Figuras de mérito para os modelos PLS construídos na região de segundo sobretom após seleção de variáveis.....	66
Tabela 3. Reprodutibilidades dos métodos de referência.....	67
Tabela 4. Figuras de mérito após realizar a padronização reversa utilizando padrões experimentais ou padrões reais.....	72
Tabela 5. Comparação entre os modelos PLS para prever octanagem em gasolinas do Tipo 1, antes e depois da atualização.....	91
Tabela 6. Comparação entre os modelos PLS para prever parâmetros de octanagem em gasolinas do Tipo 2.....	93
Tabela 7. Comparação entre os valores de octanagem do método de referencia e previsões PLS a partir de espectros experimentais e simulados.....	97
Tabela 8. Previsões PLS (NIR de bancada) dos valores de octanagem nos espectros simulados para as variações nas proporções das correntes de hidrocarbonetos	100
Tabela 9. Previsões PLS (NIR ultracompacto) dos valores de octanagem nos espectros simulados para as variações nas proporções das correntes de hidrocarbonetos.....	101

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
AEAC	Álcool Etílico Anidro Combustível
ANNs	Redes Neurais Artificiais
ANP	Agência Nacional de Petróleo
ASTM	<i>American Society for Testing and Materials</i>
Bias	Erro Sistemático
BX	Mistura diesel/biodiesel com X% de biodiesel.
CCA	<i>Canonical Correlation Analysis</i>
CIMA	Conselho Interministerial do Açúcar e do Álcool
CLS	<i>Classical Least Squares</i>
C.R.F	<i>Cooperative Fuel Research</i>
DOP	<i>Dynamic Orthogonal Projection</i>
DS	<i>Direct Standardization</i>
DWPDS	<i>Double Window Piecewise Direct Standardization</i>
EPO	<i>External Parameter Orthogonalisation</i>
EROS	<i>Error Removal by Orthogonal Subtraction</i>
FT-NIR	Infravermelho Próximo por Transformada de Fourier
IAD	Índice Antidetonante
IP	Instrumento Primário
iPLS	<i>Interval Partial Least Squares</i>
IS	Instrumento Secundário
JK	<i>Jack-Knife</i>
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
MCR-ALS	<i>Multivariate Curve Resolution - Alternate Least Squares</i>
MIR	Espectroscopia no Infravermelho Médio
MLR	Regressão Linear Múltipla
MON	<i>Motor Octane Number</i>
MSC	Correção Multiplicativa do Sinal
MU	<i>Model Update</i>
NIR	Espectroscopia no Infravermelho Próximo
O-PLS	<i>Orthogonal Projections to Latent Structures</i>
OSC	<i>Orthogonal Signal Correction</i>
PAT	<i>Process Analytical Technologies</i>
PC	Componente Principal
PCA	Análise de Componentes Principais
PCR	Regressão por Componentes Principais
PDS	<i>Piecewise Direct Standardization</i>
PLS	Mínimos Quadrados Parciais
PLS-DA	Análise Discriminante por Mínimos Quadrados Parciais
PMQC	Programa de Monitoramento da Qualidade dos Combustíveis
PRS	<i>Piecewise Reverse Standardization</i>

PV	<i>Padrões Virtuais</i>
RF	<i>Random Forest</i>
RMSEP	Raiz Quadrada do Erro Médio Quadrático de Previsão
RMSEC	Raiz Quadrada do Erro Médio Quadrático de Calibração
RMSECV	Raiz Quadrada do Erro Médio Quadrático de <i>Cross-Validation</i>
RON	<i>Research Octane Number</i>
RS	<i>Reverse Standardization</i>
SBC	<i>Slope/bias correction</i>
SDW	<i>Spectral difference by wavelengths</i>
SNV	Varição Normal Padrão
SPXY	<i>Sample Set Partitioning based on Joint x-y Distances</i>
SVD	Decomposição de Valores Singulares
SVM	<i>Support Vector Machines</i>
SW-NIR	<i>Short-wavelength Near Infrared</i>
TOP	<i>Transfer by Orthogonal Projection</i>
TPF	Temperatura no ponto final de ebulição
UV-vis	Ultravioleta-Visível
VL	Variável Latente
WHDS	<i>Wavelet Hybrid Direct Standardization</i>

SUMÁRIO

1	INTRODUÇÃO	20
1.1	Aspectos sobre a produção e fiscalização da gasolina automotiva no Brasil	21
1.1.1	Predição de parâmetros da qualidade da gasolina automotiva no processo de formulação utilizando amostras simuladas	21
1.1.2	Características e controle de qualidade da gasolina e do BX no Brasil	22
1.1.3	Técnicas de transferência de calibração envolvendo instrumentos NIR ultracompactos para análise de combustíveis	24
1.2	Objetivos	26
1.2.1	Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais	26
1.2.2	Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR	26
2	FUNDAMENTAÇÃO TEÓRICA	27
2.1	Revisão da literatura	27
2.1.1	Predição de propriedades químicas em misturas de tipos diferentes de petróleo utilizando espectros NIR experimentais e simulados	27
2.1.2	Determinação de parâmetros de qualidade de gasolina e misturas diesel/biodiesel utilizando a espectroscopia no infravermelho próximo ..	28
2.1.3	Transferência de calibração utilizando espectrômetros de infravermelho próximo de bancada e portátil	28
2.1.4	Introdução do conceito de padrões virtuais	31
2.2	Espectroscopia de Infravermelho Próximo (NIR)	32

2.3	Calibração Multivariada	34
2.3.1	Pré-processamento de dados	35
2.3.1.1	PRÉ-PROCESSAMENTO NAS VARIÁVEIS	35
2.3.1.2	PRÉ-PROCESSAMENTO NAS AMOSTRAS	35
2.3.2	Seleção de Variáveis	37
2.3.2.1	MÉTODO <i>JACK-KNIFE</i>	37
2.3.2.2	<i>INTERVAL</i> PLS (<i>i</i> PLS)	38
2.3.3	Seleção de Amostras (SPXY)	38
2.3.4	Regressão por Mínimos Quadrados Clássicos (CLS)	39
2.3.5	Regressão por Mínimos Quadrados Parciais (PLS)	40
2.3.6	Avaliação dos modelos de calibração multivariada	41
2.4	Transferência de Calibração	45
2.5	Programa de monitoramento da qualidade de combustíveis (PMQC)	46
3	MATERIAIS E MÉTODOS	49
3.1	Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais	49
3.1.1	Instrumentação	49
3.1.2	Amostras e reagentes	49
3.1.3	Aquisição dos espectros NIR	50
3.1.4	Modelos de regressão PLS	50
3.1.5	Método de transferência de calibração	51
3.1.5.1	PROCEDIMENTO PARA CRIAÇÃO DOS PADRÕES VIRTUAIS	52

3.2	Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR	55
3.2.1	Instrumentação	55
3.2.2	Amostras	55
3.2.2.1	GASOLINAS	56
3.2.2.2	CORRENTES DE HIDROCARBONETOS	56
3.2.3	Aquisição dos espectros NIR	56
3.2.4	Modelos de regressão PLS	57
3.3	Softwares	57
4	RESULTADOS E DISCUSSÃO	58
4.1	Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais	58
4.1.1	Avaliação da faixa espectral	60
4.1.2	Avaliação do efeito da seleção de variáveis no desempenho dos modelos PLS na região de segundo sobretom, para os instrumentos de bancada e ultracompacto	66
4.1.3	Comparação dos modelos PLS entre os dois instrumentos e com os métodos de referência	68
4.1.4	Procedimento para transferência de calibração	69
4.2	Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR	76
4.2.1	Experimentos preliminares utilizando gasolinas formuladas em laboratório	76

4.2.1.1	VARIABILIDADE ESPECTRAL PARA REPETIÇÕES DE UMA GASOLINA FORMULADA EM LABORATÓRIO	77
4.2.1.2	AVALIAÇÃO DO COMPORTAMENTO DOS ESPECTROS SIMULADOS EM RELAÇÃO AOS EXPERIMENTAIS	79
4.2.2	Apresentação e análise exploratória dos conjuntos de dados espectrais nos espectrômetros de bancada e ultracompacto	81
4.2.3	Prevendo a formulação a partir dos espectros de gasolinas e das respectivas correntes de hidrocarbonetos	86
4.2.4	Desenvolvimento dos modelos PLS para prever os valores de octanagem, MON e RON	89
4.2.4.1	MODELO PLS (GASOLINA DO TIPO 1)	90
4.2.4.2	MODELO PLS (GASOLINA DO TIPO 2)	93
4.2.5	Efeito da variação das proporções entre as correntes de hidrocarbonetos na octanagem predita pelos modelos PLS	96
4.2.6	Desenvolvimento do simulador de octanagem para gasolina automotiva	102
5	CONCLUSÕES	107
5.1	Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais	107
5.2	Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR	108
6	PERSPECTIVAS	110

6.1	Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais	110
6.2	Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR	110
	REFERÊNCIAS	111
	APÊNDICE A - Rotina em Matlab da função “simugas” do simulador NIRSIM	119
	ANEXO A – Tabela dos parâmetros de qualidade da gasolina automotiva comum	121
	ANEXO B – Tabela dos parâmetros de qualidade do óleo diesel BX a B30	122

1 INTRODUÇÃO

Nesta tese está descrito o desenvolvimento de dois trabalhos envolvendo a utilização de métodos quimiométricos e a espectroscopia NIR para determinação de parâmetros de qualidade de combustíveis, sob duas diferentes perspectivas.

O primeiro trabalho surgiu a partir do interesse em análise de campo utilizando instrumentos NIR portáteis. A ideia central deste trabalho está diretamente ligada à possibilidade de transferência de calibração entre um instrumento NIR de bancada de alta resolução e um instrumento NIR ultracompacto. A maioria das técnicas de transferência necessita de um conjunto de amostras de transferência cujos espectros devem ser analisados em ambos os espectrômetros. Esse é um ponto crítico quando se trata de amostras de combustíveis, uma vez que a alta volatilidade de seus componentes torna a composição instável. A alternativa encontrada para este problema foi a substituição deste conjunto de amostras por espectros criados a partir da mistura matemática dos espectros de compostos puros presentes na composição da gasolina e de misturas de óleo diesel/biodiesel. O método desenvolvido com a utilização destes espectros, chamados de padrões virtuais, pode evitar problemas com volatilização e mudanças de composição das amostras durante o armazenamento e transporte de amostras de combustíveis.

O segundo trabalho envolve o processo industrial de formulação de gasolina automotiva, que consiste em uma mistura entre correntes de hidrocarbonetos que são derivadas do processo de craqueamento da nafta. Nesse processo, a qualidade e rentabilidade da gasolina produzida são diretamente afetadas pela composição das correntes de hidrocarbonetos e pela proporção em que elas são utilizadas. Os diversos fatores que podem afetar as características físico-químicas das correntes de hidrocarbonetos utilizadas na formulação, como a origem e o tipo de processamento do petróleo, fazem com que a predição da octanagem da gasolina produzida seja bastante complicada. Com a utilização da espectroscopia NIR, é possível detectar mudanças de composição rapidamente, e considerá-las durante a formulação de uma gasolina. O foco deste trabalho foi no desenvolvimento de um simulador de octanagem de gasolina automotiva off-line baseado na espectroscopia NIR e em ferramentas de quimiometria com a finalidade de otimizar o processo de formulação. Com a utilização da espectroscopia NIR, o tempo de análise dos parâmetros de interesse para o processo pode ser significativamente reduzido, comparando com os métodos analíticos que são normalmente utilizados.

1.1 Aspectos sobre a produção e fiscalização da gasolina automotiva no Brasil

1.1.1 Predição de parâmetros da qualidade da gasolina automotiva no processo de formulação utilizando amostras simuladas

A gasolina automotiva é um produto tradicionalmente obtido como um subproduto da destilação fracionada do petróleo. Entretanto, atualmente é obtida tanto nas refinarias como nas centrais petroquímicas e em formuladores, por meio de um processo de formulação. Em outras palavras, a gasolina formulada é o produto de uma mistura entre diferentes correntes de hidrocarbonetos, obtidas em fases intermediárias do processo (no caso das refinarias e petroquímicas) ou adquiridas no mercado (no caso dos formuladores). Essas misturas são realizadas de forma criteriosa para que o padrão de qualidade desejado seja obtido, levando em conta os parâmetros de desempenho e eficiência em motores de combustão interna com ignição por centelha. Além disso, fatores como o custo de produção e a rentabilidade do processo também influenciam o processo de formulação da gasolina.

De acordo com a ANP (Resolução ANP Nº 40 DE 25/10/2013), a gasolina tipo A “é um produto produzido a partir de processos utilizados no refino de petróleo, nas centrais petroquímicas e formuladores, destinado aos veículos automotivos dotados de motores ciclo Otto, isento de componentes oxigenados”, ou seja, antes de receber a adição de etanol. A composição final de uma gasolina tipo A é bastante complexa e variável, uma vez que compreende uma grande quantidade de hidrocarbonetos contendo grupos funcionais e tamanhos de cadeia diferentes. Além disso, também existe variabilidade nas características físico-químicas das correntes de hidrocarbonetos utilizadas na formulação dependendo da região de origem ou do processo pelo qual foram obtidas. Esses fatores, entre outros, representam um obstáculo importante no desenvolvimento de procedimentos para estimar com antecedência os parâmetros de qualidade e desempenho de uma gasolina final com base nas correntes de hidrocarbonetos utilizadas na formulação. Apesar disso, existem alguns simuladores que podem ser comercialmente encontrados, desenvolvidos para estimar parâmetros de qualidade da gasolina empregando características físico-químicas das correntes petroquímicas utilizadas na formulação.

O custo para aquisição desses simuladores é geralmente elevado, no entanto, essa não é a principal desvantagem de sua aplicação às gasolinas brasileiras. Geralmente, eles são calibrados com base em dados obtidos experimentalmente em países cujas gasolinas

formuladas e correntes petroquímicas possuem composições químicas com características muito diferentes das que são processadas no Brasil. O desempenho de um simulador é prejudicado neste caso, resultando em um procedimento onde as formulações simuladas precisam ser validadas em laboratório antes de serem utilizadas definitivamente na planta industrial. Em outras palavras, este tipo de procedimento exige mão-de-obra laboratorial e mantém um tempo de resposta elevado para definição da formulação mais adequada. O desenvolvimento de um simulador construído com base nas características das gasolinas comercializadas no Brasil é um objeto de pesquisa importante e representa uma contribuição relevante para o setor de produção de combustíveis no país.

Com base na propriedade aditiva da lei de Beer, espectros NIR de gasolina podem ser simulados a partir dos espectros das correntes de hidrocarbonetos e, dessa forma, modelos de regressão PLS podem ser utilizados para realizar previsões de parâmetros de qualidade a partir dessas simulações para potenciais formulações de gasolinas. Partindo desse ponto fundamental, algumas ferramentas podem ser desenvolvidas com base em estratégias de Tecnologias Analíticas de Processos (Process Analytical Technologies - PAT) [1], para auxiliar o controle de qualidade da gasolina e reduzir o tempo de resposta para determinação dos parâmetros de qualidade. Apesar deste estudo estar sendo conduzido em amostras de gasolina, as metodologias propostas possuem potencial para aplicação em outros tipos de produtos formulados derivados de petróleo, como o diesel.

1.1.2 Características e controle de qualidade da gasolina e do BX no Brasil.

De acordo com o último Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis, divulgado em 2016 [2], um volume de 124 milhões de m³ de derivados de petróleo foi comercializado no Brasil em 2015. A gasolina C que é responsável por 33% desse montante, ficando atrás apenas do óleo diesel com 46,2%, é um combustível utilizado principalmente como fonte de energia para automóveis. Fatores como o tipo de processo de produção, a origem do petróleo, a finalidade do uso e, até mesmo, a legislação local exercem grande influência sobre a composição final de uma gasolina. Apesar de o Brasil possuir uma capacidade nominal para produção de biodiesel de aproximadamente 7,4 milhões de m³, apenas cerca de 3,9 milhões de m³ foram produzidos em 2015. O crescimento na produção de biodiesel no Brasil é impulsionado principalmente por fatores econômicos e ambientais. Esse crescimento é bastante notável, saindo de um montante de 0,07 milhões de m³ em 2006,

após a introdução do biodiesel na matriz energética do país, para quase 4 milhões de m³ em 2015, quando o biodiesel já era obrigatoriamente adicionado em uma percentagem de 7% ao óleo diesel combustível.

Um grupo de diversos hidrocarbonetos com tamanhos de cadeia que variam entre 4 e 12 átomos de carbono representam a maior parte da composição da gasolina. Essa mistura de hidrocarbonetos faz com que a temperatura de ebulição normal (1atm) da gasolina varie ao longo de uma ampla faixa que vai de 30 até 220 °C. Esses compostos são essencialmente classificados como hidrocarbonetos parafínicos, olefínicos, naftênicos ou aromáticos, porém, outros compostos contendo enxofre, oxigênio e nitrogênio, também estão presentes em baixas concentrações.

A composição do biodiesel, ou B100, é formada por ésteres mono-alquílicos de ácidos graxos de cadeia longa derivados de óleo vegetais ou gorduras animais. A produção do biodiesel é feita, em geral, por meio de uma reação entre um óleo vegetal ou gordura animal e um álcool, como o etanol e o metanol, na presença de um catalisador. Assim, os principais produtos dessa reação são os ésteres mono-alquílicos. A glicerina também é obtida como um subproduto e separada do biodiesel. O óleo diesel, por sua vez, é um produto mineral obtido na destilação do petróleo. Dessa forma, o termo BX é utilizado para designar uma mistura de X% de biodiesel no óleo diesel derivado de petróleo. O B7 é o óleo diesel automotivo comercializado atualmente no Brasil e a percentagem de biodiesel no diesel é estabelecida pela **Lei Nº 13263, de 23/03/16**, que também estabelece que o percentual de biodiesel no diesel deve atingir 10% em 2019.

Além de todos esses compostos, alguns aditivos também são adicionados à gasolina e ao BX, com a finalidade de aumentar o desempenho e vida útil dos motores [3,4]. Na gasolina comum, o percentual de etanol passou a ser de 27% a partir de 16 de março de 2015, de acordo com o que estabelece a Portaria Nº 75, de 05/03/2015 do Ministério da Agricultura, Pecuária e Abastecimento (MAPA) e pela Resolução Nº 1, de 04/03/2015, do Conselho Interministerial do Açúcar e do Álcool (CIMA) [5].

Para realizar o controle de qualidade da gasolina e do BX distribuídos e comercializados nos postos de abastecimento, a Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP), responsável pela fiscalização da qualidade dos combustíveis no Brasil, coordena um programa de monitoramento (Programa de Monitoramento da Qualidade dos Combustíveis - PMQC) [6] que conta com diversos laboratórios conveniados espalhados pelo país. A ANP estabelece, através de portarias e resoluções, especificações

para uma série de parâmetros físico-químicos que têm impacto direto nas propriedades desses combustíveis. Dessa forma, os limites de especificação destes parâmetros garantem a qualidade adequada para promover o desempenho desejado nos motores automotivos e atingir os níveis estabelecidos de emissão de poluentes.

De acordo com a ANP, as análises dos parâmetros de qualidade da gasolina e do óleo diesel combustível devem seguir métodos normalizados estabelecidos pela Associação Brasileira de Normas Técnicas (ABNT) e pela *American Society for Testing and Materials* (ASTM) [7]. Esses métodos envolvem várias técnicas analíticas que, apesar de serem métodos bem estabelecidos e com desempenhos comprovados, em sua maioria são demorados, consomem quantidades consideráveis de reagentes e exigem que a amostra seja transportada até o laboratório. Esses fatores, entre outros, estimulam o interesse pelo desenvolvimento de métodos analíticos alternativos potencialmente viáveis para determinação destes parâmetros.

A aplicação de espectrômetros de infravermelho próximo (NIR) de bancada para determinar parâmetros de qualidade da gasolina e misturas diesel/biodiesel (d/b) já é muito bem estabelecida e está descrita em diversos trabalhos na literatura [3,8-18]. Por outro lado, a utilização dos espectrômetros NIR portáteis de baixo custo para desempenhar essa tarefa em campo é pouco frequente.

1.1.3 Técnicas de transferência de calibração envolvendo instrumentos NIR ultracompactos para análise de combustíveis.

Um dos requisitos para se desenvolver modelos de calibração multivariada com um nível adequado de confiabilidade é a aquisição de um número representativo de amostras. Porém, isso nem sempre é possível dentro de um curto período de tempo, especialmente quando se trata de matrizes complexas, como é o caso dos combustíveis. O custo atribuído à determinação dos parâmetros de qualidade de amostras de combustíveis a partir dos métodos normalizados é, geralmente, elevado e se torna um fator ainda mais crítico considerando que os modelos de calibração multivariada ocasionalmente necessitam de atualização. Além disso, o modelo construído em um determinado equipamento não pode ser diretamente aplicado a outro antes que seja realizada uma prévia adaptação entre os equipamentos, como é explicado no tópico 2.4.

Os espectrômetros portáteis NIR, bem como o ultracompacto MicroNIR, estão cada vez mais ganhando espaço no mercado ao mesmo tempo em que o interesse por métodos de baixo custo capazes de realizar medidas em campo vem aumentando. A utilização de espectrômetros para estimar parâmetros de qualidade de combustíveis diretamente nos postos de abastecimento representa uma aplicação com alto potencial como uma ferramenta para o monitoramento e supervisão da qualidade dos combustíveis comercializados no Brasil. Para ter uma visão geral da viabilidade desse método, a seguinte situação prática pode ser proposta: Inicialmente adquirir espectros de um grupo de amostras representativas para construir um modelo em um instrumento de bancada (primário) e em seguida transferir esses espectros para equipamentos portáteis (secundários) que podem ser utilizados para realizar análises nos diversos pontos de fiscalização do país.

A maioria das técnicas de transferência de calibração requer um grupo de amostras cujos espectros sejam adquiridos em ambos os instrumentos, primário e secundário. Essas amostras são chamadas de amostras de transferência e devem representar a variabilidade inerente a todo o conjunto de calibração. Dessa forma, os conjuntos de espectros adquiridos caracterizam as diferenças entre as respostas destes instrumentos que, por sua vez, podem ser modeladas e corrigidas. Portanto, o sucesso do procedimento de transferência de calibração depende diretamente da representatividade das amostras de transferência, o que também torna essencial a estabilidade da composição dessas amostras. Contudo, quando se trata de análise de combustíveis, a presença de hidrocarbonetos altamente voláteis nessas amostras pode causar mudanças na sua composição durante o transporte ou ao longo do tempo. A exposição à luz, condições de temperatura e contato com o oxigênio presente no ar, também podem contribuir para a degradação natural da gasolina por meio do processo de oxidação. Além disso, o procedimento para transporte de amostras de combustíveis dentro do país é bastante complicado devido a restrições impostas pelas autoridades de segurança.

Com base na metodologia descrita por Cooper *et.al.* [19-21], é possível utilizar espectros de combustíveis matematicamente construídos como uma alternativa para substituir o conjunto de amostras de transferência. Todos os instrumentos envolvidos nos trabalhos descritos por Cooper *et al.* são da mesma marca e possuem a mesma configuração instrumental. A metodologia proposta nesta tese, por outro lado, envolve um espectrômetro de bancada e um portátil que possuem configurações instrumentais e espectrais bastante diferentes. Devido a isso, a padronização das respostas instrumentais desses equipamentos é

um pouco mais complexa e alguns desafios precisam ser alcançados durante a utilização dos padrões virtuais.

Dessa forma, tornar este procedimento viável significa evitar questões inconvenientes relacionadas ao transporte e armazenamento de amostras reais de combustíveis. Até o presente momento, não foram encontrados na literatura trabalhos envolvendo a aplicação de técnicas de transferência de calibração, seja utilizando amostras de transferência reais ou virtuais, para transferir modelos de calibração multivariada, de parâmetros de gasolina e misturas diesel/biodiesel, de equipamentos NIR de bancada para NIR portáteis.

1.2 Objetivos

1.2.1 Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais

- Avaliar o desempenho de modelos PLS desenvolvidos em um NIR ultracompacto para prever parâmetros de qualidade da gasolina e de misturas diesel/biodiesel.
- Empregar padrões virtuais como amostras de transferência no método de Padronização Reversa (RS) para corrigir respostas instrumentais de dois espectrômetros de infravermelho próximo com características instrumentais bastante diferentes.

1.2.2 Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR

- Desenvolver um simulador off-line para explorar potenciais formulações de gasolina, de acordo com os parâmetros de qualidade e desempenho especificados, a partir dos espectros NIR das correntes petroquímicas e das proporções apropriadas.
- Criar e manipular espectros NIR de gasolinas por meio da adição matemática linear dos espectros NIR das correntes de hidrocarbonetos utilizadas no processo de produção de gasolina de uma companhia petroquímica brasileira.
- Desenvolver uma interface de fácil utilização incorporando os procedimentos quimiométricos criados, visando possibilitar a utilização por usuários que não possuem conhecimento específico em quimiometria.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Revisão da literatura

2.1.1 Predição de propriedades químicas em misturas de tipos diferentes de petróleo utilizando espectros NIR experimentais e simulados

Em um trabalho publicado em 2009, Peinder *et al.* [22] demonstraram a viabilidade da predição de propriedades do resíduo atmosférico de petróleo em 16 espectros no infravermelho médio de misturas de petróleo obtidos por uma adição matemática linear entre espectros de 10 diferentes petróleos brutos, utilizando modelos de regressão por mínimos quadrados parciais (PLS). Além disso, o estudo mostra que os valores obtidos nesta predição são semelhantes aos valores preditos a partir dos espectros adquiridos nas 16 misturas correspondentes que foram fisicamente preparadas. Este trabalho demonstra que o procedimento de mistura entre espectros de infravermelho de petróleo pode ser usado como ferramenta para explorar misturas potenciais de forma mais rápida e limpa, o que representa uma alternativa de baixo custo para os métodos normalmente utilizados que envolvem a preparação de misturas reais em laboratório.

Os procedimentos para adição matemática dos espectros de infravermelho médio foram realizados de acordo com a propriedade aditiva estabelecida pela Lei de Beer. Com base nisso, em princípio, é possível reproduzir matematicamente o espectro de uma gasolina, que é uma mistura controlada entre correntes de hidrocarbonetos, a partir de espectros de infravermelho próximo (NIR) das correntes e das proporções apropriadas para a formulação. Algumas diferenças, entretanto, podem ser observadas entre o espectro adquirido nas misturas físicas e os espectros simulados, mesmo após realizar o pré-processamento espectral. Isso pode estar relacionado com efeitos não lineares, entre os quais, as mudanças de índice de refração que ocorrem naturalmente durante as misturas fisicamente preparadas. Portanto, para minimizar possíveis diferenças relacionadas a esses efeitos, algumas técnicas de transferência de calibração podem ser utilizadas.

2.1.2 Determinação de parâmetros de qualidade de gasolina e misturas diesel/biodiesel utilizando a espectroscopia no infravermelho próximo

Na literatura já existem diversos trabalhos que utilizam espectrômetros de infravermelho próximo (NIR) de bancada para determinar parâmetros de qualidade da gasolina e misturas diesel/biodiesel (d/b). Dessa forma, a utilização da espectroscopia NIR para esta aplicação é bem estabelecida e está associada à diversas técnicas de análise multivariada. Entre elas, Regressão por Componentes Principais (*Principal Component Regression - PCR*), Regressão Linear Múltipla (*Multiple Linear Regression - MLR*), Regressão por Mínimos Quadrados Parciais (*Partial Least Squares - PLS Regression*), Redes Neurais Artificiais (*Artificial Neural Networks - ANN*), *Random Forest (RF)*, Máquina de Vetores de Suporte (*Support Vector Machines - SVM*), etc. [3,8-18]. A espectroscopia NIR é uma técnica rápida, não destrutiva e dispensa o pré-tratamento de amostras, portanto, representa uma alternativa em potencial para os métodos normalizados utilizados atualmente para determinação de parâmetros de qualidade da gasolina.

Em contrapartida, a quantidade de trabalhos que utilizam espectrômetros NIR portáteis neste contexto ainda é pequena até o momento. Sendo encontrados apenas alguns trabalhos que descrevem a viabilidade dos NIR portáteis para este tipo de aplicação. No trabalho desenvolvido por Amat-Tosello *et al.* [23], quatro espectrômetros portáteis SW-NIR semelhantes foram utilizados para prever os parâmetros RON (*Research Octane Number*) e MON (*Motor Octane Number*) em gasolinas utilizando regressão PLS. Lutz *et al.* [24] avaliaram os efeitos de quatro modos de medida em uma célula de transflectância personalizada utilizando um espectrômetro ultra portátil MicroNIRTM2200 para quantificar o teor de etanol em gasolinas por regressão PLS. O trabalho de Paiva *et al.* [25] avaliou o desempenho de dois espectrômetros, um FT-NIR de bancada e um MicroNIRTM1700 ultracompacto para determinar o teor de biodiesel e detectar adulteração por óleo vegetal em misturas diesel/biodiesel, comparando duas técnicas de regressão multivariada, PLS e MLR.

2.1.3 Transferência de calibração utilizando espectrômetros de infravermelho próximo de bancada e portátil

De modo geral, o objetivo das técnicas de transferência de calibração é minimizar as diferenças entre as respostas instrumentais, ou valores preditos, de equipamentos diferentes

[26,27]. Existem diversos métodos de transferência de calibração descritos na literatura envolvendo a espectroscopia NIR, principalmente aplicados a fármacos, alimentos e combustíveis. A Padronização Direta (*Direct Standardization* - DS) e a Padronização Direta por Partes (*Piecewise Direct Standardization* - PDS) são as técnicas de transferência de calibração mais usadas.

No trabalho desenvolvido por Boiret *et al.* [28], o método de PDS é utilizado para transferir modelos PLS entre dois instrumentos FT-NIR do mesmo modelo, para estimar a atividade do insumo farmacêutico ativo em comprimidos revestidos. O trabalho de Sulub *et al.* [29], que abordou a determinação de uniformidade em comprimidos, demonstrou a transferência de calibração de modelos PLS de um instrumento NIR dispersivo para quatro instrumentos FT-NIR, utilizando o método PDS e *Wavelet Hybrid Direct Standardization* (WHDS). Pereira *et al.* [30] propôs o uso de padronização direta por partes em janela dupla (*Double Window Piecewise Direct Standardization* - DWPDS) entre espectros de reflectância difusa adquiridos em um espectrômetro FT-NIR para as mesmas amostras em formas físicas diferentes. No trabalho foi desenvolvido um método para determinação do fármaco nevirapina em formulações farmacêuticas sólidas, utilizando regressão PLS. Peng *et al.* [31] realizaram um estudo de comparação entre o método de regressão espectral, proposto em seu trabalho, e o método PDS para transferir modelos de regressão PLS e PCR para amostras de milho (entre 3 espectrômetros NIR diferentes) e de fármacos (entre dois espectrômetros NIR diferentes). Liu *et al.* [32] avaliaram o desempenho de 5 diferentes técnicas de transferência de calibração (Correção de *slope/bias* - SBC, *Local Centering*, *Orthogonal Signal Correction* - OSC, *Direct Standardization* - DS e PDS) entre dois instrumentos FT-NIR e um *scanning grating*, na determinação de proteína bruta em silagens. Fernández-Ahumada *et al.* [33] descreveram um método envolvendo a aplicação da padronização PDS na indústria de alimentação animal para transferir um modelo PLS de um espectrômetro NIR situado em um laboratório de controle de qualidade (*at-line*) para um espectrômetro Vis-NIR conectado à linha de produção (online). Igne *et al.* [34] realizaram um estudo comparando técnicas de transferência de calibração baseadas na remoção de sinal ortogonal (*Dynamic Orthogonal Projection* - DOP, *Transfer by Orthogonal Projection* - TOP, *Error Removal by Orthogonal Subtraction* - EROS, *Orthogonal Signal Correction* - OSC and *Orthogonal Projections to Latent Structures* - O-PLS) para transferir modelos de calibração PLS do teor de proteína e óleo em soja utilizando quatro analisadores NIR. Fan *et al.* [35] propuseram utilizar o método de análise de correspondência canônica (*Canonical*

Correlation Analysis – CCA) e realizaram uma comparação com o método PDS para transferir modelos de calibração PLS do teor de nitrogênio total em tabaco entre dois espectrômetros FT-NIR de marcas diferentes.

Os trabalhos envolvendo aplicação em combustíveis também são publicados com bastante frequência. Em um trabalho abordando a determinação de parâmetros de qualidade da gasolina (naftênicos e RON), Pereira *et al.* [26] realizou uma comparação do desempenho de sete diferentes métodos (DS, PDS, OSC, *Reverse Standardization* - RS, *Piecewise Reverse Standardization* - PRS, SBC e *Model Update* - MU) na transferência de calibração entre três espectrômetros FT-NIR, de duas marcas diferentes. Amat-Tosello *et al.* [23] utilizaram outro método de padronização, *External Parameter Orthogonalisation* - EPO, para transferir calibração entre quatro analisadores portáteis de combustíveis na região do NIR em um trabalho abordando a determinação dos indicadores de desempenho da gasolina RON e MON por regressão PLS.

O trabalho de Galvão *et al.* [36] propõe um método de transferência de calibração envolvendo variáveis individuais e compara-o com a técnica PDS, para uso em equipamentos de infravermelho próximo dedicados para determinação de massa específica, RON e naftênicos em gasolina. Lira *et al.* [37] empregaram a técnica DS em espectros de infravermelho próximo e médio para prever parâmetros de qualidade em misturas diesel/biodiesel por regressão PLS, com resultados comparáveis com a calibração completa do equipamento. A técnica de correção multiplicativa de sinal foi utilizada por Kramer *et al.* [38] para padronizar espectros NIR de dois espectrômetros, sem utilizar amostras de transferência, na determinação de densidade em combustíveis de aviação.

Em contrapartida, poucos trabalhos relataram métodos de transferência de calibração de equipamentos de bancada para equipamentos portáteis. As metodologias descritas nesses trabalhos são aplicadas nos campos agroalimentar, monitoramento on-line industrial, bioenergia e fármacos, em que o método de padronização direta por partes é o mais utilizado. Este método é utilizado no trabalho de Krapf *et al.* [39] para transferir a calibração entre um instrumento de laboratório e um analisador *on-line* de processo no monitoramento *in situ* da digestão anaeróbica de culturas energéticas e resíduos pecuários, utilizando a espectroscopia NIR e regressão PLS.

Os métodos SBC, Transferência por Projeções Ortogonais (TOP) e o PDS, foram utilizados por Salguero-Chaparro *et al.* [40] em um trabalho que aborda a utilização de dispositivos portáteis na região do NIR para o monitoramento on-line industrial. Foram

utilizados um instrumento NIR (pré-dispersivo) off-line e um NIR portátil (pós-dispersivo) com arranjo de diodos, para avaliar a qualidade de azeitonas. O método PDS também apresentou melhor resultado comparado com OSC e TOP de acordo com o trabalho de Barreiro *et al.* [41], que avalia o desempenho dessas técnicas na transferência de calibração, de um espectrômetro NIR de laboratório para NIR portáteis utilizados em campo, no controle de qualidade de alimentação em um programa de reprodução de animais.

Pierna *et al.* [42] descreveram um estudo semelhante ao trabalho de Barreiro *et al.*, porém, propondo um método em que os espectros de um equipamento são corrigidos adicionando as diferenças entre os dois instrumentos para cada comprimento de onda individualmente. Neste trabalho, foi realizada uma transferência de calibração entre um instrumento NIR dispersivo de bancada para um NIR portátil para análise em campo de parâmetros de qualidade em alimentos agrícolas utilizando regressão PLS.

Em Zamora-Rojas *et al.* [43], três técnicas de transferência de calibração, DS, PDS e *Spectral difference by wavelengths* (SDW), foram avaliadas para transferir dados espectrais de um espectrômetro NIR de alto desempenho *at-line* para um portátil no controle de qualidade in-situ de carnes. Hoffman *et al.* [44] também utilizaram a técnica PDS para transferir espectros adquiridos em instrumentos FT-NIR de laboratório para NIR portáteis e avaliou a padronização com base no desempenho de modelos PLS para determinar a fração dos componentes em diferentes misturas entre benzeno, ciclohexano e etilbenzeno e PLS-DA para discriminar diferentes tipos de polímeros. Silva, V., Silva, J. e Pereira. [45] utilizaram três instrumentos NIR portáteis para quantificar três diferentes formas cristalinas de mebendazol em matéria prima farmacêutica, comparando o desempenho entre esses três instrumentos e também com um instrumento de bancada FT-NIR. Nesse trabalho, foi demonstrada uma metodologia para transferência de calibração entre o instrumento de bancada e um dos instrumentos portáteis.

2.1.4 Introdução do conceito de padrões virtuais

Cooper e colaboradores [19-21] demonstraram recentemente em alguns de seus trabalhos a aplicação, na análise de combustíveis, de espectros matematicamente construídos a partir dos espectros de componentes puros presentes na composição dos combustíveis, chamados de padrões virtuais. Estes são capazes de simular espectros de amostras reais. Nestes trabalhos, os padrões virtuais foram utilizados com o método de correção *slope/bias*

(SBC^P) para corrigir valores preditos por modelos PLS entre instrumentos NIR. Uma vez que os padrões virtuais são obtidos a partir de um determinado instrumento, seus equivalentes também podem ser reconstruídos em um segundo equipamento com base apenas em uma matriz de composição e nos espectros dos solventes adquiridos no novo equipamento. Esse procedimento é descrito em detalhes no tópico 3.1.5.1. Porém, todos os instrumentos envolvidos nos trabalhos de Cooper e colaboradores são idênticos, analisadores portáteis de combustível (FUELex NIR, *Bruker Optics, The Woodlands, TX, USA*) com detector de arranjo de diodo de InGaAs de 512 elementos e resolução de 3 nm. A técnica de padronização dos valores preditos utilizada geralmente funciona com sucesso em casos em que as medidas no novo equipamento apresentam pequenas diferenças [16].

No caso do FT-NIR de bancada e do NIR ultracompacto, empregados neste trabalho de tese, que apresentam configurações instrumentais e características espectrais bastante diferentes, é necessário realizar uma padronização das respostas espectrais e a padronização direta (DS) é a técnica mais utilizada.

2.2 Espectroscopia de Infravermelho Próximo (NIR)

A absorção de energia proveniente da radiação infravermelha provoca mudanças nos estados vibracionais e rotacionais das moléculas. Porém, isso só acontece quando os movimentos de vibração e rotação naturais da molécula causam variações no seu momento dipolar resultando na geração de um campo elétrico oscilante [46]. E, além disso, quando a energia proveniente dos fótons da radiação coincide com a diferença de energia entre os níveis vibracionais no estado molecular excitado e no estado fundamental. Essa energia (E_f) pode ser descrita em função da frequência (ν), como:

$$E_f = h\nu \quad (2.1)$$

$$\nu = \frac{c}{\lambda} \quad (2.2)$$

onde, h é a constante de Planck, c é a velocidade da luz e λ é o comprimento de onda em uma determinada frequência ν [47].

A região do infravermelho próximo está situada no espectro eletromagnético no intervalo entre 12.800 cm^{-1} e 4.000 cm^{-1} [48]. Fatores como rapidez, robustez, não

destrutividade e a análise sem pré-tratamento das amostras, entre outros fizeram da espectroscopia no infravermelho próximo, ao longo das últimas décadas, uma ferramenta muito atrativa para um extenso campo de aplicações incluindo o monitoramento de processos no setor industrial [49,50a]. Porém, isso não teria acontecido sem o desenvolvimento de ferramentas estatísticas avançadas no campo da quimiometria e o avanço tecnológico sofrido pelo setor eletrônico, para superar as dificuldades encontradas com a grande complexidade dos espectros e as necessidades instrumentais da região do NIR.

Na região do infravermelho próximo, a absorção de energia acontece principalmente devido à ocorrência de dois fenômenos característicos: os sobretons e as bandas de combinação. Os sobretons são absorções em que a molécula sofre transições do estado fundamental para o segundo ou terceiro estado energético excitado (1° e 2° sobretons). As bandas de sobretons possuem intensidade até 1000 vezes menor que a intensidade das bandas de absorção correspondentes às transições entre o estado fundamental e o primeiro estado excitado (transições fundamentais). As bandas de combinação, por sua vez, surgem quando ocorre interação entre as frequências de vibrações de duas, ou mais, ligações. Quando há essa combinação, surgem bandas de absorção em frequências aproximadamente iguais à soma ou diferença das vibrações das ligações. Portanto, nesse caso a energia de um determinado comprimento de onda é absorvida por duas ou mais ligações. Assim como os sobretons, as bandas de combinação também apresentam intensidades menores, na ordem de 10 vezes em relação às bandas fundamentais [48].

As transições fundamentais referentes a ligações O-H, C-H, N-H e S-H, por exemplo, correspondem a números de onda entre 3.330 e 2.500 cm^{-1} . Portanto, as bandas de sobreton e combinações relacionadas a estas ligações comumente estão situadas em regiões de maiores frequências, ou seja, maiores números de onda. A região do infravermelho próximo é, de fato, rica em bandas de absorção provenientes de sobretons e combinações entre frequências de vibração de ligações entre o átomo de H e átomos mais pesados, como o O, C, N e S [47].

Em geral, os equipamentos que operam na faixa espectral do infravermelho próximo apresentam um custo menor, comparados aos equipamentos MIR [47]. A instrumentação também é mais simples e bastante semelhante aos equipamentos que trabalham na região do UV-Vis. Isso contribuiu para o desenvolvimento de instrumentos portáteis com ótima robustez que são uma ferramenta valiosa para aplicações em análises de campo e monitoramento de processos. Com a utilização de conexões de fibra ótica também é possível

realizar medidas em locais distantes do equipamento, de grande interesse para análises de monitoramentos *in-line* [50b,47].

A sobreposição de bandas, tipicamente observada nos espectros NIR, os torna muito mais complexos em relação aos espectros MIR, diminuindo a seletividade estrutural e dificultando a interpretação [51]. Entretanto, essa complexidade atribui, aos espectros NIR, informações específicas relacionadas com diferentes tipos de ligação C-H, que são extremamente úteis. Portanto, para poder extrair dos espectros a maior quantidade possível da informação útil, a espectroscopia NIR para fins analíticos está normalmente associada a técnicas quimiométricas multivariadas [50].

2.3 Calibração Multivariada.

De modo geral, a calibração multivariada é o processo em que se relaciona múltiplas respostas de um instrumento, como os espectros na região do infravermelho próximo, por exemplo, com uma ou mais propriedades de um determinado conjunto de amostras [52].

Os métodos de calibração multivariada podem ser classificados como clássicos ou inversos. Nos métodos inversos, como a regressão PLS, a concentração (ou propriedade de interesse) é tratada como uma função das respostas instrumentais, como na equação a seguir [50]:

$$\mathbf{c} = \mathbf{Rb} \quad (2.3)$$

em que, o vetor \mathbf{c} ($n_{\text{amostras}} \times 1$) contém as concentrações de cada amostra, \mathbf{R} ($n_{\text{amostras}} \times q_{\text{variáveis}}$) é a matriz de respostas instrumentais e o vetor \mathbf{b} ($q_{\text{variáveis}} \times 1$) contém os coeficientes do modelo.

Entre as técnicas de calibração multivariada normalmente utilizados para extrair informações de espectros na região NIR, as mais utilizadas são a Regressão Linear Múltipla (MLR), Regressão em Componentes Principais (PCR) e Regressão por Mínimos Quadrados Parciais (PLS). A técnica de regressão PLS foi aplicada neste trabalho e será descrita em mais detalhes no tópico 2.3.5.

2.3.1 Pré-processamento de dados.

As técnicas de pré-processamento consistem de uma série de manipulações que são realizadas nos dados previamente à utilização de técnicas de análise multivariada, com a finalidade de eliminar informações irrelevantes que não estão relacionadas com o problema estudado.

2.3.1.1 PRÉ-PROCESSAMENTO NAS VARIÁVEIS.

Dentre os pré-processamentos nas variáveis, a centralização na média é o mais comum, e é realizado subtraindo-se os valores de cada vetor variável, representado pelos elementos em uma mesma variável ao longo de todas as amostras, pelo valor médio desse vetor. Esse procedimento é realizado em todas as variáveis, portanto, é o equivalente a subtrair a média espectral, obtida a partir de todas as amostras, de todo o conjunto de dados [52].

2.3.1.2 PRÉ-PROCESSAMENTO NAS AMOSTRAS.

Os métodos de correção de linha de base são normalmente utilizados em dados espectroscópicos, nos quais, tipicamente algumas variáveis espectrais contêm sinal devido apenas ao *background*. Assumindo que o espectro de uma amostra qualquer pode ser representado em função da soma entre o sinal de interesse e a linha de base, então, uma equação pode ser utilizada para descrever a contribuição de ambos os componentes nesse espectro [52]:

$$\mathbf{r} = \tilde{\mathbf{r}} + \alpha + \beta x + \gamma x^2 + \delta x^3 \dots \quad (2.4)$$

onde, $\tilde{\mathbf{r}}$ é o sinal de interesse e os outros termos da equação representam uma aproximação da linha de base.

Uma vez que se consegue descrever a linha de base de um determinado espectro, então, um espectro corrigido pode ser obtido subtraindo-a diretamente do espectro original (\mathbf{r}). Se a variação na linha de base resulta apenas de um efeito aditivo, por exemplo, então ela pode ser descrita como na Equação 2.5. Quando a linha de base também resulta de um efeito multiplicativo, então um segundo termo deve ser adicionado ao modelo (Equação 2.6).

$$\mathbf{r} = \tilde{\mathbf{r}} + \boldsymbol{\alpha} \quad (2.5)$$

$$\mathbf{r} = \tilde{\mathbf{r}} + \boldsymbol{\alpha} + \boldsymbol{\beta}x \quad (2.6)$$

Para correção de efeitos de espalhamento da radiação, a variação normal padrão (SNV) [53] pode ser aplicada a cada espectro isoladamente e não utiliza nenhuma referência associada ao conjunto de dados, assim, não é influenciada pelas características do conjunto de dados como um todo. Ao sofrer essa transformação cada espectro passa a ter média igual a zero e variância igual a um.

No método SNV, inicialmente é realizada uma centralização na média, ou seja, em cada espectro todos os seus valores (r_i) são subtraídos pelo valor médio (\bar{r}_i). Depois de centralizados, os valores espectrais são divididos pelo desvio padrão (σ_i) dos valores de seu respectivo espectro (S_i), de acordo com a Equação 2.7:

$$S_i = \frac{(r_i - \bar{r}_i)}{\sigma_i} \quad (2.7)$$

A técnica *Multiplicative Signal Correction* (MSC) [54] também é normalmente utilizada para corrigir variações de linha de base decorrentes de efeitos de espalhamento da radiação. No entanto, em casos onde o espalhamento não é a principal fonte de variabilidade, a informação química pode ser influenciada, comprometendo o ajuste da regressão [52,55].

Na técnica MSC, inicialmente, é ajustada uma regressão de cada espectro contra um espectro de referência, estimando ambos os coeficientes de regressão, linear (b_0) e angular ($b_{ref,1}$), como mostra a Equação 2.8:

$$r_{orig} = b_0 + b_{ref,1} \cdot r_{ref} + e \quad (2.8)$$

Os coeficientes b_0 e $b_{ref,1}$ são, então, aplicados ao espectro original e um espectro corrigido é obtido:

$$r_{corr} = \frac{r_{orig} - b_0}{b_{ref,1}} = r_{ref} + \frac{e}{b_{ref,1}} \quad (2.9)$$

onde, r_{orig} é o espectro original, r_{ref} é o espectro utilizado como referência para todos os espectros originais, e é o resíduo da modelagem de r_{orig} , r_{corr} é o espectro corrigido, e, b_0 e $b_{ref,1}$ são os coeficientes linear e angular, respectivamente, calculados para cada r_{orig} .

As funções derivativas [56], por sua vez, geralmente são utilizadas para remover variações de linha de base e aumentar a resolução de dados espectrais, melhorando a distinção entre bandas sobrepostas, como mostra o exemplo da Figura 1.

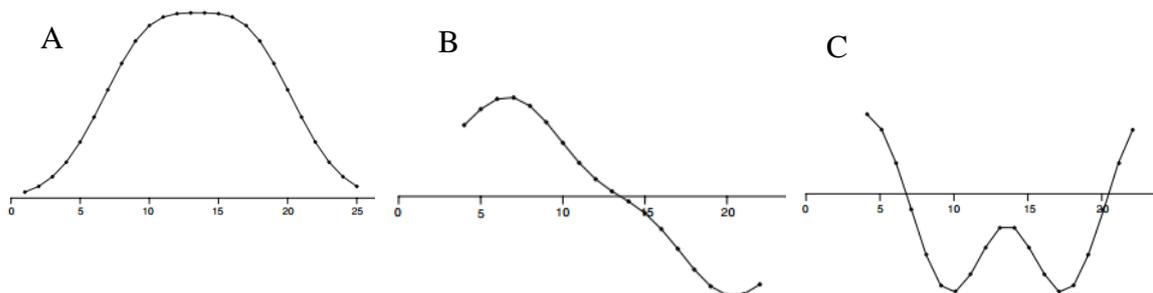


Figura 1. (A) Bandas sobrepostas, (B) primeira derivada e (C) segunda derivada. (Fonte: BRERETON, 2003 [56]).

A aplicação da derivada, porém, pode provocar um aumento significativo dos ruídos espectrais. Para contornar esse problema geralmente são usados alguns filtros de suavização, como por exemplo o filtro *Savitzky-Golay* [55,56]. Esse filtro utiliza uma janela de pontos centrada em um determinado comprimento de onda e um polinômio de pequena ordem ajustado aos pontos dessa janela por mínimos quadrados. O tamanho da janela é um fator importante, pois apesar de eliminar os ruídos, janelas com pontos demais podem suavizar (atenuar) as bandas a ponto de provocar perda de informações. Por outro lado, janelas muito pequenas podem amplificar os ruídos presentes no espectro original.

2.3.2 Seleção de Variáveis.

2.3.2.1 MÉTODO *JACK-KNIFE*.

O método *Jack-Knife* [57] é utilizado com o objetivo de eliminar variáveis que diminuem a confiabilidade ou que não são importantes para os modelos PLS. A redução do número de variáveis originais geralmente resulta na utilização de um número menor de variáveis latentes, o que torna o modelo mais simples. Nesse método, inicialmente, a

variância dos coeficientes de regressão (\mathbf{b}) do modelo PLS para cada variável, é calculada da seguinte forma:

$$s^2 \mathbf{b} = \left[\sum_{j=1}^J (\mathbf{b} - \mathbf{b}_j)^2 \right] \left[\frac{N-1}{N} \right] \quad (2.10)$$

em que, N é o número de amostras; $s^2 \mathbf{b}$ é a incerteza estimada do coeficiente de regressão (\mathbf{b}), individualmente; \mathbf{b} é o coeficiente de regressão com para o número ótimo de fatores com todas as amostras de calibração (N); e \mathbf{b}_j é o coeficiente de regressão para o número ótimo de fatores na validação cruzada. Em seguida, as variáveis são selecionadas por meio de testes de significância. Um teste t avalia a significância dos valores estimados para cada coeficiente de regressão (\mathbf{b}), obtendo-se assim o nível de significância correspondente a cada variável. Às variáveis que não forem significativas é atribuído o coeficiente de regressão igual a zero.

2.3.2.2 INTERVAL PLS (iPLS).

O algoritmo iPLS [58] constrói iterativamente modelos de regressão PLS locais em intervalos espectrais resultantes da divisão da faixa espectral total. Com isso, a principal propriedade do algoritmo é fornecer uma visão geral sobre quais das diferentes regiões espectrais ao longo de todo o espectro possuem informações relevantes. Dessa forma é possível identificar as regiões que mais contribuem para o ajuste da regressão PLS. A comparação entre a capacidade preditiva dos modelos locais em relação ao modelo global, construído com todo o espectro, é baseada no valor da raiz quadrada do erro médio quadrático de validação cruzada (**RMSECV**).

2.3.3 Seleção de Amostras (SPXY).

O algoritmo SPXY [59] é extensivamente utilizado no desenvolvimento de modelos de calibração multivariada para dividir grupos de amostras entre conjuntos de calibração e validação externa.

O SPXY calcula a distância Euclidiana na matriz \mathbf{R} e na variável dependente (\mathbf{y}), iterativamente, para cada par de amostras do conjunto de dados. A partir daí, um par de

amostras é selecionado toda vez que a distância calculada, $d_{ry}(\mathbf{p}, \mathbf{q})$, é maior que todas as que foram calculadas nas interações anteriores, garantindo que as amostras do conjunto de calibração descrevem de forma representativa todo o conjunto de dados.

A distância $d_{ry}(\mathbf{p}, \mathbf{q})$ é a soma das distâncias Euclidianas em \mathbf{R} e \mathbf{y} divididas pelo máximo valor, como mostra a Equação 2.11.

$$d_{ry}(\mathbf{p}, \mathbf{q}) = \frac{d_r(\mathbf{p}, \mathbf{q})}{\max_{p, q \in [1, N]} d_r(\mathbf{p}, \mathbf{q})} + \frac{d_y(\mathbf{p}, \mathbf{q})}{\max_{p, q \in [1, N]} d_y(\mathbf{p}, \mathbf{q})}; \quad \mathbf{p}, \mathbf{q} \in [1, N] \quad (2.11)$$

2.3.4 Regressão por Mínimos Quadrados Clássicos (CLS).

O método dos mínimos quadrados clássicos (*Classical Least Squares*) [52] pode ser utilizado quando o sistema estudado obedece uma relação linear entre os vetores de medidas e a concentração, como por exemplo a relação estabelecida pela Lei de Beer, em que a absorvância espectral em um determinado comprimento de onda é proporcional à concentração do componente químico. Além disso, o método CLS considera que a resposta de um instrumento para uma mistura de componentes é igual à soma das respostas deste instrumento para os componentes puros da mistura. Assim, a notação matricial para os métodos clássicos de calibração multivariada (CLS) tem a forma da Equação 2.12:

$$\mathbf{R} = \mathbf{C}\mathbf{S} \quad (2.12)$$

onde, \mathbf{R} é a matriz de respostas instrumentais para misturas de componentes, \mathbf{S} é a matriz com as respostas instrumentais para cada componente puro e \mathbf{C} é a matriz de concentrações de cada componente puro para cada mistura. Portanto, os espectros de infravermelho das misturas em \mathbf{R} resultam de uma combinação linear dos espectros dos componentes puros em \mathbf{S} , cujos coeficientes são as concentrações na matriz \mathbf{C} . Na Equação 2.12 a matriz \mathbf{C} pode ser calculada de acordo com as Equações 2.13 a 2.15:

$$\mathbf{R}\mathbf{S}^T = \mathbf{C}\mathbf{S}\mathbf{S}^T \quad (2.13)$$

$$\mathbf{R}\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1} = \mathbf{C}(\mathbf{S}\mathbf{S}^T)(\mathbf{S}\mathbf{S}^T)^{-1} \quad (2.14)$$

$$\mathbf{R}\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1} = \hat{\mathbf{C}} \quad (2.15)$$

onde $\hat{\mathbf{C}}$ é a matriz contendo as estimativas das concentrações.

Então, a Equação 2.15 pode ser utilizada para prever as concentrações dos componentes em novas amostras (espectros) desconhecidas que contém os mesmos componentes. Também é possível reconstruir a estimativa dos espectros das misturas $\hat{\mathbf{R}}$ multiplicando a matriz de concentrações $\hat{\mathbf{C}}$, calculada na Equação 2.15, pela matriz dos espectros dos componentes puros \mathbf{S} :

$$\hat{\mathbf{R}} = \hat{\mathbf{C}}\mathbf{S} \quad (2.16)$$

2.3.5 Regressão por Mínimos Quadrados Parciais (PLS)

Para estabelecer o modelo de regressão a técnica PLS calcula fatores estabelecidos com base em ambas as matrizes, \mathbf{R} e \mathbf{y} . Esses fatores são chamados de variáveis latentes (VL's) e estabelecem as direções de máxima covariância entre \mathbf{R} (medidas) e \mathbf{y} (concentração ou propriedade de interesse). Portanto, as variáveis latentes estão diretamente relacionadas com a variabilidade no vetor \mathbf{y} [52,55].

Os métodos inversos de calibração multivariada têm a forma descrita na Equação 2.3. No modelo PLS, uma nova matriz (\mathbf{U}) é calculada e substitui a matriz \mathbf{R} :

$$\mathbf{y} = \mathbf{U}\mathbf{b} \quad (2.17)$$

As colunas que compõem a matriz \mathbf{U} são combinações lineares das colunas da matriz \mathbf{R} . A matriz \mathbf{U} é calculada com base em ambas as matrizes, \mathbf{R} e \mathbf{y} , da seguinte forma:

$$\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (2.18)$$

onde, \mathbf{V} é a matriz de loadings e \mathbf{S} é a matriz diagonal de valores singulares. Cada elemento da diagonal representa a quantidade de covariância descrita por cada variável latente. Resolvendo a Equação 2.18 para \mathbf{U} , utilizando a propriedade ortonormal de \mathbf{V} ($\mathbf{V}^T\mathbf{V} = \mathbf{I}$), tem-se que:

$$\mathbf{U} = \mathbf{R}\mathbf{V}\mathbf{S}^{-1} \quad (2.19)$$

A Equação 2.17 pode, então, ser resolvida para o vetor \mathbf{b} (utilizando-se também a propriedade ortonormal de \mathbf{U}), estabelecendo-se a regressão, como mostra a Equação 2.20:

$$\hat{\mathbf{b}} = \mathbf{U}^T \mathbf{y} \quad (2.20)$$

Um novo vetor de escores (\mathbf{u}_{desc}) pode ser calculado a partir de um vetor de medidas de amostras desconhecidas (\mathbf{r}_{desc}) e das matrizes \mathbf{V} e \mathbf{S} , estabelecidas pela regressão:

$$\mathbf{u}_{desc} = \mathbf{r}_{desc} \mathbf{V} \mathbf{S}^{-1} \quad (2.21)$$

Portanto, o valor da concentração (\hat{y}) para uma amostra desconhecida (vetor de medidas desconhecido - \mathbf{r}_{desc}) pode ser estimado a partir de \mathbf{u}_{desc} e do vetor $\hat{\mathbf{b}}$:

$$\hat{y} = \mathbf{u}_{desc} \hat{\mathbf{b}} \quad (2.22)$$

2.3.6 Avaliação dos modelos de calibração multivariada

Os parâmetros analíticos indicadores de desempenho, **RMSE** (*Root Mean Square Error*) e **bias** [55], são utilizados para avaliar a capacidade preditiva dos modelos PLS, levando em conta os valores preditos pelo modelo e os valores de referência para os parâmetros modelados. As comparações do desempenho de diferentes modelos são normalmente baseadas nos testes estatísticos de Fischer e t de *student*.

Considerando uma única amostra, o **RMSE**(\hat{y}) é matematicamente representado pela raiz quadrada do quadrado da diferença entre o valor predito (\hat{y}) e o valor de referência (como na Equação 2.23). O **RMSE** possui a mesma unidade dos valores originais medidos (\mathbf{y}).

$$\mathbf{RMSE}(\hat{y}) = \sqrt{(\hat{y} - y)^2} \quad (2.23)$$

A raiz quadrada do erro médio quadrático calculado para a etapa de calibração de um modelo multivariado (**RMSEC**), em que as mesmas amostras do conjunto de calibração são utilizadas para avaliar a sua capacidade preditiva, é definida como:

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{N_{cal}} (\hat{y}_i - y_i)^2}{(N_{cal} - A - 1)}} \quad (2.24)$$

onde, \hat{y} é o valor predito para o seu respectivo valor de referência (y), N_{cal} é o número de amostras de calibração, e A o número de variáveis latentes.

Na etapa de validação interna, utilizando a técnica de validação cruzada *Venetian Blinds* para calcular o número ideal de variáveis latentes, o conjunto de dados com \mathbf{n} amostras é dividido em s subconjuntos e o número de subconjuntos deve ser menor que $\mathbf{n}/2$. Cada conjunto de teste é determinado começando da amostra 1 e selecionando as próximas a cada s amostras no conjunto de dados de calibração. Inicialmente, o primeiro conjunto de amostras é retirado do conjunto e um modelo é construído com as restantes. O modelo é aplicado às amostras que foram deixadas de fora e os valores de referência (\mathbf{y}) e preditos ($\hat{\mathbf{y}}$) são comparados. O primeiro subconjunto retorna ao conjunto de calibração e o novo conjunto é selecionado iniciando da amostra 2. Dessa forma, o procedimento se repete até que s subconjuntos tenham sido deixados de fora do conjunto de calibração. Com isso, o $RMSE(\hat{\mathbf{y}})$ para esta etapa é a raiz quadrada do erro médio quadrático de *Cross-validation* [55]. O $RMSECV$ é obtido pela Equação 2.25:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{N_{cal}} (\hat{y}_{CV,i} - y_i)^2}{N_{cal}}} \quad (2.25)$$

onde, $\hat{y}_{CV,i}$ é o valor previsto na etapa de validação cruzada e y_i é o valor de referência para a amostra i .

Da mesma forma, na etapa de validação externa é calculada a raiz quadrada do erro médio quadrático de previsão ($RMSEP$), dada por:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{N_{pred}} (\hat{y}_i - y_i)^2}{N_{pred}}} \quad (2.26)$$

sendo, N_{pred} o número de amostras, \hat{y}_i o valor predito e y_i o valor de referência, para cada amostra do conjunto de validação externa.

O teste estatístico de Fischer (Teste F) [60,61] pode ser utilizado para testar hipóteses sobre a relação entre variâncias populacionais, utilizando a Equação 2.27:

$$\frac{s_A^2}{s_B^2} \cong \frac{\sigma_A^2}{\sigma_B^2} F_{calc} \quad (2.27)$$

onde, s_A^2 e s_B^2 são as variâncias amostrais e σ_A^2 e σ_B^2 as variâncias populacionais.

Em um caso particular, pode ser testada a possibilidade de as duas variâncias populacionais serem idênticas, ou seja, no caso em que a razão $\frac{\sigma_A^2}{\sigma_B^2}$ é igual a 1. Nesse caso, é necessário testar a hipótese nula em que as variâncias amostrais s_A^2 e s_B^2 são estimativas da mesma variância populacional. Assim o teste F pode ser realizado obtendo-se o valor de F calculado de acordo com a Equação 2.28:

$$F_{calc} = \frac{s_A^2}{s_B^2} \quad (2.28)$$

Em seguida, é necessário comparar o valor de F_{calc} com o valor crítico tabelado (F_{V_A, V_B}), considerando os respectivos números de graus de liberdade, V_A e V_B . Se F_{calc} for maior que F_{V_A, V_B} , então a hipótese nula é rejeitada. O teste F é, em geral, utilizado para avaliar a semelhança entre os valores de RMSEP (Erro Médio Quadrático na etapa de validação externa) em modelos PLS.

Enquanto o **RMSEP** avalia a exatidão das previsões na etapa de validação externa, um valor de **bias** distante de zero indica a presença de erros sistemáticos [55]. O **bias** é calculado pela Equação 2.29:

$$bias = \frac{\sum_{i=1}^{N_p} (\hat{y}_i - y_i)}{N_{pred}} \quad (2.29)$$

A avaliação da significância estatística do **bias** dos modelos PLS pode ser realizada utilizando o teste t de *student*, de acordo com a norma ASTM E 1655-05 [62]. Nesse caso o cálculo do **t** envolve o desvio padrão dos erros de validação (SDV - *standard deviation of validation*). O SDV e o valor de **t** são calculados pelas equações 2.30 e 2.31:

$$SDV = \sqrt{RMSEP^2 - Bias^2} \quad (2.30)$$

$$t_{calc} = \frac{|Bias|\sqrt{N_{pred}}}{SDV} \quad (2.31)$$

onde, N_{pred} é o número de amostras de previsão do modelo.

Em seguida o valor de t_{calc} é comparado com o t crítico tabelado para distribuição t de *student*. O *bias* do modelo PLS é estatisticamente significativo se o valor de t_{calc} for maior que o respectivo valor crítico, considerando v graus de liberdade no nível de 95% de confiança.

Os testes de significância também podem ser utilizados para comparar valores de *bias* entre dois diferentes modelos PLS. O teste estatístico para comparação de duas médias experimentais, de acordo com Miller e Miller [63], está descrito a seguir.

Inicialmente, é verificado se as variâncias dos respectivos modelos são significativamente diferentes, de acordo com o teste F descrito anteriormente nesta seção. Caso os desvios padrões não sejam significativamente diferentes, uma estimativa conjunta do desvio padrão, s_c , é calculada utilizando a equação 2.32:

$$s_c = \sqrt{\frac{(N_{pred\ 1}-1)SDV_1^2 + (N_{pred\ 2}-1)SDV_2^2}{N_{pred\ 1} + N_{pred\ 2} - 2}} \quad (2.32)$$

Então, para avaliar se os dois *bias* são significativamente diferentes o valor de t é calculado como segue:

$$t_{calc} = \frac{Bias_1 - Bias_2}{s_c \sqrt{\frac{1}{N_{pred\ 1}} + \frac{1}{N_{pred\ 2}}}} \quad (2.33)$$

e comparado com o valor de t crítico com $N_{pred\ 1} + N_{pred\ 2} - 2$ graus de liberdade. Porém, caso as variâncias sejam significativamente diferentes, ou seja, os dois conjuntos de amostras possuam desvios padrões diferentes, a utilização de uma estimativa conjunta do desvio padrão não é adequada. Nesse caso, o teste é realizado utilizando as equações 2.34 e 2.35, para o cálculo do t e do número de graus de liberdade, v :

$$t_{calc} = \frac{Bias_1 - Bias_2}{\sqrt{\frac{SDV_1^2}{N_{pred\ 1}} + \frac{SDV_2^2}{N_{pred\ 2}}}} \quad (2.34)$$

$$v = \frac{\left(\frac{SDV_1^2}{N_{pred\ 1}} + \frac{SDV_2^2}{N_{pred\ 2}} \right)^2}{\frac{\left(\frac{SDV_1^2}{N_{pred\ 1}} \right)^2}{N_{pred\ 1}+1} + \frac{\left(\frac{SDV_2^2}{N_{pred\ 2}} \right)^2}{N_{pred\ 2}+1}} \quad (2.35)$$

O valor de graus de liberdade calculado, v , é arredondado para um número inteiro e o t crítico tabelado é encontrado considerando-se v graus de liberdade.

2.4 Transferência de Calibração

As condições de operação e o ambiente aos quais um determinado equipamento está submetido podem sofrer variações ao longo do tempo. Além disso, as amostras analisadas também podem apresentar mudanças em sua composição em relação às amostras que foram utilizadas para calibrar o equipamento. Esses fatores, portanto, afetam diretamente as respostas instrumentais. Por isso, um modelo de calibração desenvolvido em um equipamento deve ser aplicado a respostas instrumentais do mesmo equipamento e em condições de operação equivalentes àsquelas em que foi construído, do contrário, as previsões terão sua confiabilidade comprometida. Entretanto, por meio de técnicas de transferência de calibração é possível contornar essa limitação, corrigindo as respostas instrumentais de um equipamento para que possam ser submetidas a um modelo de calibração que foi desenvolvido em condições distintas ou, até mesmo, em um outro equipamento [64].

A Padronização Direta (*Direct Standardization* - DS) [26,47,64-66] é uma das técnicas de transferência de calibração mais utilizadas. Esse método corrige os espectros de um dado equipamento secundário, \mathbf{R}_s , de forma que esses espectros se tornem semelhantes aos espectros do equipamento primário, \mathbf{R}_p . Esta transformação é realizada por meio da uma matriz $\mathbf{F}(\mathbf{R}_p = \mathbf{F}\mathbf{R}_s)$, chamada de matriz de correção. Quando o número de variáveis da matriz de dados é menor que o número de amostras, então a matriz \mathbf{F} pode ser obtida por meio do cálculo da matriz inversa de \mathbf{R}_s . Caso contrário, um espaço coluna reduzido de \mathbf{R}_m é normalmente calculado por decomposição de valores singulares (*Singular Value Decomposition* - SVD) [55]. A decomposição de valores singulares é uma técnica de fatorização de matrizes que consiste em definir as direções de variabilidade mais importantes na matriz de dados, assim como suas respectivas variâncias. Então, os vetores singulares (semelhantes aos escores) resultantes da projeção da matriz \mathbf{R}_s neste mesmo espaço coluna

são usados na padronização ao invés dos valores originais. Uma vez que a matriz \mathbf{F} foi calculada, basta multiplicá-la pelo espectro de uma nova amostra adquirido no instrumento secundário para corrigi-lo.

O método de Padronização Reversa (*Reverse Standardization* - RS) envolve os mesmos procedimentos matemáticos usados no DS, porém, corrige os espectros do equipamento primário, \mathbf{R}_p , de forma que eles se tornem semelhantes aos espectros do equipamento secundário, \mathbf{R}_s . Então, um novo modelo é desenvolvido com os espectros corrigidos e aplicado diretamente aos espectros secundários originais. Este procedimento é bastante flexível uma vez que permite a detecção de *outliers* e seleção de variáveis nos novos modelos construídos após a padronização [26,66].

2.5 Programa de monitoramento da qualidade de combustíveis (PMQC)

De acordo com o histórico fornecido pela ANP [6], a redução no número de não conformidades entre as amostras de gasolina coletadas pelo PMQC caiu de 7.098 em 2001 (três anos após o início do programa) para 1.094 em 2010. E isso ocorreu ao mesmo tempo em que aumentou o número de amostras coletadas pelo programa. Em 2010, as principais não conformidades encontradas nas amostras de gasolina estavam relacionadas ao teor de etanol e a destilação. Nas amostras de óleo diesel o percentual de não conformidade caiu de 6,5% em 2001 para 3,6% em 2010. Neste último ano (2010) o teor de biodiesel ainda foi responsável por 1,1% das amostras. Em 2016, o índice de conformidade no mês de dezembro atingiu 98,6% nas amostras de gasolina, com o teor de etanol e a destilação sendo responsáveis respectivamente por 46% e 35% das não conformidades. Este índice para o mesmo mês foi de 95,7% nas amostras de óleo diesel, com o teor de biodiesel sendo de longe o maior responsável, com 40% das não conformidades [67].

Os parâmetros físico-químicos da qualidade para a gasolina comum (ANEXO 1) e óleo diesel (ANEXO 2), monitorados pelo PMQC, são atualmente especificados pelas resoluções ANP N° 30 de 29/06/2015 e ANP N° 30 de 23/06/2016, respectivamente.

No presente trabalho, foram modelados oito parâmetros de qualidade da gasolina, as temperaturas de destilação a 10%, 50%, 90% de evaporado e no ponto final de ebulição (TPF), a massa específica a 20°C, Índice Antidetonante (IAD), MON (*Motor Octane Number*) e RON (*Research Octane Number*). Para as misturas diesel/biodiesel, apenas o teor de biodiesel foi modelado.

O método de destilação atmosférica da gasolina, de acordo com as normas NBR9619 e D86, realiza a determinação quantitativa das faixas de ebulição características da gasolina. As temperaturas determinadas para as diferentes percentagens de combustível evaporado ao longo da faixa de ebulição são um indicador da tendência de evaporação da gasolina. O gráfico das diferentes temperaturas contra a quantidade de volume evaporado caracteriza a curva de destilação. As temperaturas recomendadas pela ANP para compor a curva de destilação da gasolina são a 10%, 50%, 90% de evaporado e no ponto final de ebulição (TPF). A informação sobre a tendência de volatilidade da gasolina, fornecida nas análises de destilação atmosférica está, de maneira geral, relacionada com o desempenho do combustível. Por exemplo, o balanceamento adequado entre os parâmetros pressão de vapor e as temperaturas a 10% de evaporado, além de influenciar no desempenho, é importante no controle de emissão de gases. A temperatura a 50% de evaporado é um indicador da capacidade de aquecimento do combustível “*warm-up*” e do desempenho de aceleração, principalmente em condições de baixa temperatura. Baixas temperaturas a 50% favorecem a diminuição na emissão de hidrocarbonetos. As temperaturas em 90% de evaporado e no ponto final de ebulição estão relacionadas com a percentagem de compostos pesados na composição da gasolina. Altas temperaturas em 90% normalmente contribuem para o aumento na economia de combustível e octanagem, entretanto, temperaturas muito altas em 90% e no ponto final podem causar efeitos como o aumento na emissão de hidrocarbonetos [68].

Na determinação da densidade da gasolina, de acordo com a norma ASTM D4052, recomendada pela ANP, a análise de um pequeno volume de amostra é realizada em um analisador de densidade digital. A densidade é uma relação entre a massa e o volume de uma substância em uma temperatura específica. Os combustíveis dentro da especificação possuem densidades bastante típicas, de forma que o aumento ou diminuição dessa propriedade pode ser um indício de adulteração. O uso das informações complementares da densidade e de outras propriedades da gasolina também é importante na caracterização das percentagens de componentes leves e pesados em sua composição [68].

O índice antidetonante (IAD) é um parâmetro bastante aceito em muitos países para medir a capacidade antidetonante de um combustível. Em geral, a utilização de um combustível com índice antidetonante maior do que o especificado para o motor não provoca aumento no desempenho. Porém, combustíveis com IAD menores que o especificado para o motor causam perdas de potência. Os motores possuem uma capacidade de detonação que

varia especificamente de acordo com as características físico-químicas estabelecidas pelas condições operacionais e projeto. Essa relação é bastante complexa, por isso, o desempenho antidetonante do combustível, determinado pelo IAD, depende da medida de outros dois parâmetros, o MON (*Motor Octane Number*) e o RON (*Research Octane Number*). Esses dois parâmetros determinam a resistência do combustível à detonação em um motor com alta rotação e com baixa rotação, respectivamente [68]. O IAD é determinado pela média aritmética entre os valores de MON e RON.

O teor de biodiesel em óleo diesel é determinado por espectroscopia de infravermelho médio com base na norma EN14078, recomendada pela ANP. Segundo a norma, a mistura diesel/biodiesel deve ser diluída em hexano e os espectros adquiridos no modo de transmitância. A absorbância do grupo carbonila é então utilizada para determinar o teor de biodiesel por meio de uma curva analítica univariada construída utilizando diluições de um padrão de biodiesel regulamentado em hexano. O grupo carbonila presente no éster (biodiesel) absorve intensamente a radiação na região de 1745 cm^{-1} , região na qual o óleo diesel não apresenta absorção.

3 MATERIAIS E MÉTODOS

3.1 Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais

3.1.1 Instrumentação

Os espectros de transmitância foram adquiridos em dois espectrômetros de infravermelho próximo, um FT-NIR de bancada e um NIR ultracompacto, utilizando uma cubeta de quartzo com caminho óptico de 20 mm. O instrumento primário é o espectrômetro de bancada FT-NIR de alta resolução, modelo FrontierTM da marca Perkin Elmer, equipado com um detector de sulfato de triglicina (NIR-TGS). Este equipamento foi operado na faixa de comprimento de onda de 680 nm até 2500 nm (12000 - 4000 cm^{-1}), com resolução espectral de 4 cm^{-1} . O instrumento secundário é um espectrômetro ultracompacto MicroNIR1700 da marca JDSU *Corporation* equipado com um filtro variável linear (LVF) e um detector InGaAs com arranjo de 128 pixels. A faixa operacional do MicroNIR vai de 950 nm até 1650 nm (11012,0 - 5965,8 cm^{-1}) e possui uma resolução nominal espectral de 6,25 nm. Um acessório de transmitância construído em laboratório descrito por Maia *et al.* (2015) [25] foi utilizado para adquirir os espectros de absorbância no MicroNIR1700.

3.1.2 Amostras e reagentes

O conjunto de dados foi composto por 103 amostras de gasolina automotiva tipo C e 130 misturas diesel/biodiesel, coletadas nos estados de Pernambuco, Sergipe e Alagoas. As 103 amostras de gasolina e 49 misturas diesel/biodiesel (óleo diesel B7) foram fornecidas pelo Programa de Monitoramento da Qualidade de Combustíveis, regulamentado pela Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Além disso, diesel e biodiesel puros foram fornecidos pela distribuidora Dislub Combustíveis LTDA. As 81 amostras de diesel/biodiesel restantes foram obtidas adicionando diferentes volumes de biodiesel ao diesel puro e às amostras de diesel/biodiesel B7, a fim de produzir misturas com teor de biodiesel menor e maior que 7% v/v, respectivamente. Portanto, a faixa do teor de biodiesel observada nas 130 misturas diesel/biodiesel foi de 1% a 9,9% v/v.

Adicionalmente, 12 componentes químicos puros tipicamente presentes na composição desses combustíveis foram adquiridos comercialmente de diferentes fornecedores: etanol (Dinâmica, pureza = 99,50%), pentano (Dinâmica, pureza = 99,00%), 1-Penteno (Aldrich Chemistry, pureza \geq 98,50%), hexano (Dinâmica, pureza = 99,00%), 1-hexeno (Aldrich Chemistry, pureza \geq 99,00%), heptano (Dinâmica, pureza = 99,50%), tolueno (Sigma Aldrich, pureza \geq 99,50%), iso-octano (Vetec Química Fina, pureza = 99,50%), xileno (Dinâmica, pureza = 98,50%), etilbenzeno (Fluka Analytical, pureza \geq 99,00%), n-hexadecano (Sigma Aldrich, pureza = 99,00%) e padrão de biodiesel B100 (AccuStandard).

3.1.3 Aquisição dos espectros NIR

No equipamento NIR de bancada, os espectros de absorvância foram adquiridos na região de 12000 até 4000 cm^{-1} com média de 16 varreduras por espectro e medindo absorvância a cada 4 cm^{-1} . Os espectros do NIR ultracompacto foram adquiridos na região de 11012,0 cm^{-1} até 5965,8 cm^{-1} com média de 50 varreduras por espectro e tempo de integração de 50 ms e medindo absorvância a cada 6 nm aproximadamente. O espectro de *background* foi adquirido no equipamento de bancada com a cubeta de quartzo vazia e limpa no compartimento de amostra. No ultracompacto, os espectros foram adquiridos com o auxílio do acessório de transmitância feito em laboratório [25]. A medida em 0% de transmitância foi feita impedindo a passagem de luz entre a fonte (a mini lâmpada de tungstênio do acessório de transmitância, alimentada por uma porta USB) e o detector, enquanto que a medida em 100% de transmitância foi adquirida da mesma maneira como foi medido o *background* no equipamento de bancada.

3.1.4 Modelos de regressão PLS

A regressão PLS foi utilizada para construir modelos quantitativos de 5 parâmetros de qualidade da gasolina (temperaturas de destilação em 10%, 50%, 90% de volume recuperado e no ponto final de ebulição e densidade a 20 °C) e do teor de biodiesel em misturas diesel/biodiesel. As faixas de calibração para esses parâmetros foram, respectivamente, 49 a 57,5°C, 64 a 74°C, 130 a 165°C, 180 a 215°C, 710 a 750 $\text{kg}\cdot\text{m}^{-3}$ e 1 a 9,9% v/v.

Os valores de referência para as temperaturas de destilação em 10%, 50%, 90% de volume recuperado e no ponto final de ebulição (PFE) foram determinados de acordo com a norma ASTM D86. A determinação dos valores de referência para densidade e teor de biodiesel foi realizada de acordo com as normas ASTM D4052 e EN14078, respectivamente.

As amostras e as respectivas propriedades foram separadas em dois grupos, calibração e validação externa, numa proporção de aproximadamente 70% e 30%, para cada uma das propriedades da gasolina e das misturas diesel/biodiesel. Os conjuntos de calibração e validação externa foram selecionados pelo algoritmo SPXY (*sample set partitioning based on joint x-y distances*) utilizando os espectros pré-processados [59]. Para selecionar o número ótimo de variáveis latentes, a validação interna dos modelos PLS foi realizada utilizando o método de validação cruzada *venetian blinds*. Algumas técnicas de pré-processamento de dados espectrais, como a variação normal padrão (SNV), correção multiplicativa de sinal (MSC), primeira derivada e correção de linha de base, foram testadas para corrigir desvios de linha de base presente nos espectros. O desempenho dos modelos foi avaliado e comparado com base nos valores obtidos para os parâmetros de performance analítica RMSEP, *bias*, R^2 e nos testes de significância de Fischer e *t-student*.

3.1.5 Método de transferência de calibração

Como já foi mencionado, o espectrômetro de bancada foi considerado como instrumento primário (IP) e o ultracompacto como instrumento secundário (IS). O grupo de amostras do qual os espectros foram adquiridos em ambos os equipamentos, primário e secundário, compõe o chamado conjunto de transferência, empregado no cálculo da matriz de diferença (matriz \mathbf{F}). A partir das amostras de calibração de cada propriedade, o algoritmo SPXY [59] foi utilizado para selecionar um número de amostras para compor o conjunto de transferência. Após realizar a padronização reversa (RS), como descrito no tópico 2.4, os modelos PLS construídos com os espectros do IP que foram corrigidos foram validados com os espectros originais do conjunto de validação externa adquiridos no IS.

Neste trabalho, o desempenho do método de padronização reversa foi comparado entre duas abordagens, utilizando um conjunto de transferência composto por espectros experimentais (espectros adquiridos a partir de amostras reais de combustíveis) e por espectros virtuais.

3.1.5.1 PROCEDIMENTO PARA CRIAÇÃO DOS PADRÕES VIRTUAIS

Devido à enorme diferença entre características espectrais dos instrumentos primário e secundário, todos os espectros do instrumento primário, um FT-NIR de alta resolução, foram submetidos a um procedimento para diminuir sua resolução tornando os números de onda exatamente os mesmos dos espectros do equipamento secundário. Essa etapa foi realizada antes do cálculo dos padrões virtuais. O procedimento consiste em uma interpolação dos valores do espectro primário, correspondentes ao respectivo vetor de números de onda do instrumento primário, nos pontos do vetor de números de onda do instrumento secundário.

O procedimento que foi realizado para obter matematicamente os padrões virtuais está brevemente explicado abaixo, de acordo com o trabalho de Cooper e colaboradores em 2011 [19], e resumido no esquema mostrado na Figura 2. O trabalho de Cooper demonstra uma abordagem segmentada na qual os espectros virtuais resultam da realização do cálculo em cada segmento espectral, ou seja, em cada uma das partes nas quais o espectro foi dividido, separadamente. Os espectros do conjunto de calibração adquiridos no instrumento primário (FT-NIR) foram considerados os espectros-alvo (f). Os espectros-alvo e os espectros dos componentes químicos foram segmentados como mostram as Equações 3.1 e 3.2:

$$f_{alvo} = [f_1 \dots f_j] \quad (3.1)$$

$$x_{componentes\ químicos} = [x_1 \dots x_j] \quad (3.2)$$

onde f_{alvo} é o espectro de uma amostra de gasolina, f_j é um segmento desse espectro, $x_{componentes\ químicos}$ é o espectro de um componente químico e x_j é um segmento desse espectro. Dessa forma, cada segmento do padrão virtual foi construído por vez e depois concatenados formando o espectro inteiro.

A primeira etapa do procedimento consiste em uma regressão linear envolvendo os espectros-alvo (F) e os espectros dos componentes puros (X) (Equação 3.3):

$$F_j = X_j \beta_j + \varepsilon \quad (3.3)$$

onde \mathbf{F}_j é a matriz em que o número de colunas é igual ao número de espectros-alvo e o número de linhas corresponde a quantidade de números de onda do segmento j ; \mathbf{X}_j é uma matriz cujo número de colunas é igual ao número de componentes químicos puros e as linhas correspondem aos números de onda do segmento j e $\boldsymbol{\beta}_j$ é a matriz de coeficientes de regressão. Em outras palavras, a matriz $\boldsymbol{\beta}_j$ contém a quantidade necessária de cada componente puro para compor o segmento j de cada espectro-alvo.

Ao realizar uma decomposição de valores singulares da matriz de componentes puros (\mathbf{X}_j) para obter uma nova matriz \mathbf{C}_j (com o número de linhas correspondendo aos números de onda e o número de colunas igual ao número de valores singulares) e, em seguida, atribuir valor zero para o termo residual $\boldsymbol{\varepsilon}$, a Equação 3.3 pode ser solucionada para gerar os coeficientes de regressão (matriz $\boldsymbol{\beta}_j$), como mostra a Equação 3.4. Dessa forma, o segmento j dos padrões virtuais para o instrumento primário pode ser calculado de acordo com a Equação 3.5.

$$\boldsymbol{\beta}_j = (\mathbf{C}_j^T \mathbf{C}_j)^{-1} \mathbf{C}_j^T \mathbf{F}_j \quad (3.4)$$

$$\mathbf{S}_j = \mathbf{C}_j \boldsymbol{\beta}_j \quad (3.5)$$

Uma vez que a matriz $\boldsymbol{\beta}_j$ foi determinada com base nos espectros do instrumento primário, o segmento j dos padrões virtuais para o instrumento secundário podem ser obtidos da mesma forma multiplicando a matriz $\mathbf{C}_{j,s}$ obtida a partir dos espectros dos componentes puros do instrumento secundário pela mesma matriz $\boldsymbol{\beta}_j$ obtida na Equação 3.6, como mostrado abaixo:

$$\mathbf{S}_{j,s} = \mathbf{C}_{j,s} \boldsymbol{\beta}_j \quad (3.6)$$

Então, os padrões virtuais para o instrumento primário e secundário são obtidos pela concatenação de todos os segmentos, \mathbf{S}_j e $\mathbf{S}_{j,s}$, obtidos nas Equações 3.7 e 3.8, respectivamente:

$$\mathbf{S} = [\mathbf{S}_1 \dots \mathbf{S}_j] \quad (3.7)$$

$$\mathbf{S}_s = [\mathbf{S}_{1,s} \dots \mathbf{S}_{j,s}] \quad (3.8)$$

Em que, \mathbf{S} e \mathbf{S}_s são matrizes cujo número de colunas é igual ao número de padrões virtuais.

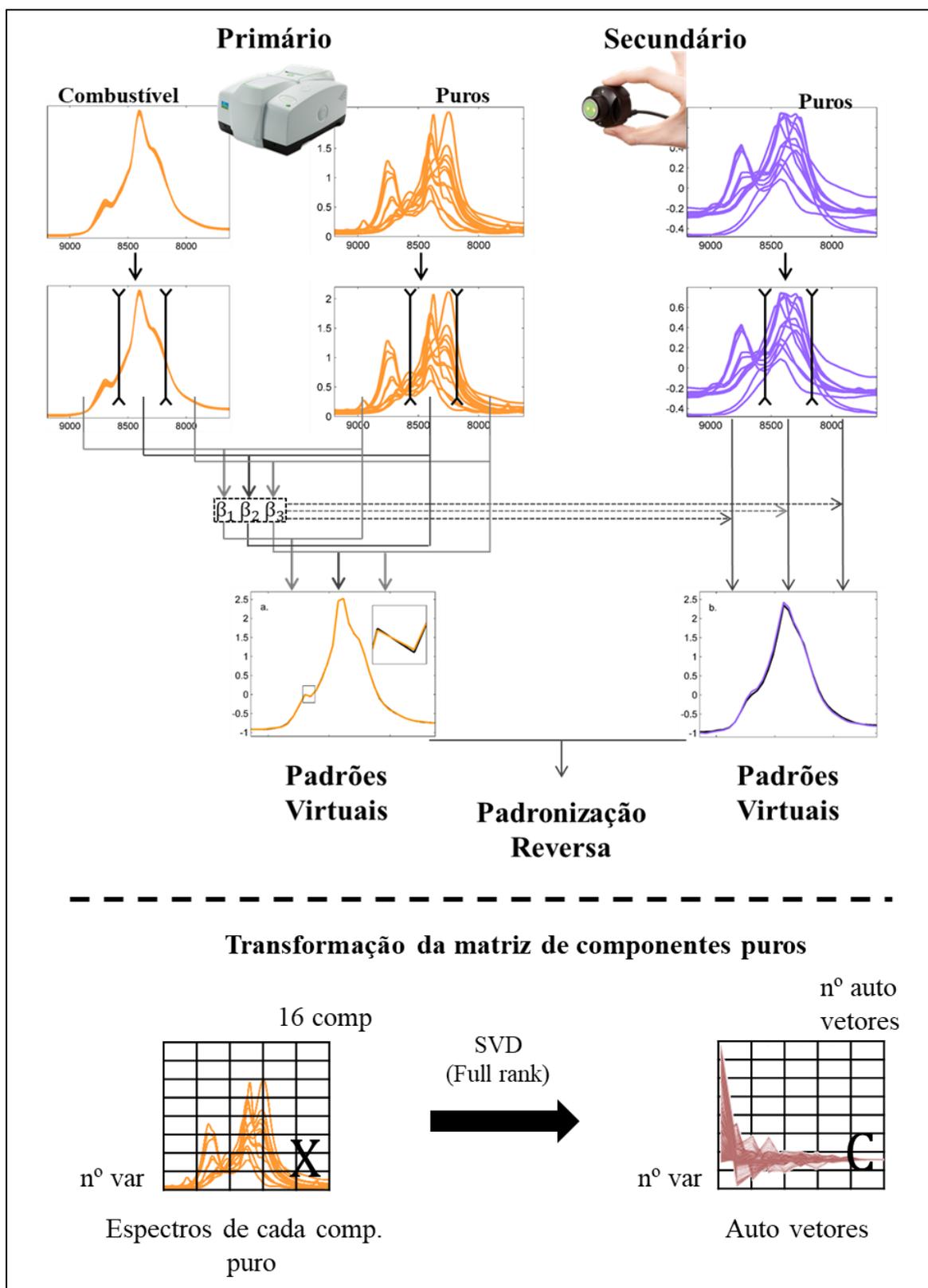


Figura 2. Esquema do procedimento matemático para criação dos padrões virtuais (acima) e transformação da matriz de componentes químicos puros por SVD (abaixo).

3.2 Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR

3.2.1 Instrumentação

Dois espectrômetros de infravermelho próximo foram utilizados para adquirir os espectros no modo de transmitância, sendo um deles o MicroNIR ultracompacto, equipado com o acessório de transmitância mencionado no tópico 3.1.1. O outro instrumento é um espectrômetro de bancada FT-NIR, modelo ABB FTLA2000 da marca Bomem, equipado com um detector de sulfato de triglicina deuterada (MIR-NIR DTGS). Este equipamento possui faixa operacional de números de onda entre 12.000 e 500 cm^{-1} (~833 e 20.000 nm). Uma célula de fluxo (ACC3-9001 Bomem) com caminho ótico de 0,5 mm e janelas de BaF₂ foi utilizada nas medidas de transmitância.

Os parâmetros de desempenho das gasolinas utilizados, *Motor octane number* (MON) e *Research octane number* (RON) foram determinados conforme as normas ASTM D2699 e D2700 em um motor C.F.R, padronizado pelo *Cooperative Fuel Research Committee*. O índice antidetonante (IAD) é determinado pela média aritmética entre os valores de MON e RON.

3.2.2 Amostras

Todas as amostras foram analisadas no laboratório de controle de qualidade de uma companhia petroquímica localizada no nordeste do Brasil. A companhia também forneceu um histórico de dados com espectros NIR, formulações e octanagem das gasolinas produzidas. Devido à disponibilidade dos espectrômetros, após início do projeto algumas das amostras foram analisadas apenas em um dos instrumentos.

As amostras de gasolina e de corrente de hidrocarbonetos foram coletadas diretamente da planta industrial durante a produção da gasolina. Todas as coletas foram realizadas seguindo o padrão de amostragem estabelecido pela empresa, visando evitar perdas e mudanças de composição por conta da alta volatilidade das amostras.

3.2.2.1 GASOLINAS

O banco de dados históricos foi composto por 804 espectros NIR de gasolina (todos adquiridos no NIR de bancada entre os anos de 2001 e 2014). Após o início do projeto, foram analisadas 259 amostras de gasolina automotiva tipo A (206 do Tipo 1 e 53 do Tipo 2) no NIR de bancada e 170 amostras de gasolina automotiva tipo A (111 do Tipo 1 e 59 do Tipo 2) no NIR ultracompacto. Além disso, no instrumento ultracompacto, também foram adquiridos 18 espectros de gasolinas que foram formuladas em laboratório utilizando correntes de hidrocarbonetos coletadas na linha de produção. Para realizar essas formulações foram utilizadas proporções das correntes com base no histórico de produção.

3.2.2.2 CORRENTES DE HIDROCARBONETOS

Entre os anos de 2015 e 2017, foram coletadas amostras das diferentes correntes de hidrocarbonetos normalmente utilizadas na formulação da gasolina. No NIR de bancada foram analisadas no total 57 amostras de correntes de hidrocarbonetos (16 - corrente de hidrocarbonetos pesados; 21 - corrente de hidrocarbonetos intermediários; 14 - corrente de hidrocarbonetos leves e 6 - corrente de hidrocarbonetos aromáticos) e no NIR ultracompacto 42 amostras foram analisadas (12 - corrente de hidrocarbonetos pesados; 14 - corrente de hidrocarbonetos intermediários; 10 - corrente de hidrocarbonetos leves e 6 - corrente de hidrocarbonetos aromáticos). Na discussão dos resultados, essas correntes são respectivamente chamadas de T1, T2, T3 e T4.

3.2.3 Aquisição dos espectros NIR

No espectrômetro de bancada, os espectros de absorvância foram adquiridos no modo de transmitância na região de 10.000 até 4.000 cm^{-1} (1.000-2.500 nm) com média de 16 varreduras por espectro e medindo absorvância a cada 1 cm^{-1} . A medida do espectro de *background* nesse caso foi realizada com a célula de fluxo vazia e limpa no compartimento de amostra.

No ultracompacto os espectros foram adquiridos em toda a faixa operacional, de 11012,0 cm^{-1} até 5965,8 cm^{-1} (950 - 1650 nm), com média de 50 varreduras por espectro e tempo de integração de 50 ms e medindo absorvância a cada 6 nm, aproximadamente. Nesse

instrumento, os espectros foram adquiridos com o auxílio do acessório de transmitância e as medidas em 0% de transmitância e em 100% de transmitância foram adquiridas da maneira como foi descrito no tópico 3.1.3.

3.2.4 Modelos de regressão PLS

Foram construídos modelos PLS para os parâmetros de octanagem MON, RON e IAD. Os valores de referência para MON e RON neste trabalho foram obtidos utilizando um motor C.F.R., padronizado pelo *Cooperative Fuel Research Committee*, conforme estabelecem as normas ASTM D2699 e D2700. O algoritmo SPXY foi utilizado, com os espectros previamente pré-processados, para selecionar os conjuntos de amostras de calibração e validação externa, para cada um dos parâmetros de octanagem, numa proporção de 70% e 30%, respectivamente. O número ótimo de variáveis latentes foi determinado utilizando o método de validação cruzada *venetian blinds* para realizar a validação interna dos modelos PLS. Dentre as técnicas de pré-processamento de dados espectrais, foram testadas a variação normal padrão (SNV), primeira derivada e correção de linha de base. O desempenho dos modelos foi avaliado com base nos valores obtidos para os parâmetros de performance analítica RMSECV, RMSEP, *bias* e R^2 .

3.3 Softwares

Todos os cálculos envolvendo o pré-processamento dos dados, seleção de amostras por SPXY e de variáveis, teste de Fischer (função “FINV”) e *t-student*, aplicação do método de padronização, incluindo o procedimento matemático para criação dos padrões virtuais e o desenvolvimento da interface para o simulador de gasolina automotiva, foram realizados por meio de rotinas em ambiente MATLAB® (*The Math Works, Inc*). O teste de significância *Jack-Knife* foi realizado no software *The Unscrambler*® v 9.7 (*Camo Software AS*). Os modelos PLS foram desenvolvidos na interface PLS_Toolbox 5.2 (*Eigenvector Research, Inc.*) para MATLAB.

4 RESULTADOS E DISCUSSÃO

4.1 Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais

Os espectros brutos adquiridos em toda a faixa operacional dos instrumentos primário e secundário estão apresentados na Figura 3, para as amostras de gasolinas e misturas diesel/biodiesel, e na Figura 4, para os componentes químicos puros. Devido à saturação observada nos espectros do instrumento de bancada, abaixo de 6000 cm^{-1} , a região espectral que vai de $11012,0$ até $5965,8\text{ cm}^{-1}$ (faixa operacional do instrumento ultracompacto) será considerada a faixa espectral completa, para ambos os instrumentos. Portanto, para uma melhor comparação entre os espectros dos dois instrumentos, a faixa espectral completa é mostrada nas Figuras 5 e 6, para os combustíveis e para os componentes puros, respectivamente.

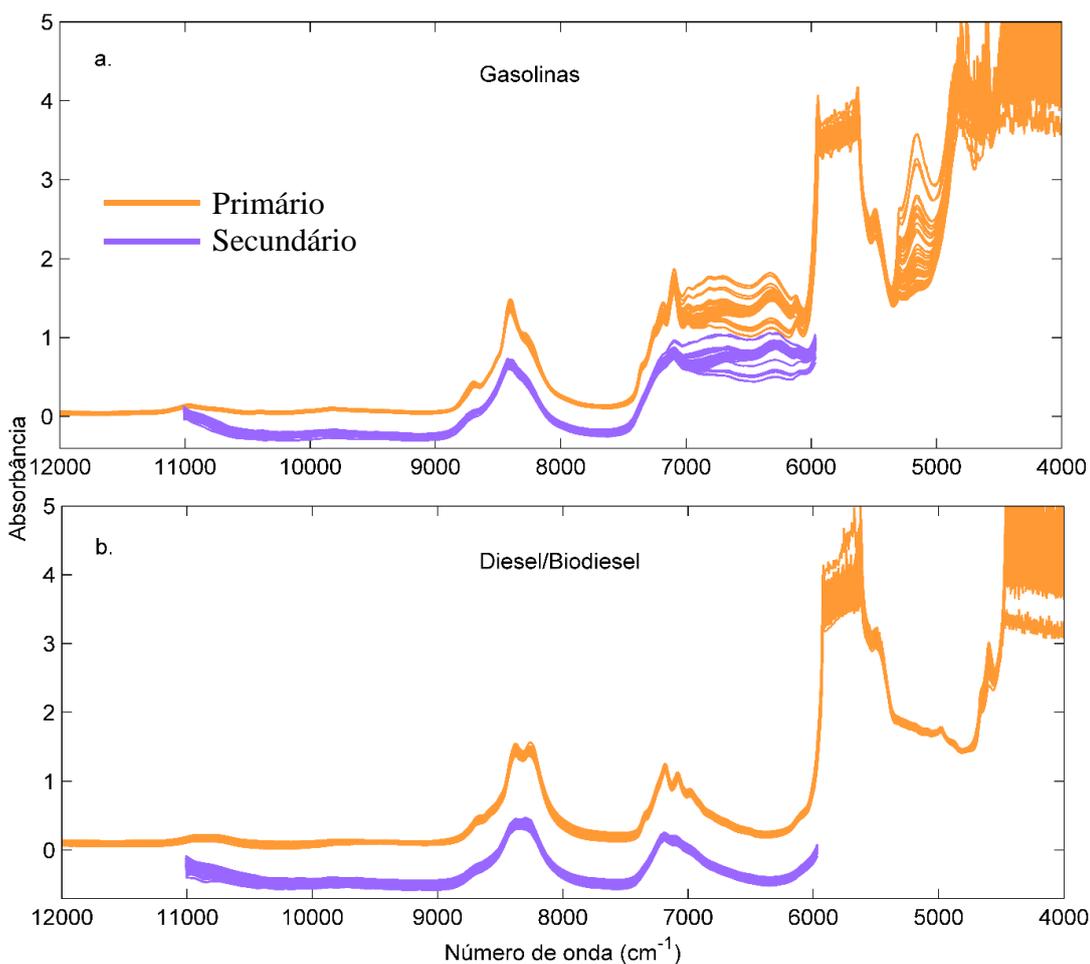


Figura 3. Espectros brutos das amostras de gasolina adquiridos nos instrumentos primário e secundário para as amostras de gasolina (a.) e misturas diesel/biodiesel (b.).

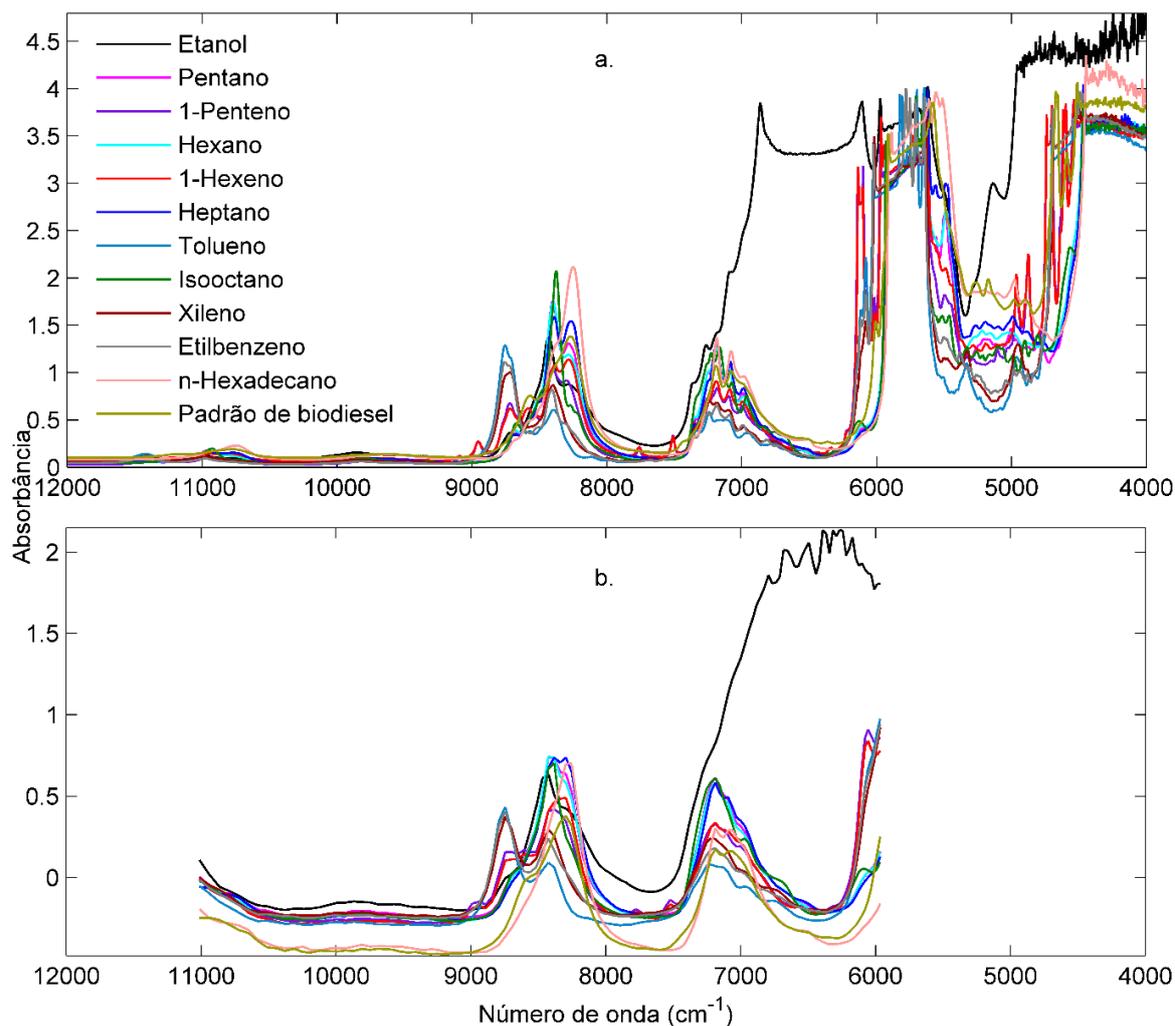


Figura 4. Espectros brutos dos componentes químicos puros presentes na composição da gasolina e misturas diesel/biodiesel, adquiridos nos instrumentos primário (a.) e secundário (b.).

Nos espectros de gasolina (Figura 5a), a banda de formato alargado na região de 7000 a 6100 cm^{-1} , que corresponde ao primeiro sobretom do grupamento O-H do etanol presente na gasolina brasileira, aparece sobreposta com as bandas de primeiro sobretom das combinações de estiramento C-H (em torno de 7150 cm^{-1}). Além disso, as bandas de segundo sobretom de estiramento C-H nos grupamentos CH, CH₂ e CH₃ podem ser observadas entre 9000 e 7800 cm^{-1} . Esses grupamentos estão essencialmente presentes nas cadeias de carbono das substâncias que compõem a gasolina. Nos espectros das misturas diesel/biodiesel, também mostrados na Figura 5, os primeiros sobretoms das bandas de combinação (7500 a 6500 cm^{-1}) e os segundos sobretoms (9000 a 7600 cm^{-1}) das ligações C-H são também observados. A absorção devido ao quarto sobretom de estiramento do grupo C=O ocorre

tipicamente próximo a 8600 cm^{-1} . A absorção do terceiro sobreton do estiramento das ligações C-H também pode ser observada, embora com uma intensidade muito pequena, logo acima de 10500 cm^{-1} , nos espectros do instrumento primário [47,69].

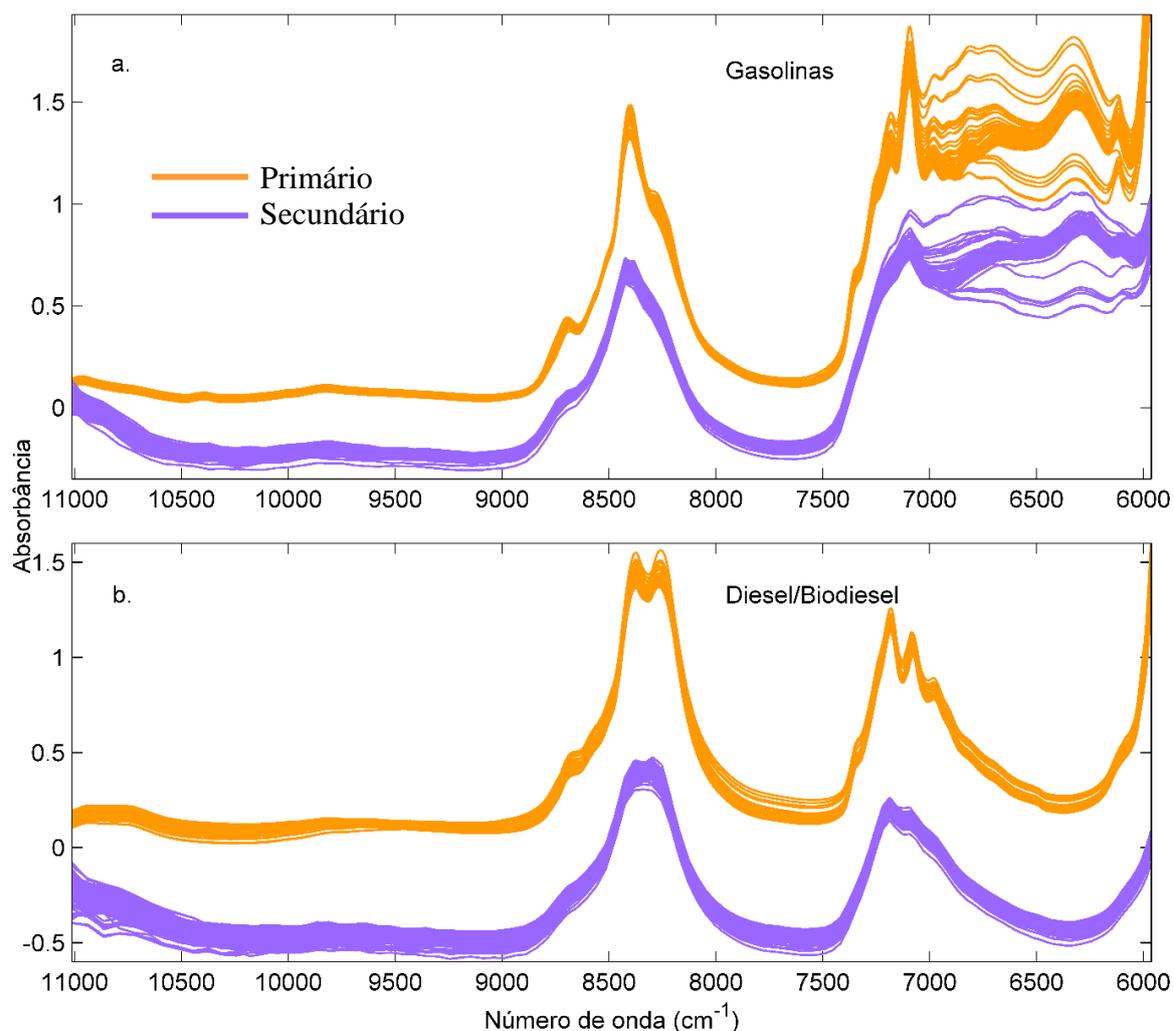


Figura 5. Espectros brutos para todas as amostras de gasolina (a.) e de misturas diesel/biodiesel (b.) adquiridos nos instrumentos primário e secundário na faixa de 11000 a 6000 cm^{-1} .

4.1.1 Avaliação da faixa espectral

No caso dos espectros das misturas diesel/biodiesel, a abordagem para realizar a padronização reversa (RS) utilizando os padrões virtuais atingiu o melhor desempenho quando os modelos foram desenvolvidos na região de segundo sobretons das ligações C-H (9000 a 7600 cm^{-1}). O desempenho do método RS utilizando padrões virtuais depende

fortemente da similaridade entre os espectros experimentais e os virtuais. Considerando a complexidade da composição das amostras e levando em conta os componentes químicos utilizados, essa similaridade foi maior em certas regiões do espectro, como a região de segundo sobretons de estiramento C-H. Além disso, os modelos PLS construídos apenas na região de segundo sobreton atingiram um desempenho muito similar aos modelos construídos na faixa espectral completa (Figura 5), tanto para os parâmetros das misturas diesel/biodiesel como para os da gasolina. As figuras de mérito para essa comparação serão discutidas mais adiante. Ainda, os coeficientes de regressão do modelo PLS para o teor de biodiesel construído com a faixa espectral completa indicam que os valores mais significativos ocorrem na região de segundo sobreton (Figura 6).

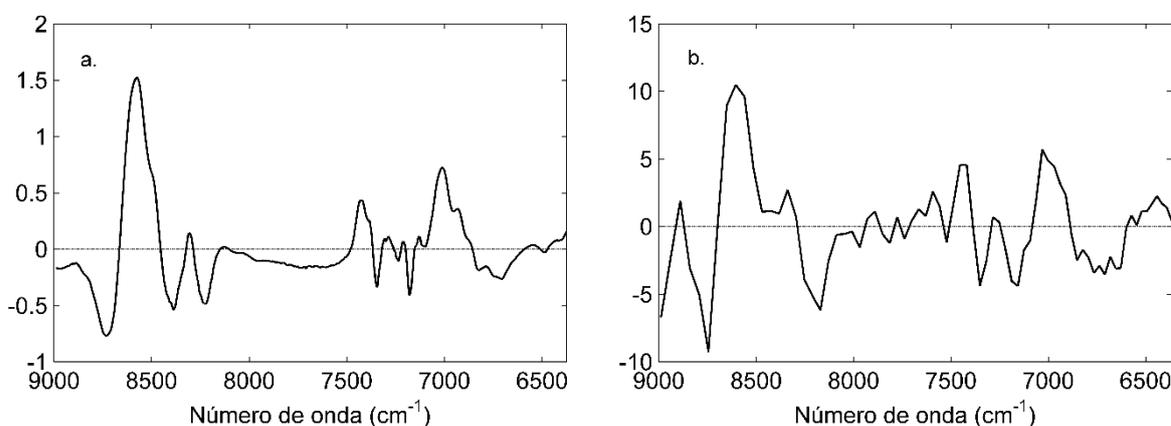


Figura 6. Coeficientes de regressão dos modelos PLS do teor de biodiesel desenvolvidos com os espectros derivativos dos instrumentos primário (a.) e secundário (b.) para a região espectral completa.

Os espectros dos 12 componentes puros analisados com os instrumentos primário e secundário, na região de 11012,0 a 5965,8 cm^{-1} , estão apresentados na Figura 7. Em relação à transferência de calibração para as amostras de gasolina, o espectro do etanol puro apresentou sinal com saturação na região de primeiro sobreton da ligação O-H (em números de onda menores que 7500 cm^{-1}).

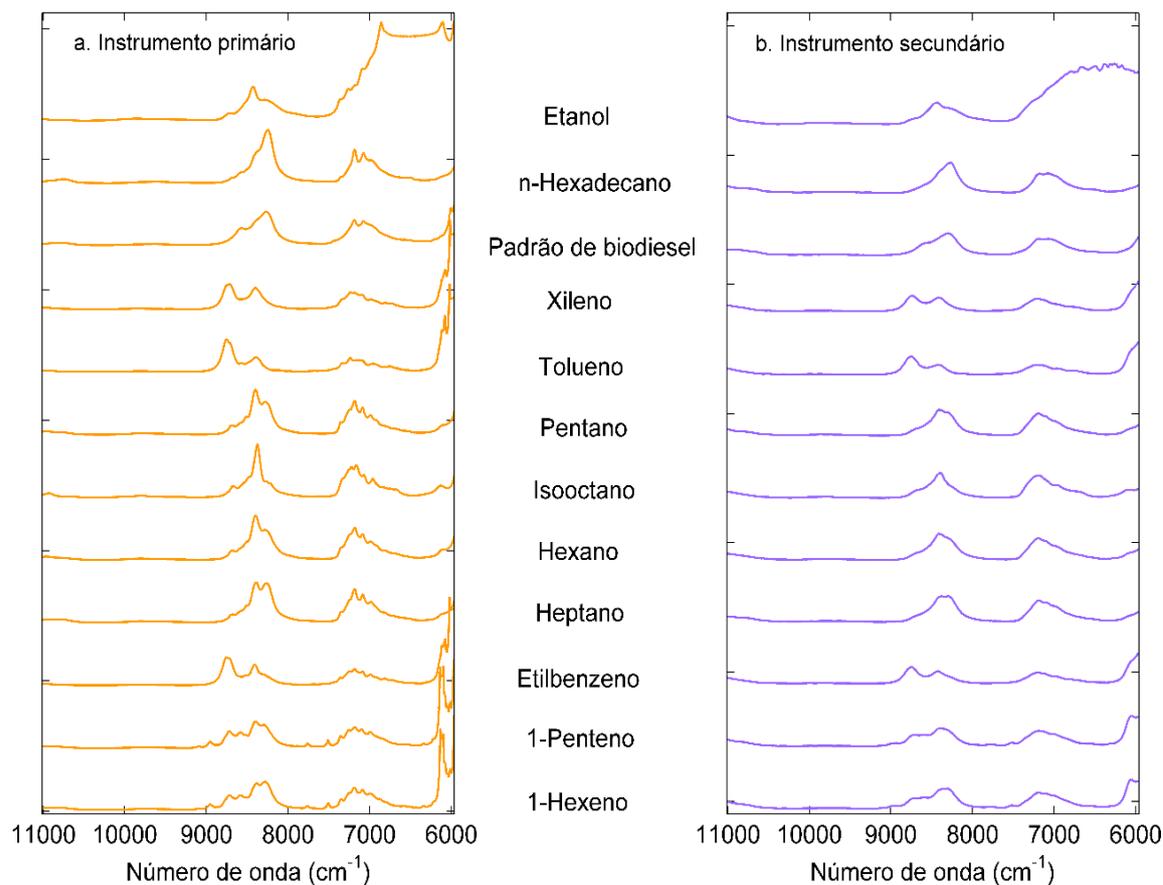


Figura 7. Espectros brutos para os 12 componentes químicos puros adquiridos com os instrumentos primário (a.) e secundário (b.).

Portanto, todos os modelos PLS e os procedimentos de transferência de calibração para ambas as matrizes, gasolina e misturas diesel/biodiesel, foram realizadas nas suas respectivas regiões de segundo sobretom das ligações C-H. As regiões de segundo sobretom nos espectros da gasolina e das misturas diesel/biodiesel estão mostradas na Figura 8.

Para poder comparar apropriadamente os desempenhos das previsões dos parâmetros de qualidade da gasolina e das misturas diesel/biodiesel utilizando toda a região espectral ou apenas na região de segundo sobretom das ligações C-H, foram construídos modelos PLS para cada instrumento. Dentre as técnicas de pré-processamento que foram testadas, como SNV, MSC, correção de linha de base e técnicas derivativas, SNV mostrou o melhor desempenho. A Tabela 1 apresenta as figuras de mérito para esses modelos PLS.

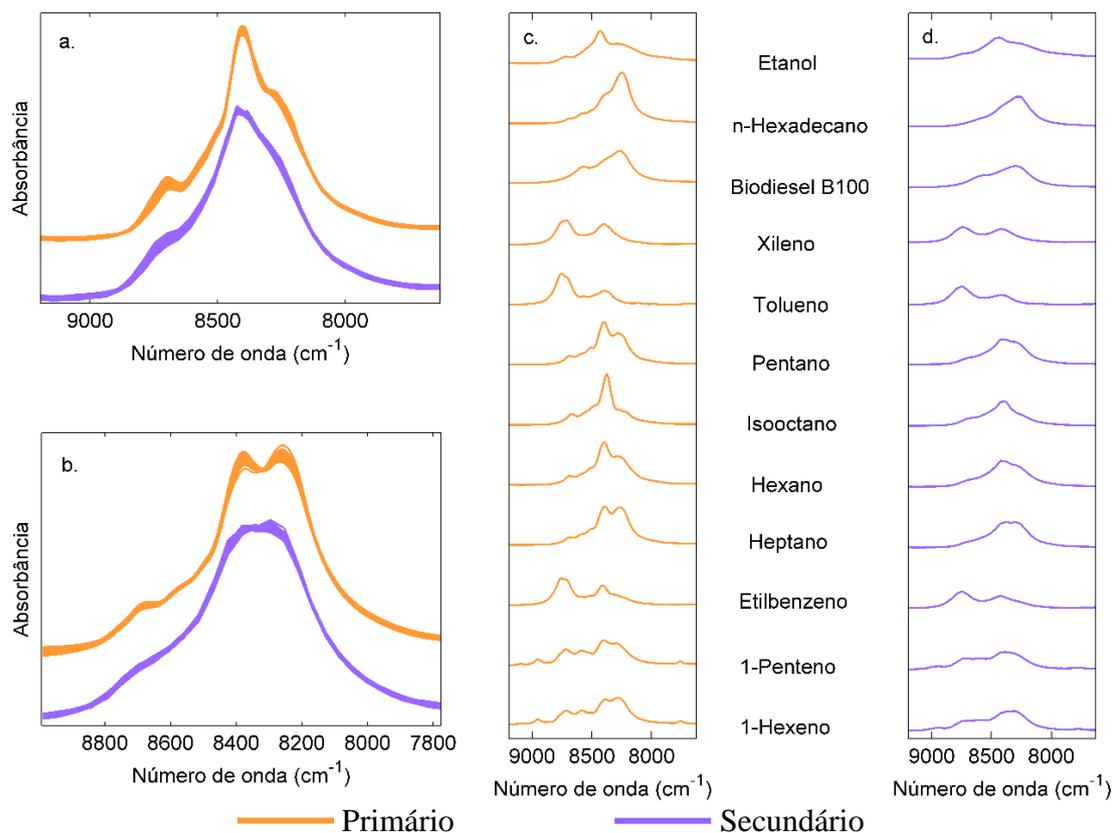


Figura 8. Região de segundo sobretom de ligações C-H para os espectros pré-processados com SNV para todas as amostras de gasolina (a.) e diesel/biodiesel (b.). A região de segundo sobretom nos espectros dos componentes puros é mostrada em (c.) e (d.).

Tabela 1. Figuras de mérito para os modelos PLS construídos com a faixa espectral completa e com a região de segundo sobretom.

	Parâmetro*	Faixa espectral completa				Região de Segundo sobretom de ligações C-H			
		VL	R ²	RMSEP	<i>bias</i>	VL	R ²	RMSEP	<i>bias</i>
Bancada	T10 (°C)	4	0,80	0,4	-0,22**	4	0,80	0,4	-0,20**
	T50 (°C)	4	0,96	0,5	0,062	4	0,96	0,4	0,024
	T90 (°C)	6	0,95	1,9	0,34	6	0,95	2,0	-0,16
	PFE (°C)	6	0,80	2,4	0,35	5	0,79	2,5	0,31
	D (kg/m ³)	5	0,99	0,9	0,25	3	0,99	0,9	0,26
	%Bio (% v/v)	4	0,96	0,4	0,038	3	0,95	0,5	0,08
Ultracompacto	T10 (°C)	4	0,78	0,5	-0,21**	4	0,78	0,6	-0,22**
	T50 (°C)	4	0,93	0,6	0,013	6	0,94	0,7	0,14
	T90 (°C)	4	0,86	3,4	0,86	4	0,83	4,1	0,78
	PFE (°C)	4	0,65	3,7	-0,36	5	0,59	4,4	-0,52
	D (kg/m ³)	4	0,97	1,8	0,29	4	0,92	2,4	1,03**
	%Bio (% v/v)	5	0,88	1,01	0,30	6	0,87	1,06	0,25

* T10, T50, T90 e PFE - temperaturas de destilação a 10%, 50%, 90% de volume recuperado; PFE – ponto final de ebulição; D - densidade; %Bio – teor de biodiesel.

** *bias* estatisticamente significativo (em um nível de significância de 95%).

Os valores de RMSEP obtidos para os modelos construídos na região de segundo sobretom podem ser considerados semelhantes aos valores obtidos para os modelos construídos na região espectral completa, com base no teste F em um nível de 95% de significância. Em relação aos valores de *bias*, de acordo com um teste t (como recomenda a norma ASTM E1655-05 [62]), apenas o parâmetro T10 apresentou *bias* significativo para ambos os instrumentos, bancada e ultracompacto, utilizando a faixa espectral completa ou a região de sobretom. Contudo, quando os valores de *bias* dos modelos construídos na região de sobretom são comparados (teste t) com os valores de *bias* dos modelos construídos na faixa completa, não há diferenças estatisticamente significativas. Um *bias* significativo também foi observado para o modelo de densidade (D) construído no NIR ultracompacto e utilizando apenas a região de segundo sobretom. Também é importante notar que uma vez que os valores de RMSEP foram considerados similares, uma previsão com *bias* pode ser facilmente corrigida utilizando uma correção de *slope* e *bias* simples. Os gráficos dos valores preditos vs. referência para os modelos PLS construídos na região de segundo sobretom, para ambos os instrumentos, são mostrados na Figura 9.

O número de graus de liberdade utilizado em todos os testes F (para verificar a significância estatística dos valores de RMSEP) e testes t (para verificar a significância estatística dos valores de *bias*) realizados neste trabalho correspondem ao número de amostras de previsão para cada parâmetro: 34 (T10), 34 (T50), 33 (T90), 33 (PFE), 33 (D) e 40 (%Bio). O número de graus de liberdade para os testes t realizados para comparar o valor de *bias* entre diferentes modelos foram calculados como recomenda Miller e Miller (teste t para comparação de duas médias) [63]. Dessa forma, o desempenho dos modelos PLS construídos apenas na região de segundo sobretom não foi comprometido em nenhum dos dois instrumentos.

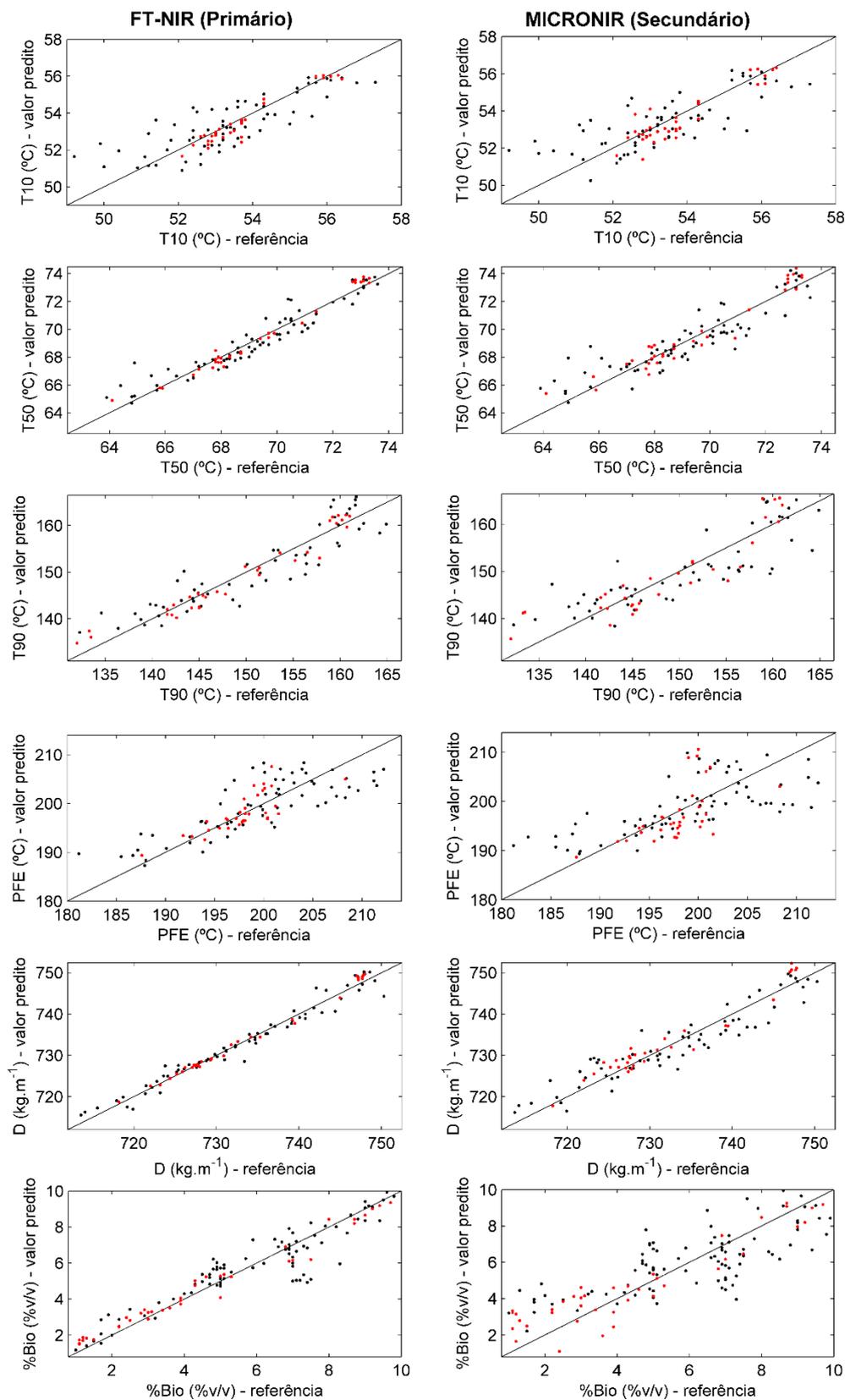


Figura 9. Gráficos de valores preditos vs. referência para os modelos PLS construídos no instrumento primário e secundário na região de segundo sobretom de estiramentos de ligações C-H. As amostras de calibração estão em preto e de validação em vermelho.

4.1.2 Avaliação do efeito da seleção de variáveis no desempenho dos modelos PLS na região de segundo sobretom, para os instrumentos de bancada e ultracompacto.

Duas técnicas de seleção de variáveis foram utilizadas na região de segundo sobretom com o objetivo de melhorar o desempenho dos modelos em ambos os instrumentos, o método *Jack-Knife* e o método *interval-PLS*. A Tabela 2 apresenta as figuras de mérito e as características dos modelos PLS desenvolvidos nos instrumentos de bancada e ultracompacto, com as variáveis selecionadas.

Tabela 2. Figuras de mérito para os modelos PLS construídos na região de segundo sobretom após seleção de variáveis.

		PLS- <i>Jack-Knife</i>					<i>i</i> -PLS				
		VE	VL	R ²	RMSEP	<i>bias</i>	VE	VL	R ²	RMSEP	<i>bias</i>
Bancada	T10	314	3	0,78	0,4	-0,26**	160	3	0,79	0,4	-0,15**
	T50	247	3	0,96	0,5	-0,049	120	4	0,97	0,3	-0,018
	T90	220	6	0,94	2,4	0,21	40	3	0,94	2,0	-0,47
	PFE	224	4	0,78	2,5	0,086	60	3	0,78	2,5	0,15
	D	280	3	0,99	0,9	0,27	80	6	0,99	0,6	0,062
	%Bio	269	3	0,96	0,43	0,066	120	6	0,97	0,34	-0,081
Ultracompacto	T10	11	2	0,76	0,5	-0,13	8	4	0,76	0,7	-0,17
	T50	9	4	0,88	0,9	0,088	20	5	0,93	0,7	0,035
	T90	8	2	0,78	4,2	0,96	12	2	0,77	4,3	1,64**
	PFE	9	4	0,66	3,3	0,79	12	4	0,56	4,6	-0,17
	D	14	2	0,94	2,5	1,049**	25	3	0,94	2,7	1,26**
	%Bio	12	4	0,80	1,39	0,41	18	6	0,87	0,94	0,30**

* VE, Variáveis espectrais; VL, Variáveis latentes; T10, T50, T90 e PFE - temperaturas de destilação a 10%, 50%, 90% de volume recuperado (°C); PFE – ponto final de ebulição (°C); D – densidade (kg/m³); %Bio – teor de biodiesel (% v/v).

** *bias* estatisticamente significativo (em um nível de significância de 95%).

De maneira geral, os modelos construídos com as variáveis selecionadas apresentaram uma diminuição no número de variáveis latentes em ambos os instrumentos, primário e secundário. Após a seleção de variáveis pelo método *Jack-Knife*, apenas o modelo do teor de biodiesel para o NIR ultracompacto apresentou um valor de RMSEP significativamente maior do que o modelo desenvolvido com a região de segundo sobretom, com base em um teste F a um nível de 95% de confiança. Todos os outros modelos não apresentaram diferença estatisticamente significativa no valor de RMSEP. Com respeito ao *bias*, apenas em dois parâmetros foram observados valores significativos, T10 (bancada) e Densidade (ultracompacto). Entre os modelos construídos com as variáveis *i*PLS, em dois

dos parâmetros (Densidade e Teor de biodiesel modelados no NIR de bancada) foram observados valores de RMSEP significativamente menores em relação aos respectivos modelos construídos na região de segundo sobretom, embora os valores ainda sejam próximos. Os valores de *bias*, no caso do método *i*PLS, apresentaram significância para os modelos PLS do parâmetro T10, construído no NIR de bancada, e dos parâmetros T90, Densidade e Teor de biodiesel, construídos no NIR ultracompacto. Os valores de RMSEP dos modelos PLS de T90 construídos no ultracompacto, particularmente, não apresentaram diferença significativa após a seleção de variáveis, porém são mais altos do que a reprodutibilidade do método de referência, utilizando tanto o *Jack-Knife* como o *i*PLS. As reprodutibilidades dos respectivos métodos de referência para os parâmetros de qualidade da gasolina, assim como as normas recomendadas, estão mostradas na Tabela 3.

Tabela 3. Reprodutibilidades dos métodos de referência.

Parâmetro de qualidade	Método de referência	Reprodutibilidade
Temperatura de destilação a 10% de volume recuperado (T10)	ASTM D 86	3,2°C
Temperatura de destilação a 50% de volume recuperado (T50)	ASTM D 86	1,9°C
Temperatura de destilação a 90% de volume recuperado (T90)	ASTM D 86	4,1°C
Ponto final de ebulição (PFE)	ASTM D 86	6,8°C
Densidade a 20 °C (D)	ASTM D 4052	1,9 - 3,2 kg/m ³

Os métodos de seleção de variáveis promoveram uma melhora pequena no desempenho de apenas dois dos modelos. Portanto, deste ponto em diante, todos os modelos PLS utilizados para determinar os parâmetros da gasolina e das misturas diesel/biodiesel, assim como os procedimentos de transferência de calibração, foram construídos com base na região espectral de segundo sobretom das ligações C-H (ver Figura 8).

4.1.3 Comparação dos modelos PLS entre os dois instrumentos e com os métodos de referência.

Uma comparação entre os modelos desenvolvidos com o NIR de bancada e com o ultracompacto mostra que o desempenho dos modelos PLS construídos com o NIR ultracompacto é inferior. Novamente com base em um teste F, em um nível de 95% de significância, todos os modelos PLS construídos com o NIR de bancada atingiram valores de RMSEP significativamente menores em relação aos modelos construídos com os espectros do NIR ultracompacto. O melhor desempenho alcançado pelos modelos desenvolvidos no NIR de bancada já era esperado devido às diferenças na resolução espectral em relação ao ultracompacto, que pode ser observado na Figura 8 (pág. 63).

Uma comparação entre os valores de RMSEP e a reprodutibilidade dos métodos de referência mostrou que os modelos PLS dos parâmetros de qualidade da gasolina em ambos os instrumentos, bancada e ultracompacto, atingiram valores de RMSEP menores ou equivalentes à reprodutibilidade dos métodos de referência, para todos os parâmetros de qualidade. O NIR ultracompacto, portanto, é capaz de gerar modelos PLS para realizar previsões dos parâmetros de qualidade da gasolina com um nível de confiabilidade comparável à reprodutibilidade dos métodos de referência.

Essa comparação não pode ser realizada para as misturas diesel/biodiesel porque essas amostras foram criadas por meio da adição de mais biodiesel (B100), considerando uma proporção volume/volume, às amostras de diesel/biodiesel coletadas pelo programa de monitoramento PMQC. Isso ocorreu devido à pouca disponibilidade de diesel e biodiesel puros. Além disso, o método de referência EN 14078 para determinação do teor de biodiesel, o qual consiste em um método de regressão univariada na região do infravermelho médio, é bastante influenciado pela presença de óleo vegetal. Por esse motivo, uma incerteza considerável está associada aos valores de referência do teor de biodiesel dessas amostras. Portanto, o foco para a parte deste trabalho que envolve as amostras de biodiesel está principalmente na avaliação do desempenho da técnica proposta para realizar transferência de calibração do instrumento de bancada para o ultracompacto.

4.1.4 Procedimento para transferência de calibração

Todos os padrões virtuais para as amostras de gasolina e de misturas diesel/biodiesel foram calculados de acordo com o tópico 3.1.5.1. Espectros de dez componentes químicos puros foram utilizados no cálculo dos padrões virtuais da gasolina: etanol, pentano, 1-penteno, hexano, 1-hexeno, heptano, tolueno, isooctano, xileno e etilbenzeno. No cálculo dos padrões virtuais das misturas diesel/biodiesel foram utilizados espectros de cinco componentes químicos puros: pentano, hexano, heptano, n-hexadecano e padrão de biodiesel (B100). A escolha desses componentes foi baseada em quão bem as bandas de absorção presentes nos seus espectros correspondem às bandas de absorção presentes nos espectros dos combustíveis. A faixa espectral foi dividida em segmentos (intervalos de números de onda) com o objetivo de obter padrões virtuais o mais semelhante possível aos espectros experimentais. O procedimento matemático descrito no tópico 3.1.5.1. foi realizado em cada um desses segmentos separadamente. Após alguns testes, verificou-se que o número adequado de segmentos foi dois ($9141,4\text{--}8425,7\text{ cm}^{-1}$; $8382,0\text{--}7629,3\text{ cm}^{-1}$) para os espectros de gasolina e três ($8889,6\text{--}8469,9\text{ cm}^{-1}$; $8425,6\text{--}8211,3\text{ cm}^{-1}$; $8169,8\text{--}7776,2\text{ cm}^{-1}$) para os espectros das misturas diesel/biodiesel, considerando que quanto melhor a aproximação entre os espectros experimentais e os padrões virtuais mais confiável é o desempenho do método de transferência de calibração.

Como mostrado na Figura 10, os padrões virtuais atingiram uma forma bastante semelhante aos respectivos espectros experimentais. As diferenças são mais pronunciadas no instrumento secundário (Figura 10b e 10d), porém, como mostra o detalhe na Figura 10a, existem pequenas diferenças mesmo nos espectros do instrumento primário (Figuras 10a e 10c). Como este comportamento foi observado tanto para o conjunto de espectros da gasolina quanto diesel/biodiesel, concluiu-se que uma representação satisfatória dos espectros instrumentais foi atingida pelos padrões virtuais.

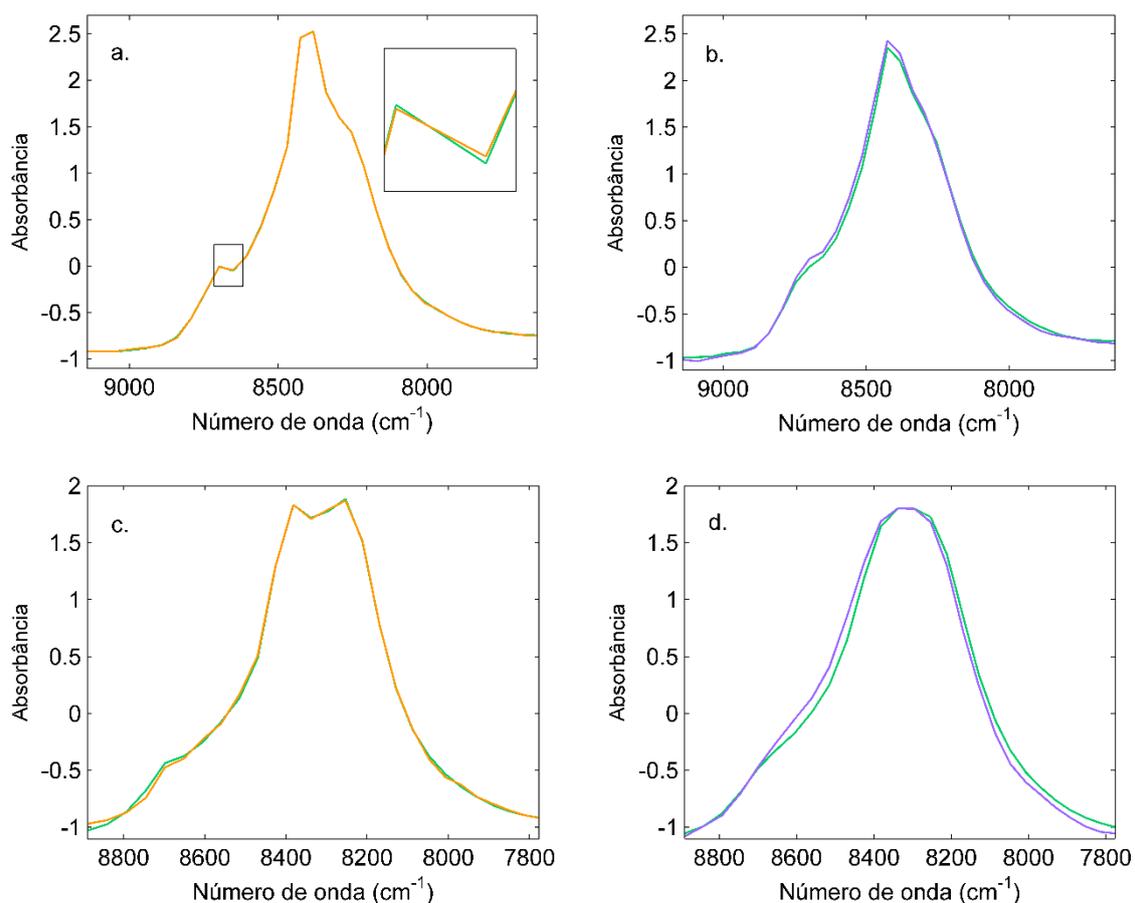


Figura 10. Padrões virtuais (verde) sobrepostos aos respectivos espectros experimentais de gasolina (a,b) e misturas diesel/biodiesel (c,d) para o instrumento primário (laranja) e secundário (azul).

O conjunto de transferência experimental foi selecionado levando-se em conta o número mínimo de amostras que geraram um valor de RMSEP menor ou o mais próximo possível da reprodutibilidade dos respectivos métodos de referência. A Figura 11 mostra como os valores de RMSEP obtidos se comportaram após realizar a padronização reversa (RS) utilizando diferentes números de espectros experimentais de transferência. É possível observar, portanto, que para a maioria dos parâmetros os valores de RMSEP atingem a reprodutibilidade dos métodos de referência na região próxima a 20 padrões experimentais.

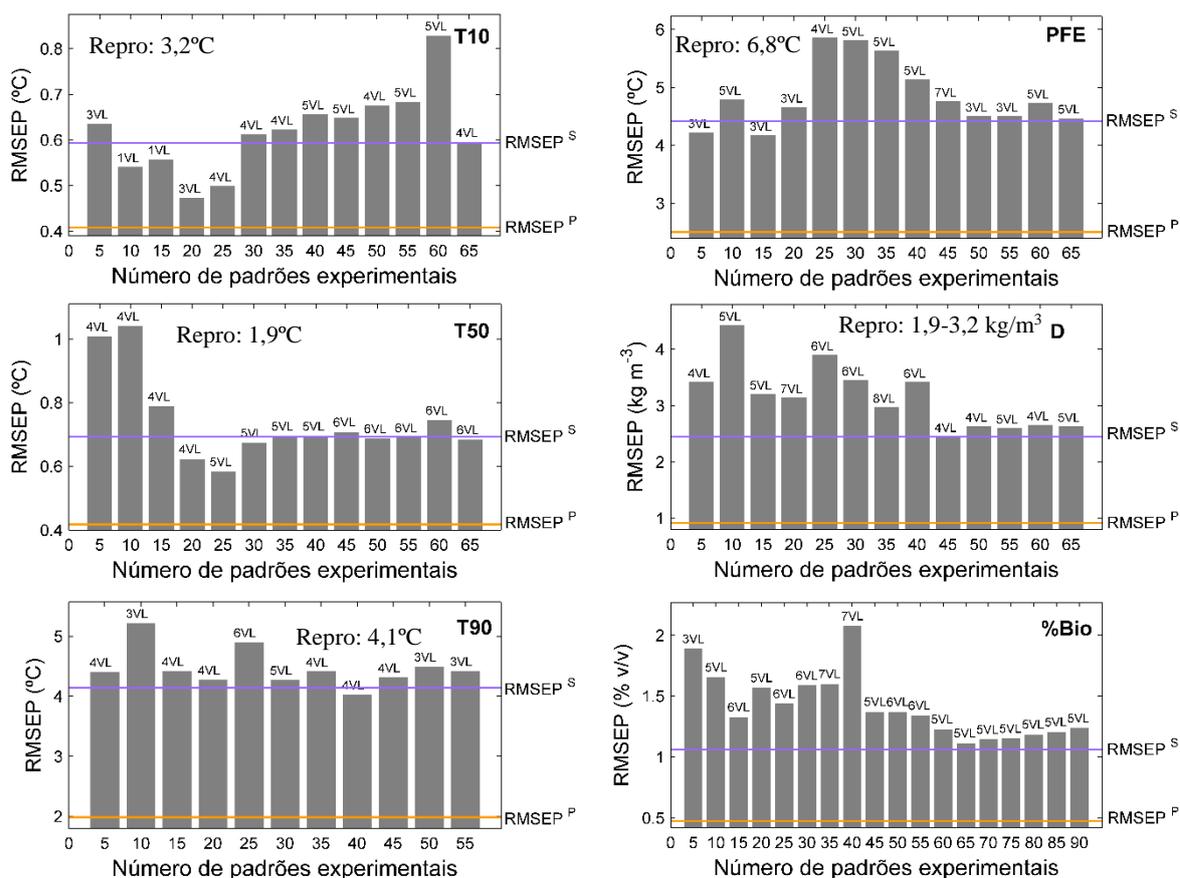


Figura 11. Comportamento dos valores de RMSEP em relação ao número de padrões após realizar o método de padronização reversa (RS) utilizando padrões experimentais. VL = Variável Latente; Repro: Reprodutibilidade do método de referência; RMSEP^P e RMSEP^S = Valores de RMSEP atingidos pelos modelos PLS construídos no instrumento primário (linha laranja) e secundário (linha azul), respectivamente.

No caso dos padrões virtuais, o número de espectros selecionados foi sempre igual ao número de amostras do conjunto de calibração de cada modelo. Pois, nesse caso, quanto maior o número de padrões virtuais menor é o RMSEP. Esse número maior de padrões virtuais, que representa pelo menos 3 vezes o número de padrões experimentais, não representa uma desvantagem para realizar RS utilizando padrões virtuais. Na verdade, uma vez que os espectros dos solventes foram adquiridos, o procedimento para calcular os padrões virtuais é baseado apenas em procedimentos matemáticos, como explica o tópico 3.1.5.1.

As amostras que compõem os conjuntos de calibração e validação externa no método RS utilizando padrões experimentais são as mesmas no método RS utilizando padrões virtuais. A Tabela 4 apresenta as figuras de mérito para os valores preditos para os espectros

de validação adquiridos no instrumento secundário, utilizando os modelos PLS construídos com os espectros de calibração adquiridos no instrumento primário, depois de realizar a padronização reversa empregando tanto os padrões experimentais como os padrões virtuais.

Tabela 4. Figuras de mérito após realizar a padronização reversa utilizando padrões experimentais ou padrões reais.

	RS (padrões virtuais)					RS (padrões experimentais)				
	VL	n° de padrões	R ²	RMSEP	<i>bias</i>	VL	n° de padrões	R ²	RMSEP	<i>bias</i>
T10 (°C)	2	69	0.68	1.120	-0.76*	3	20	0.76	0.473	-0.11
T50 (°C)	2	69	0.84	1.488	-0.84*	4	20	0.90	0.622	0.15
T90 (°C)	2	60	0.79	4.021	1.31	4	20	0.83	4.280	1.27
PFE (°C)	2	69	0.56	3.933	-0.069	3	15	0.58	4.172	-0.25
D (kg/m ³)	2	69	0.89	3.734	2.05*	7	20	0.96	3.147	0.78
%Bio (%)	3	90	0.60	2.090	0.008	6	15	0.88	1.328	0.62*

* *bias* estatisticamente significativo (em um nível de significância de 95%).

Para os parâmetros de temperatura de destilação a 90% de volume recuperado (T90), ponto final de ebulição (PFE) e densidade a 20 °C, a padronização utilizando padrões virtuais atingiu valores de RMSEP similares aos obtidos com a padronização utilizando padrões experimentais. O valor de RMSEP para o parâmetro densidade, em particular, teve um aumento relativo de pelo menos 31%, de 2,4 kg/m³, no modelo PLS do instrumento ultracompacto (Tabela 1) para 3,7 kg/m³, após a realização da padronização reversa utilizando os padrões virtuais (Tabela 4). Para o teor de biodiesel e para as temperaturas de destilação a 10% e 50% de volume recuperado, os valores de RMSEP obtidos após a padronização utilizando padrões virtuais foi significativamente maior em relação à padronização utilizando padrões experimentais. A Figura 12 faz uma comparação entre os gráficos dos valores preditos *versus* valores de referência para os modelos PLS apresentados na Tabela 4. Os gráficos de previsão vs. referência (Figura 12), após a padronização com padrões experimentais, para os parâmetros T10, T50 e principalmente %Bio, mostram o reflexo dos valores RMSEP obtidos, que foram menores em comparação com a padronização utilizando padrões virtuais. Do mesmo modo, os modelos para os parâmetros T90, PFE e Densidade, que atingiram valores de RMSEP bastante semelhantes para os dois casos, padrões experimentais e virtuais, mostraram um comportamento também semelhante nos gráficos de previsão vs. referência.

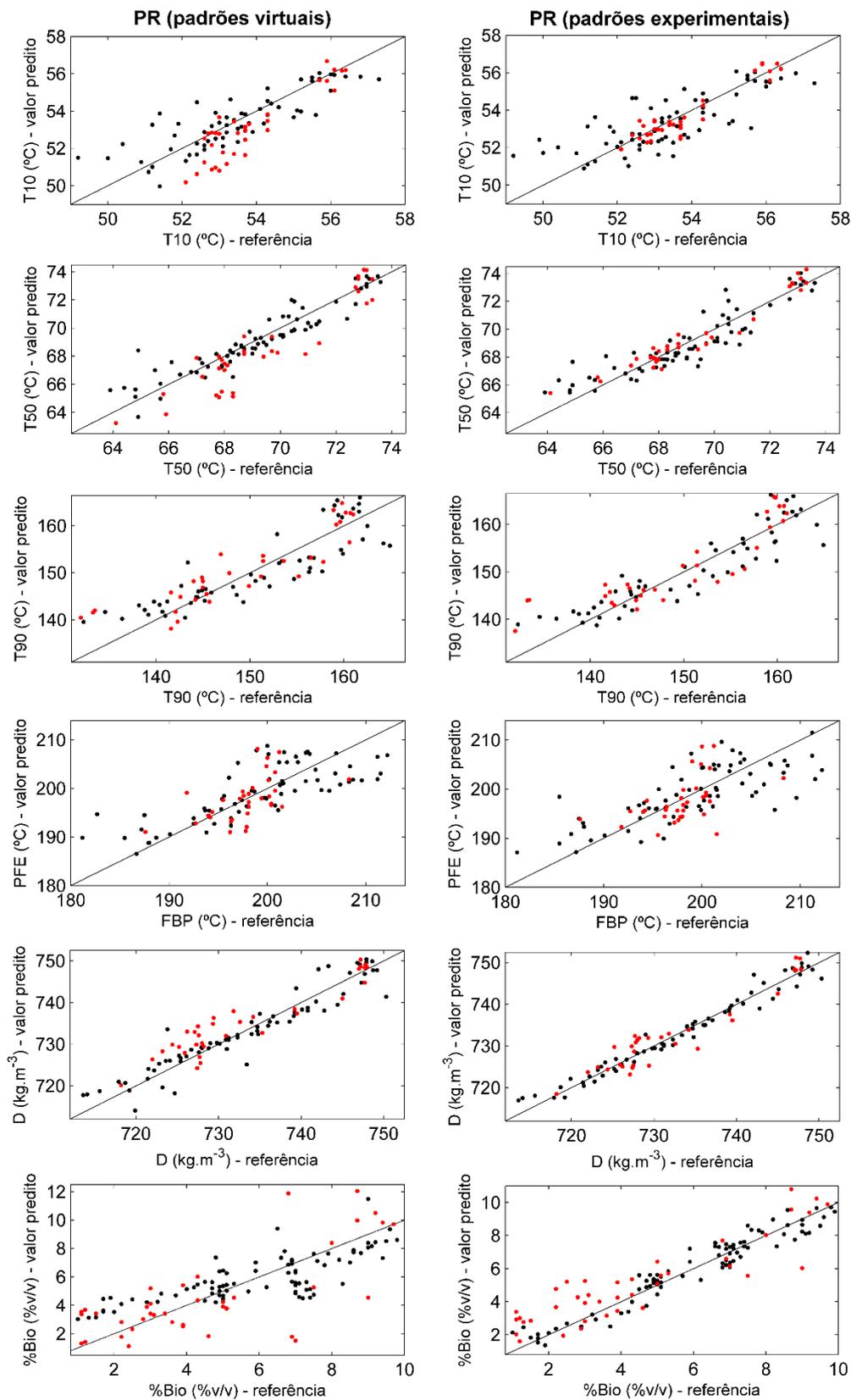


Figura 12. Valores preditos para os espectros de validação (vermelho) adquiridos no instrumento secundário, pelos modelos PLS construídos com os espectros do instrumento primário (preto).

De acordo com um teste t (ASTM E 1655-05 [62]), os parâmetros T10, T50 e densidade apresentaram *bias* estatisticamente significativos após a padronização utilizando padrões virtuais. No entanto, comparando o *bias* após a padronização com padrões virtuais com o *bias* obtido após a padronização com os padrões experimentais, foram observadas diferenças estatisticamente significativas (teste t) apenas para dois dos parâmetros, T10 e T50. Esses *bias* podem ser facilmente corrigidos por meio de uma correção de *slope* e *bias*.

Para realizar um procedimento simples de correção *slope* e *bias* (SBC) e poder provar esse ponto, 5 amostras foram excluídas do conjunto de validação externa e utilizadas para realizar a correção SBC nas amostras restantes. Os novos valores de RMSEP, para as 29 amostras de validação restantes, passaram a ser 1,15 °C e 1,48 °C, para T10 e T50, respectivamente. Enquanto os valores de *bias* passaram a ser -0,85 °C e -0,88 °C para T10 e T50, respectivamente. Os espectros medidos no NIR ultracompacto e os valores de referência para esses dois parâmetros foram utilizados para realizar a correção SBC. Esse procedimento foi realizado ajustando-se uma linha de regressão para os pontos do gráfico dos valores preditos para as cinco amostras excluídas (utilizando os espectros obtidos no NIR ultracompacto e o modelo PLS após ser aplicada a padronização com padrões virtuais) contra os seus respectivos valores de referência. Com isso, o valor do *slope* e do *bias* obtidos na regressão foram usados para ajustar os valores preditos das 29 amostras restantes no conjunto de validação externa. Portanto, é importante mencionar para aplicar esse método SBC é necessário apenas adquirir os espectros de algumas amostras no instrumento secundário e seus respectivos valores de referência. Desse modo, a vantagem de utilizar os padrões virtuais, evitando assim o transporte e armazenagem de amostras de combustíveis, é absolutamente mantida. Esse tipo de correção *slope* e *bias* é diferente do que foi utilizado nos trabalhos de Cooper e colaboradores [19-21], pois o método proposto por esses autores requer que sejam obtidos valores preditos para os padrões virtuais reconstruídos no instrumento secundário pelo modelo construído no instrumento primário. No caso do método descrito aqui, os instrumentos são extremamente diferentes e não é possível utilizar o modelo construído no NIR de bancada para realizar previsões a partir dos espectros obtidos no NIR ultracompacto, sem que antes seja feita uma transferência de calibração. Depois de fazer a correção utilizando o método SBC (utilizando as 5 amostras), portanto, os valores de RMSEP foram reduzidos para 0,60 °C e 1,05 °C para T10 e T50, respectivamente. Os valores de *bias* foram reduzidos para -0,05 °C e -0,21 °C, respectivamente. Esses valores de *bias* corrigidos, por sua vez, não são estatisticamente significativos e não há diferença

estatisticamente significativa (teste F, 95% de confiança) entre o valor de RMSEP corrigido para T10 e o valor obtido após a padronização com padrões experimentais. Entretanto, o valor de RMSEP corrigido para T50 continua estatisticamente diferente apesar de ter sido consideravelmente reduzido. E, além disso, é importante notar que para ambas as abordagens de padronização, utilizando padrões experimentais ou virtuais, o valor de RMSEP para T50 é menor do que a reprodutibilidade do método de referência. A correção SBC também foi realizada para os modelos de densidade, após padronização utilizando padrões virtuais, onde um *bias* estatisticamente significativo também foi observado. Nesse caso, os valores de RMSEP e *bias* para o conjunto de validação externa após excluir as 5 amostras (28 amostras restantes, nesse caso) foram $3,9 \text{ kg/m}^3$ e $2,0 \text{ kg/m}^3$, respectivamente. Após a correção SBC, os valores caíram para $3,3 \text{ kg/m}^3$ e $-0,13 \text{ kg/m}^3$. Apesar de o valor de *bias* corrigido não ser estatisticamente significativo, o valor de RMSEP ainda é maior do que a reprodutibilidade do método de referência para densidade. A Figura 13 faz uma comparação dos gráficos de referência vs. previsão para os parâmetros T10, T50 e densidade, antes e após a correção SBC. As 5 amostras utilizadas na correção estão circuladas nos gráficos.

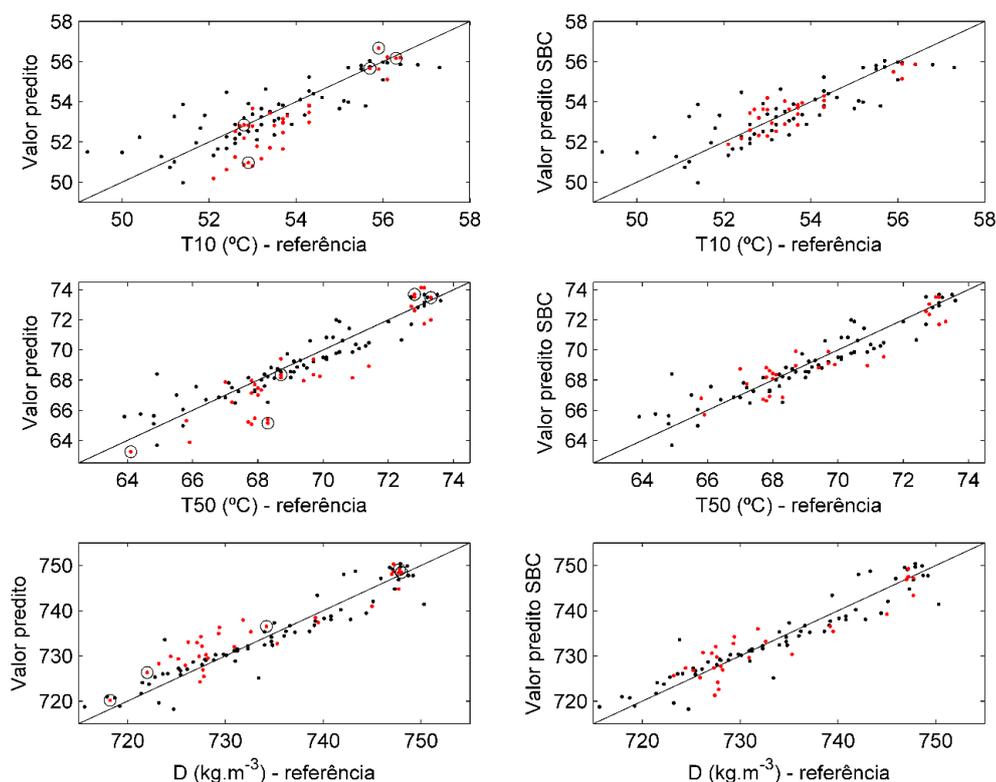


Figura 13. Gráfico de referência vs. valores preditos antes (esquerda) e após (direita) a correção SBC para os parâmetros T10, T50 e densidade. As 5 amostras utilizadas para realizar a correção estão circuladas nos gráficos. As amostras de calibração estão em preto e de validação em vermelho.

O número de graus de liberdade utilizado em todos os testes F e testes t (para verificar a significância estatística dos valores de *bias*) realizados correspondem ao número de amostras de previsão de cada parâmetro. Adicionalmente, o número de graus de liberdade para os testes t realizados na comparação entre valores de *bias* de modelos diferentes foram calculados como recomenda a referência Miller e Miller (teste t para comparação entre duas médias) [63]. Todos os testes foram realizados com um nível de 95% de confiança.

Com exceção dos valores de RMSEP obtidos para o parâmetro T90 após padronização com padrões experimentais (4.279 °C) e para a densidade após padronização com padrões virtuais (3.734 kg/m³ e 3.349 kg/m³ após correção SBC), todos os RMSEPs para os parâmetros de qualidade da gasolina foram menores do que a reprodutibilidade dos respectivos métodos de referência. O valor de RMSEP atingido pelos modelos PLS do teor de biodiesel não foram comparados com o método de referência pelas razões já discutidas anteriormente. Porém, embora a padronização reversa utilizando padrões experimentais tenha se mostrado mais eficiente em prever o teor de biodiesel, ambas as abordagens utilizando padrões experimentais e virtuais, tornaram possível utilizar os modelos PLS construídos em um espectrômetro de alta resolução para realizar previsões do teor de biodiesel a partir de espectros de um espectrômetro ultracompacto.

4.2 Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR

4.2.1 Experimentos preliminares utilizando gasolinas formuladas em laboratório

Antes de iniciar os testes com as gasolinas diretamente coletadas na linha de produção, alguns experimentos preliminares foram conduzidos com gasolinas formuladas em laboratório. Esses experimentos foram focados em: 1) verificar qual a variabilidade espectral obtida para uma mesma formulação de gasolina sendo reproduzida mais de uma vez e; 2) avaliar se a alta volatilidade das correntes de hidrocarbonetos, especialmente a de hidrocarbonetos leves, pode provocar diferenças relevantes entre os espectros da mistura física e os respectivos espectros simulados.

4.2.1.1 VARIABILIDADE ESPECTRAL PARA REPETIÇÕES DE UMA GASOLINA FORMULADA EM LABORATÓRIO

Uma formulação de gasolina, M1, definida com base no histórico de formulações da empresa, foi reproduzida no laboratório utilizando três correntes de hidrocarbonetos coletadas no mesmo dia. As misturas foram feitas em triplicata por dois analistas, totalizando 6 amostras de gasolina da mesma formulação. Espectros NIR das 6 gasolinas foram adquiridos no NIR ultracompacto e estão apresentados na Figura 14. A derivada foi o pré-processamento utilizado pois se mostrou o mais adequado para os espectros do NIR ultracompacto, de acordo com o desempenho dos modelos PLS que serão discutidos posteriormente.

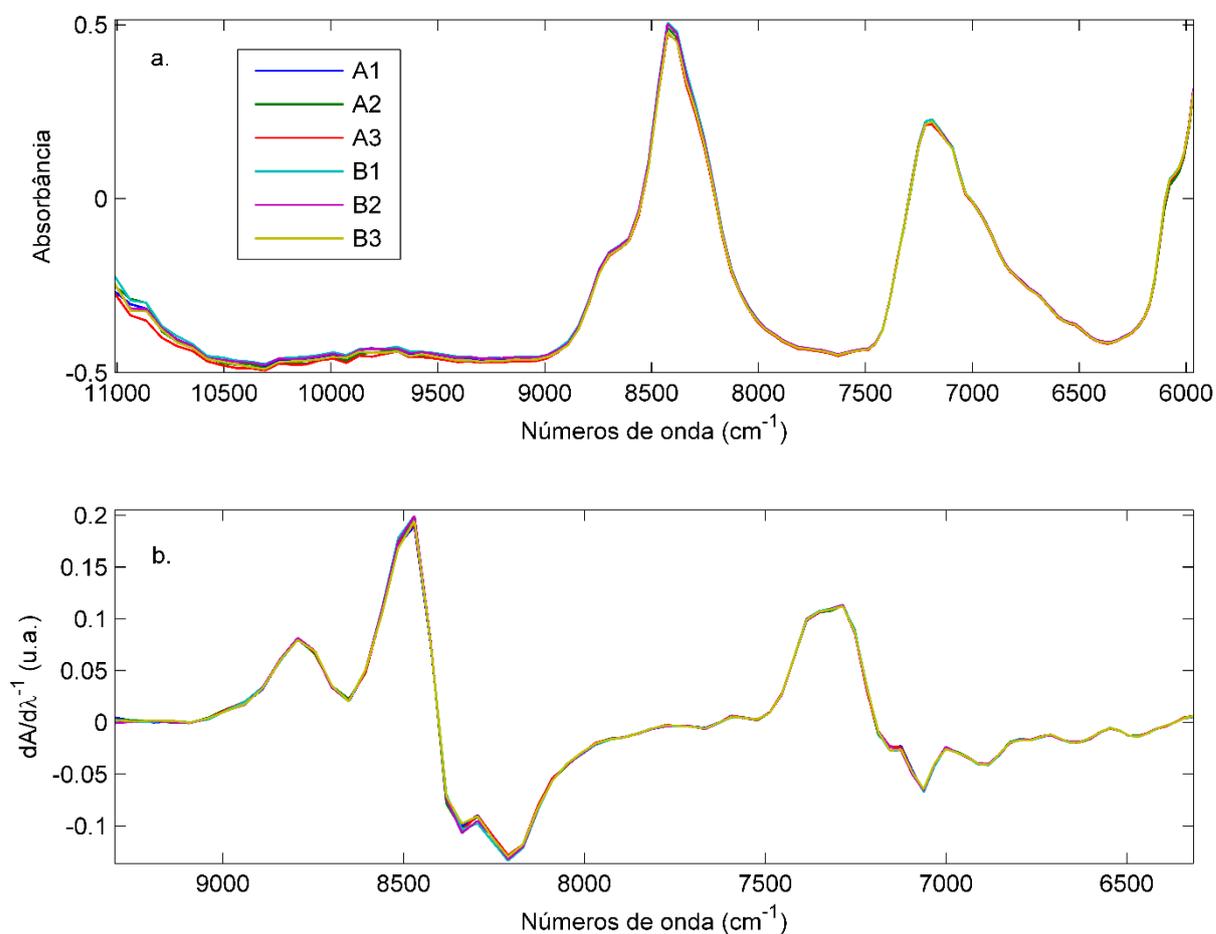


Figura 14. Espectros NIR das repetições em triplicata para a gasolina formulada em laboratório pelos analistas A e B. (a.) Espectros brutos e (b.) derivativos (1ª derivada com polinômio de 2º grau e janela de suavização de 3 pontos).

Embora os espectros pré-processados sejam visualmente muito semelhantes, uma avaliação mais objetiva foi feita realizando-se uma análise de componentes principais nos espectros das repetições, como mostra a Figura 15. Não é possível notar uma tendência clara de agrupamento entre as amostras no gráfico dos escores de PC1 vs. PC2 (Figura 15a.). O gráfico do T^2 de Hotelling vs. Q Residual (Figura 15b.) também mostra que nenhuma repetição apresentou comportamento extremo. De acordo com as Figuras 15c. e 15d., os *loadings* mais elevados em PC1 e PC2 são devidos justamente às variações de intensidade próximas do máximo das bandas de absorção. Dessa forma, apesar da alta volatilidade das correntes de hidrocarbonetos, a variabilidade observada entre os espectros é pequena e se deve, principalmente, às pequenas diferenças de proporção entre as correntes nas repetições.

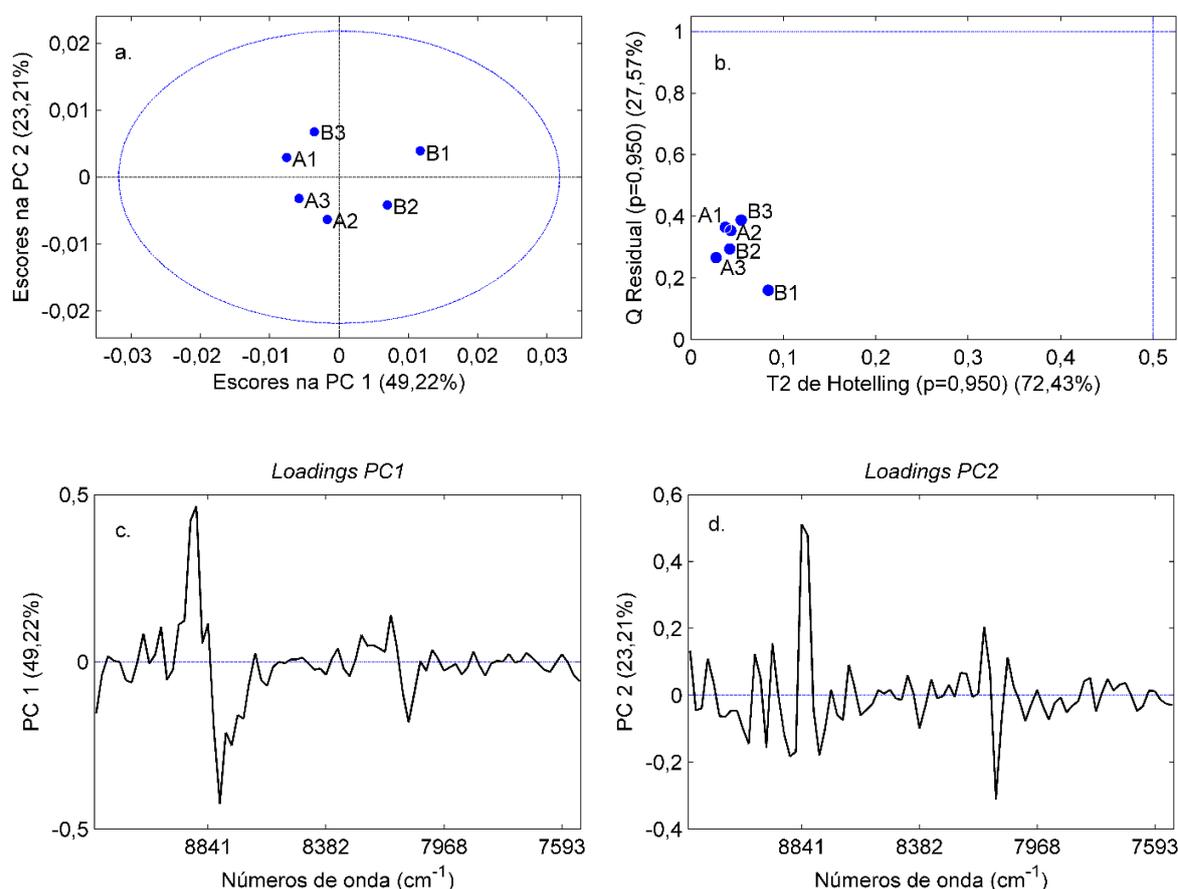


Figura 15. Análise de componentes principais das repetições da gasolina M1 formulada em laboratório.

4.2.1.2 AVALIAÇÃO DO COMPORTAMENTO DOS ESPECTROS SIMULADOS EM RELAÇÃO AOS EXPERIMENTAIS

No segundo experimento, realizado para verificar o comportamento dos espectros simulados em relação aos respectivos espectros das misturas físicas realizadas em laboratório, foram reproduzidas além da M1, mais três diferentes formulações, M2, M3 e M4. E além disso, foram adquiridos os espectros das três correntes de hidrocarbonetos utilizadas na formulação dessas gasolinas. Os espectros simulados (S1-S4) referentes a cada gasolina formulada foram obtidos por meio de uma combinação linear entre os espectros das correntes, atribuindo-lhes pesos correspondentes às percentagens utilizadas na formulação de cada amostra. A Figura 16 mostra uma comparação entre os espectros adquiridos da mistura física e os espectros simulados (Figuras 16a. a 16d.), assim como os espectros das três correntes de hidrocarbonetos utilizadas (Figura 16f.). Os espectros das quatro misturas são mostrados juntos na Figura 16e. em que é possível perceber que os espectros M1 e M2 são parecidos, assim como os espectros M3 e M4. De fato, as duas primeiras formulações são do Tipo 1 e as outras duas do Tipo 2.

Uma análise de componentes principais foi realizada nos espectros experimentais e simulados (Figura 17). No gráfico de T^2 vs. Q residual (Figura 17a.) aparece uma tendência de agrupamento entre os espectros experimentais e as simulações, porém, não é observado nenhum comportamento anômalo. De acordo com o gráfico dos escores da PCA, mostrado na Figura 17b., a variabilidade dos dados é basicamente explicada pela primeira componente (~88%). Como mostram os *loadings* de PC1 (Figura 17c.), as regiões espectrais próximas a 8750, 8200 e 7300 cm^{-1} são as responsáveis por essa variabilidade, que está diretamente relacionada com os dois diferentes tipos de gasolina, tipo 1 e tipo 2. A PC2 é responsável por apenas 7% da variabilidade, que corresponde principalmente às pequenas diferenças entre os espectros experimentais e os espectros simulados, observada principalmente nas variáveis próximas aos comprimentos de onda 8900, 8650, 8150 e 7200 cm^{-1} , aproximadamente (Figura 17d.).

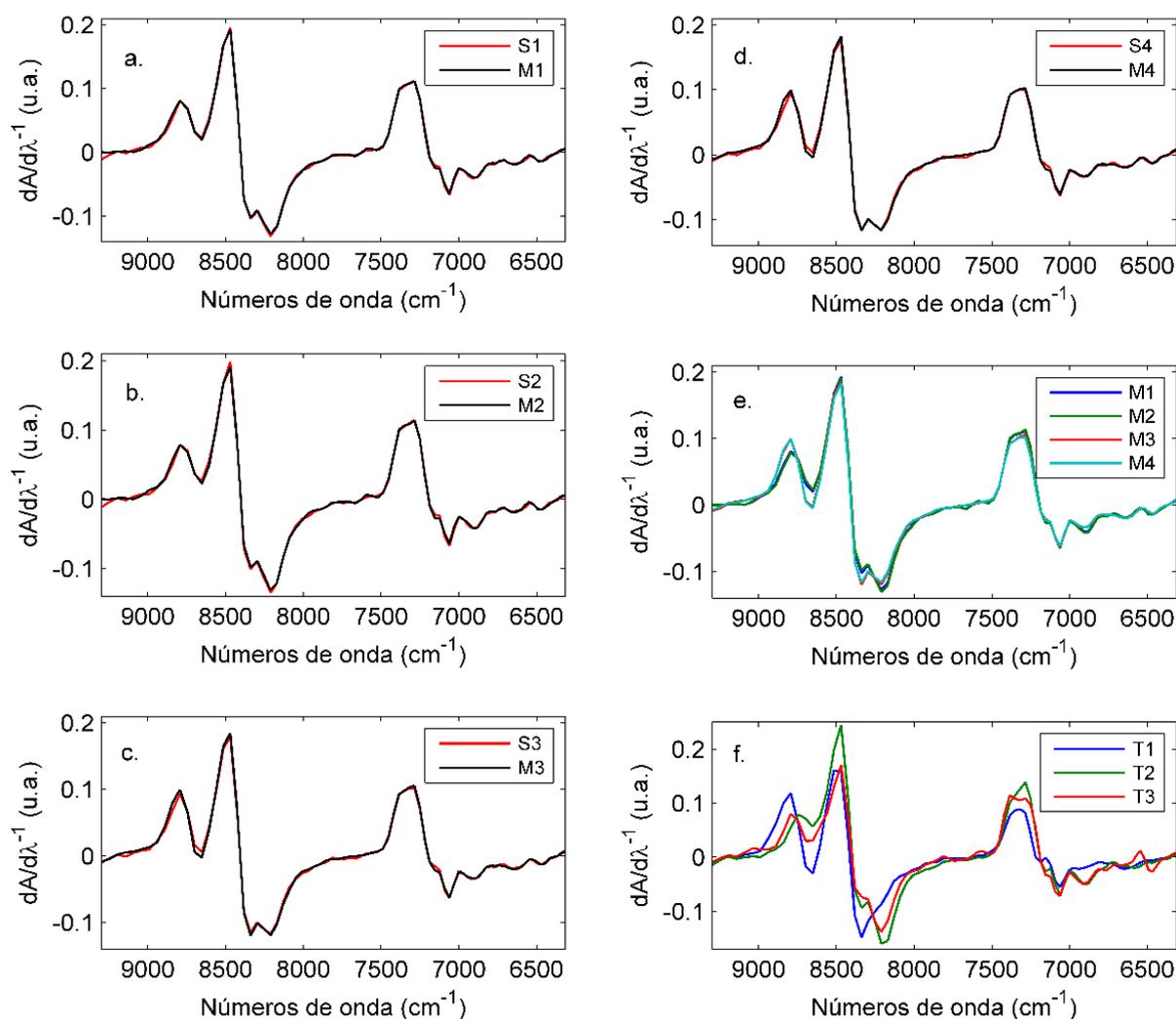


Figura 16. Comparação entre os espectros das amostras de gasolina formuladas em laboratório (M1-M4) e os espectros simulados (S1-S4) a partir dos espectros de três correntes de hidrocarbonetos (T1-T3).

As diferenças observadas entre os espectros simulados e experimentais podem ocorrer devido à volatilização de compostos durante a mistura física das correntes de hidrocarbonetos e, além disso, devido ao efeito causado pelos diferentes índices de refração das correntes de hidrocarbonetos e da gasolina em espectros de infravermelho. Embora tenham sido utilizadas apenas 4 amostras, a análise de componentes principais mostra que os espectros simulados são bastante parecidos com os experimentais. As diferenças entre eles podem ser minimizadas, se necessário, por meio de técnicas de transferência de calibração utilizando, evidentemente, um conjunto de dados maior.

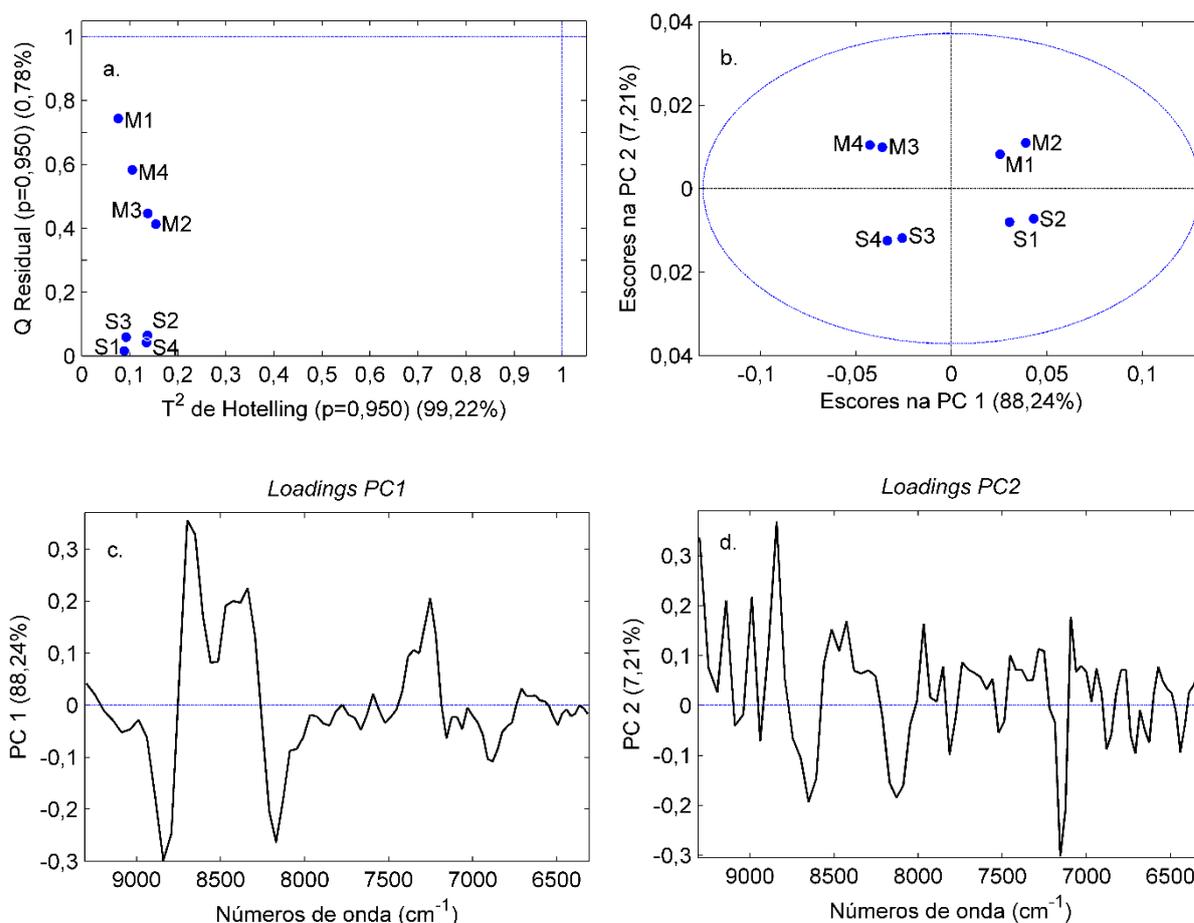


Figura 17. Análise de componentes principais dos espectros experimentais (M) e simulados (S) de quatro gasolinas formuladas em laboratório.

4.2.2 Apresentação e análise exploratória dos conjuntos de dados espectrais nos espectrômetros NIR de bancada e ultracompacto.

Os espectros brutos de todas as gasolinas automotivas e correntes de hidrocarbonetos coletadas na empresa após o início do projeto e analisadas nos espectrômetros de bancada e ultracompacto estão apresentados, no intervalo de 11000 a 4000 cm^{-1} , na Figura 18. Todos os espectros de gasolinas apresentados são de amostras do Tipo A, por isso não são observadas as bandas de primeiro sobretom e combinações de ligações O-H. Nos espectros do NIR de bancada, na região entre 5000 e 4000 cm^{-1} , onde tipicamente ocorrem bandas de combinação de hidrocarbonetos alifáticos, são observadas as bandas de combinações envolvendo ligações C-H e C-C. As bandas predominantemente atribuídas ao primeiro sobretom de ligações C-H também são facilmente observadas na região de 6.100-5.500 cm^{-1} . Em torno de 6200 cm^{-1} ocorre a banda, com pouca intensidade, de primeiro sobretom

de ligações C-H aromáticas. As bandas de primeiro sobreton das combinações de estiramento C-H (em torno de 7150 cm^{-1}) são observadas mais claramente nos espectros do NIR ultracompacto devido ao maior caminho ótico utilizado (20 mm). As bandas de segundo sobreton de estiramento C-H nos grupamentos CH, CH₂ e CH₃ também são facilmente observadas entre 9000 e 7800 cm^{-1} nos espectros do NIR ultracompacto [70].

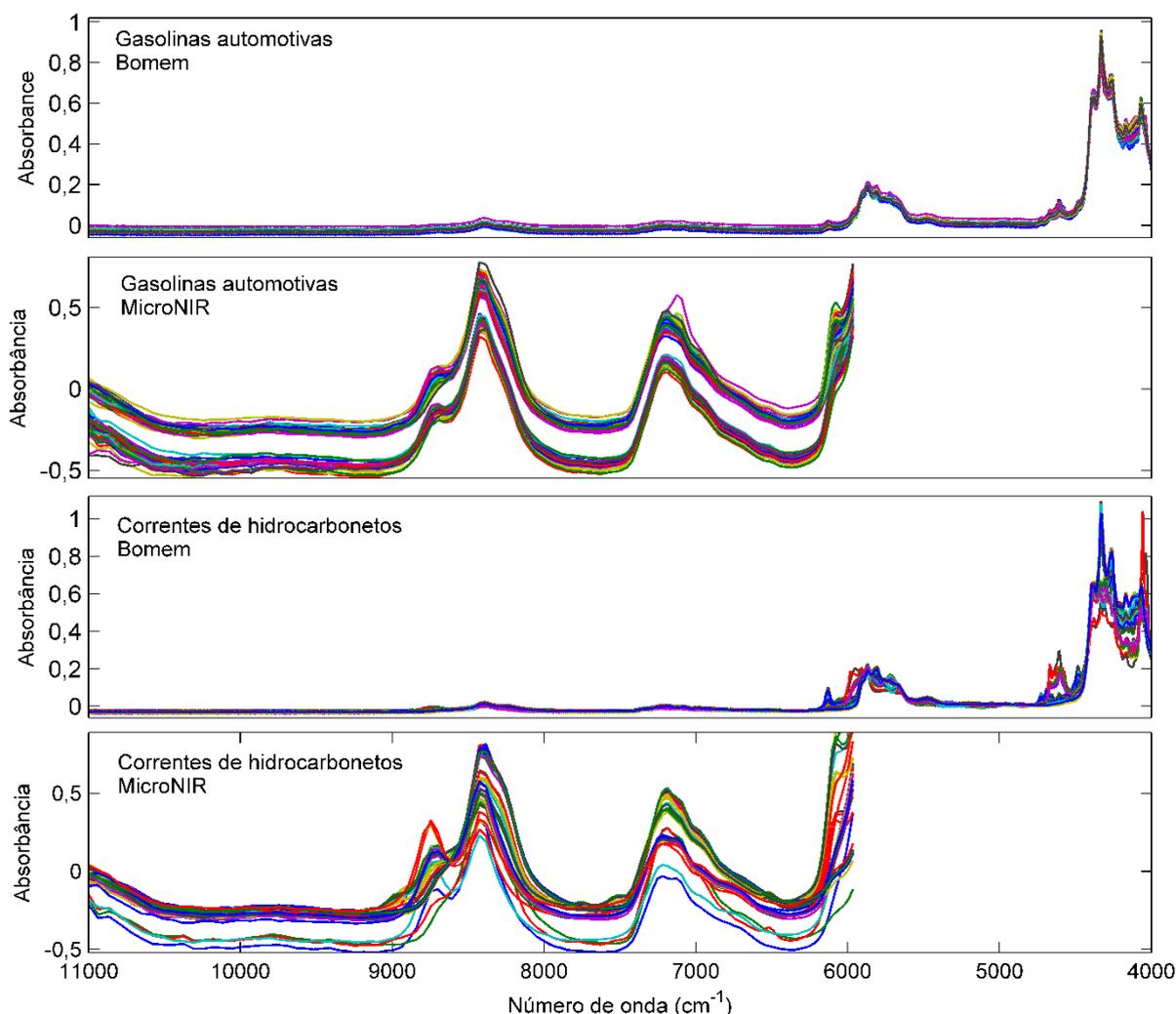


Figura 18. Espectros brutos das gasolinas e das correntes de hidrocarbonetos coletadas de 2015 a 2017 e analisadas no espectrometro de bancada e no ultracompacto.

Algumas técnicas de pré-processamentos foram testadas com base na observação do comportamento dos escores dos modelos PCA e no desempenho dos modelos CLS e de regressão PLS. O conjunto de processamentos mais adequado para os espectros do NIR de bancada foi uma suavização Savitzky-Golay, com janela de 25 pontos e polinômio de segunda ordem seguida por uma correção de linha de base, utilizando a região entre os

comprimentos de onda 9400 e 11300 cm^{-1} como referência para a linha de base. Para os espectros do NIR ultracompacto a primeira derivada com polinômio de segunda ordem e janela de 3 pontos foi a técnica mais adequada. Os espectros pré-processados das gasolinas para ambos os instrumentos estão apresentados na Figura 19. Além disso, os espectros estão coloridos de acordo com as gasolinas Tipo 1 e Tipo 2. Para entender um pouco melhor como esses dois tipos de gasolinas se comportam e o que caracteriza a diferença entre elas, foi realizada uma análise de componentes principais com os espectros de cada um dos equipamentos.

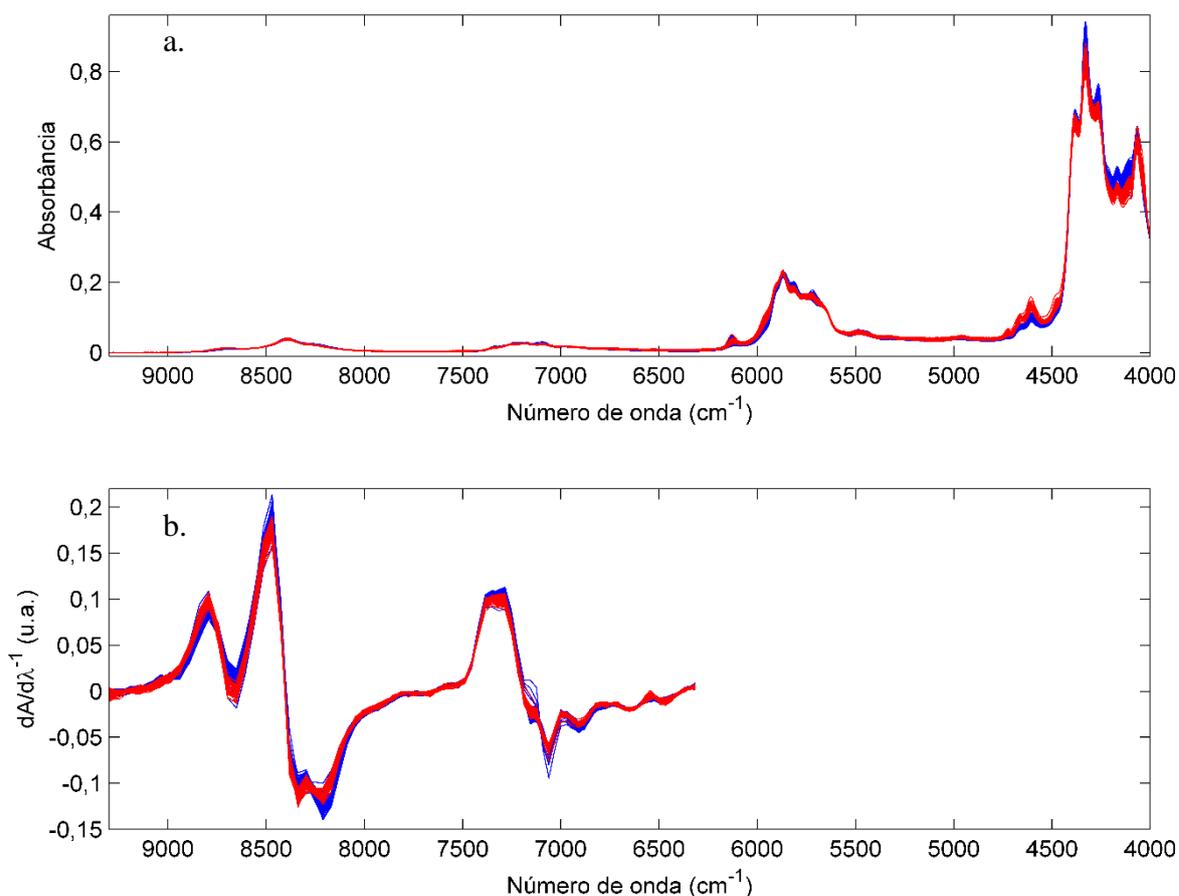


Figura 19. Espectros pré-processados das gasolinas coletados no espectrômetro de bancada (a.) e ultracompacto (b.). As cores azul e vermelho representam respectivamente as gasolinas do tipo 1 e 2.

Nos escores da PCA realizada com os espectros dos instrumentos de bancada e ultracompacto, Figuras 20 e 21 respectivamente, é possível notar que a maior fonte de variabilidade nos conjuntos de espectros é de fato devido aos dois diferentes tipos de gasolina. Isto é facilmente observado pela forte tendência de separação ao longo do eixo da

primeira componente principal, que explica ~75% (bancada) e ~60% (ultracompacto) da variabilidade dos dados. Embora algumas amostras apresentem um Q residual ou T2 de Hotelling um pouco acentuados, nenhuma amostra foi eliminada do modelo nessa etapa exploratória inicial. Esse comportamento foi observado em ambos os instrumentos. Para os espectros do NIR de bancada, os *loadings* da PC1 (Figura 20) mostram que as variáveis responsáveis por esta separação estão entre 4700 e 4000 cm^{-1} e próximas a 6000 cm^{-1} . Observando o conjunto de espectros na Figura 19a, nota-se que essas regiões possuem diferenças mais pronunciadas entre os dois tipos de gasolina.

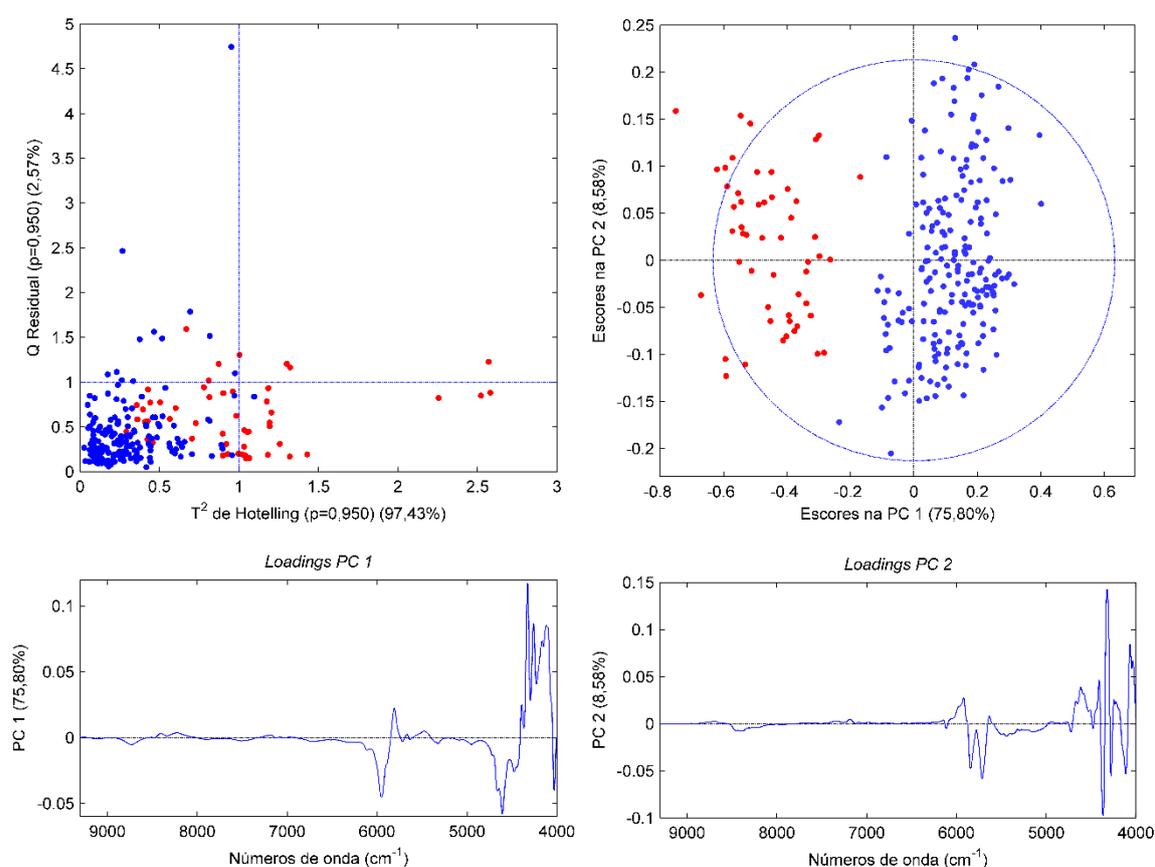


Figura 20. Análise de componentes principais dos espectros de gasolina automotiva adquiridos no espectrômetro de bancada. As cores azul e vermelho representam respectivamente as gasolinas do tipo 1 e 2.

Apesar de o NIR ultracompacto possuir uma faixa espectral mais estreita em relação ao de bancada, as diferenças entre os dois tipos de gasolina também podem ser facilmente notadas em toda a região de primeiro sobretudo das combinações e segundo sobretudo dos estiramentos C-H. Elas são mais pronunciadas na região entre 9000 e 8000 cm^{-1} , como

também mostram os *loadings* de PC1 para a análise de componentes principais realizada nos espectros do NIR ultracompacto na Figura 21. Portanto, devido a essas diferentes características espectrais observadas entre os dois tipos de gasolina assim como às diferentes faixas de octanagem, os modelos PLS discutidos mais adiante foram desenvolvidos separadamente para cada tipo de gasolina.

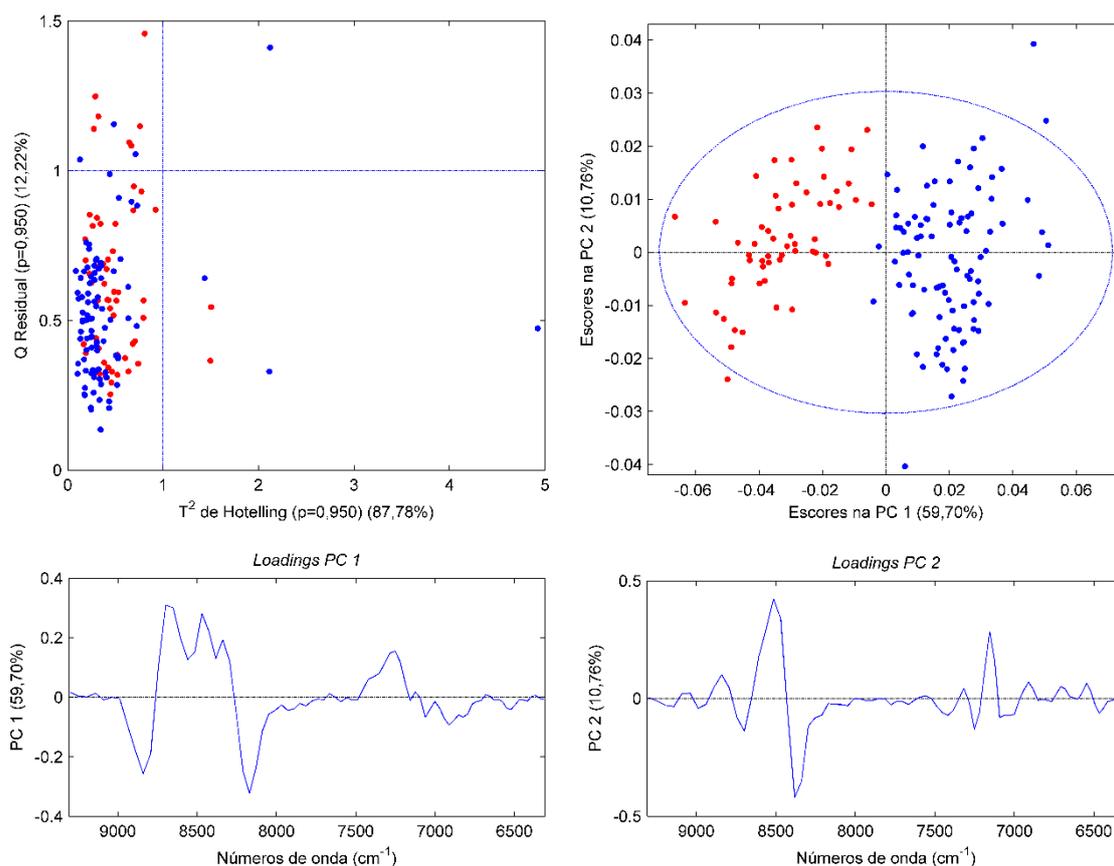


Figura 21. Análise de componentes principais dos espectros de gasolina automotiva adquiridos no espectrômetro ultracompacto. As cores azul e vermelho representam respectivamente as gasolinas do tipo 1 e 2.

4.2.3 Prevendo a formulação a partir dos espectros de gasolinas e das respectivas correntes de hidrocarbonetos

Uma vez que os experimentos com gasolinas formuladas em laboratório produziram resultados favoráveis, a estratégia imediata foi tentar estimar a proporção das correntes nas gasolinas produzidas diretamente na linha de produção. Embora, com a evolução do estudo esta estratégia não tenha sido utilizada definitivamente no simulador, os resultados obtidos com este teste são importantes para confirmar o potencial da espectroscopia NIR para a aplicação proposta neste trabalho.

Nesse caso, modelos de regressão por mínimos quadrados clássicos (*Classical Least Squares* - CLS) foram construídos utilizando como dados de entrada os espectros experimentais das gasolinas e os respectivos espectros das correntes de hidrocarbonetos. As estimativas dos espectros de gasolina e das proporções das correntes calculadas por CLS foram comparadas com os espectros experimentais e com as proporções reais utilizadas na linha de produção.

As proporções reais foram recuperadas do histórico de formulações utilizadas na produção de gasolina na empresa em um período de dois anos, de 2015 a 2017. Espectros NIR das amostras de gasolinas e das diferentes correntes de hidrocarbonetos utilizadas nas respectivas formulações foram adquiridos em diferentes pontos ao longo desse tempo. No total, incluindo gasolinas tipo 1 e 2, foram realizadas análises em 10 pontos com o espectrômetro NIR de bancada e em 13 pontos com o NIR ultracompacto. Em alguns desses pontos foram produzidas e analisadas diferentes gasolinas utilizando a mesma formulação. Os resultados obtidos pelos modelos CLS estão mostrados na Figura 22, para o NIR de bancada, e na Figura 23, para o ultracompacto.

Os espectros estimados por CLS são bastante próximos dos espectros experimentais, para ambos os espectrômetros, NIR de bancada e ultracompacto. Os valores dos resíduos mostrados nas figuras foram calculados para cada espectro de gasolina por meio da soma dos quadrados dos valores do vetor correspondente à diferença entre o espectro estimado e o espectro experimental. Os resíduos são baixos comparados aos valores espectrais, da ordem de 10^{-3} para o NIR de bancada e 10^{-4} para o NIR ultracompacto.

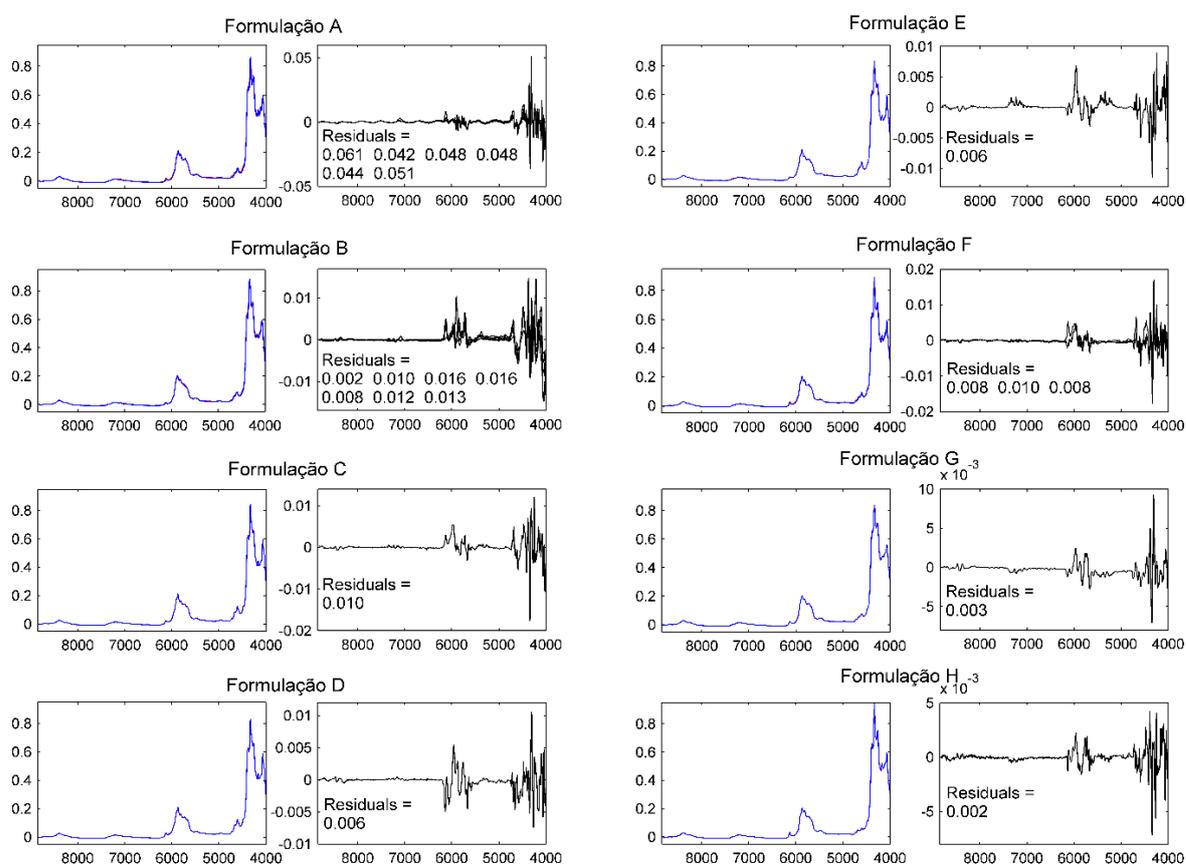


Figura 22. Comparação entre os espectros experimentais (vermelho) e as estimativas geradas pelos modelos CLS (sobrepostas em azul) no espectrômetro de bancada.

No caso dos espectros do NIR de bancada, as estimativas dos modelos CLS para as proporções das correntes de hidrocarbonetos foram bastante semelhantes aos valores reais no caso das gasolinas que foram formuladas com apenas três correntes. No caso em que foram utilizadas quatro correntes na formulação, as estimativas foram satisfatórias para apenas uma das correntes; para as outras 3 os valores estimados apresentaram uma discrepância alta em relação aos reais. Em relação aos modelos CLS realizados com os espectros do NIR ultracompacto, para apenas duas das correntes de hidrocarbonetos, as proporções estimadas foram semelhantes aos valores reais. Para as outras duas correntes e, principalmente no caso em que foram utilizadas quatro correntes na formulação, as estimativas foram bastante imprecisas. Tanto as estimativas como os valores reais para as proporções das correntes utilizadas na produção da gasolina formulada não estão explícitos aqui devido a questões de confidencialidade.

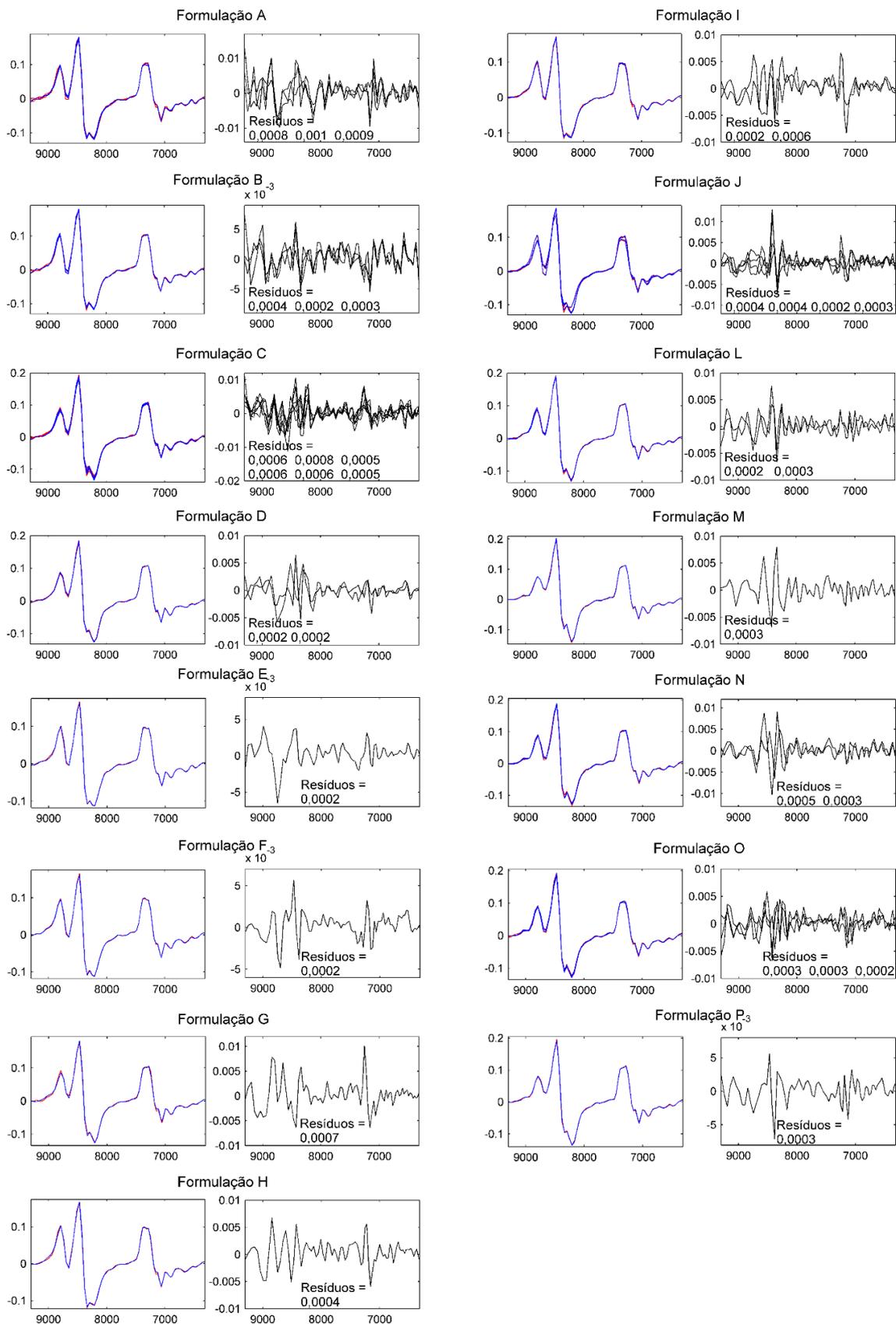


Figura 23. Comparação entre os espectros experimentais (vermelho) e as estimativas geradas pelos modelos CLS (sobrepostas em azul) no espectrômetro ultracompacto.

De maneira geral, os modelos CLS apresentaram resultados promissores, no entanto, podem ser melhorados com a utilização de métodos como o *Multivariate Curve Resolution - Alternate Least Squares* (MCR-ALS), em que é possível otimizar as estimativas espectrais e das proporções, assim como a utilização de restrições para evitar que o modelo forneça valores negativos para as proporções das correntes de hidrocarbonetos. Porém, no decorrer do trabalho uma estratégia mais adequada para atender aos requisitos do simulador foi construída, como será discutido.

4.2.4 Desenvolvimento dos modelos PLS para prever os valores de octanagem, MON e RON.

Durante o desenvolvimento dos modelos PLS foram encontrados alguns desafios em relação à faixa de octanagem das amostras. Mesmo durante os dois anos de coleta de dados, os valores de octanagem das gasolinas analisadas possuem uma diferença de aproximadamente 2 unidades, para MON, e 5 unidades para RON. Com essa variabilidade já foi possível ajustar modelos PLS para a gasolina Tipo 2, que é destinada a comercialização sem adição de etanol, portanto, os valores de octanagem de referência para essas amostras foi determinado na gasolina sem adição de etanol. A gasolina Tipo 1, entretanto, é comercializada com a adição de etanol. Nesse caso, os espectros foram adquiridos durante a produção (antes da adição de etanol) e a análise de referência para octanagem é feita após adição de etanol. A faixa de octanagem das gasolinas analisadas nos dois anos após o início do projeto não foi suficientemente ampla para desenvolver modelos PLS adequados no caso da gasolina Tipo 1.

A estratégia utilizada para contornar este problema foi realizar uma atualização do modelo PLS para gasolinas Tipo 1 utilizando um banco de dados com espectros e dados de referência de aproximadamente 800 amostras que foi fornecido pela empresa. A seguir serão apresentados os resultados obtidos para os modelos PLS para as gasolinas Tipo 2 com apenas os espectros adquiridos durante o projeto, e para gasolinas Tipo 1 após a atualização do modelo. Como os dados históricos foram adquiridos apenas no espectrômetro NIR de bancada, não foi possível fazer a atualização do modelo para o NIR ultracompacto.

4.2.4.1 MODELO PLS (GASOLINA DO TIPO 1)

Com a atualização do modelo PLS para gasolinas Tipo 1, a faixa dos valores de referência para octanagem passou a ser de aproximadamente 7 unidades para MON e 10 unidades para RON. Os espectros pré-processados (de acordo com o que foi descrito anteriormente), que foram adquiridos do histórico de dados da empresa estão apresentados na Figura 24.

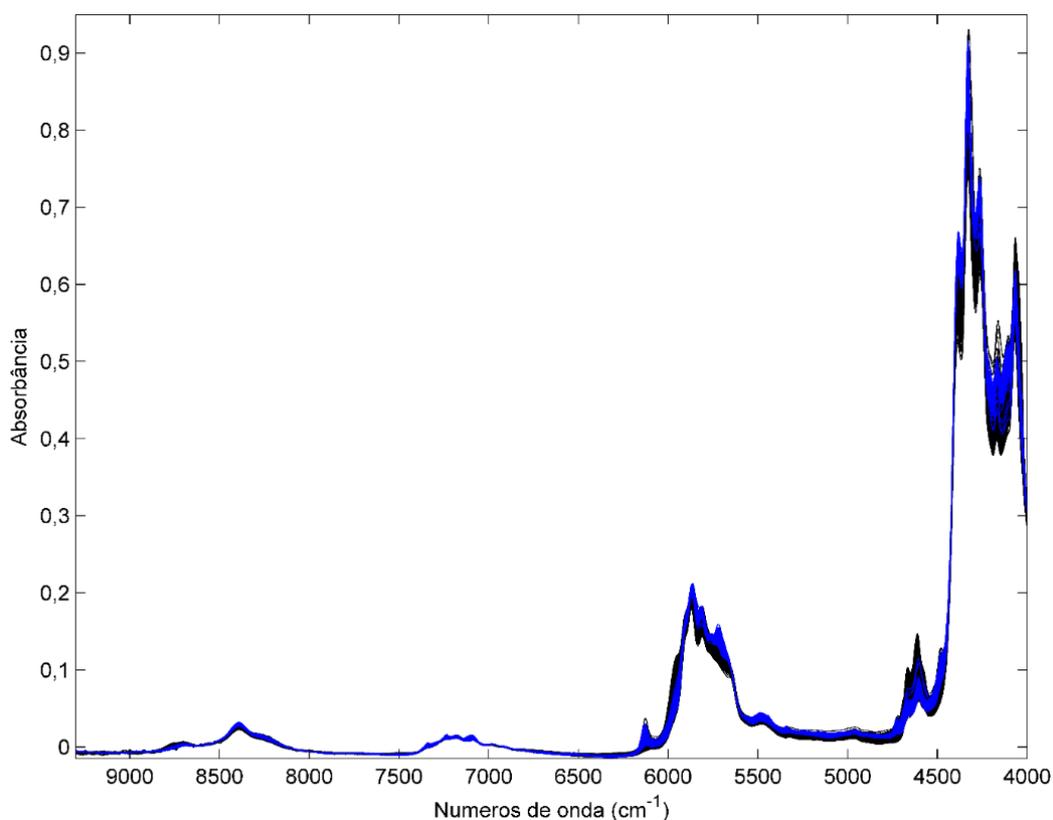


Figura 24. Espectros do NIR de bancada de gasolinas Tipo 1 adquiridos do histórico de dados da companhia (preto) e coletados no período de 2015 a 2017 (azul).

A Tabela 5 compara as figuras de mérito para os modelos PLS da octanagem em gasolinas do Tipo 1, antes e após a atualização. É possível verificar que, na etapa de validação cruzada, o R^2 antes da atualização para os modelos PLS para MON e RON é muito baixo. As Figuras 25a, 26a e 27a mostram os gráficos dos valores preditos vs. referência para as amostras do conjunto de calibração dos modelos PLS após atualização. As figuras mostram que as amostras analisadas após o início do projeto (em azul) possuem uma faixa

de octanagem correspondente apenas a cerca de 40% da faixa de octanagem do histórico da empresa.

Tabela 5. Comparação entre os modelos PLS para prever octanagem em gasolinas do Tipo 1, antes e depois da atualização.

		Validação cruzada					Validação externa			
		n° amostras	VL	R ²	RMSECV	<i>bias</i>	n° amostras	R ²	RMSEP	<i>bias</i>
Antes da atualização	MON	171	3	0,14	0,48	-0,00093	-	-	-	-
	RON	171	2	0,16	0,73	-0,0010	-	-	-	-
	IAD	167	2	0,11	0,33	0,0033	-	-	-	-
Após a atualização	MON	568	6	0,81	0,66	-0,0015	256	0,83	0,62	0,014
	RON	567	4	0,85	0,70	0,0021	194	0,85	0,78	0,11
	IAD	568	6	0,88	0,51	3,8x10 ⁻⁶	256	0,86	0,56	0,55

* VL = Variáveis Latentes

Após a atualização, os modelos PLS foram capazes de prever os valores de octanagem em gasolinas do Tipo 1 com coeficientes de determinação acima de 0,80 e erros de previsão de no máximo 0,8% em relação às faixas de octanagem modeladas. Com exceção do *bias* para a etapa de validação externa do modelo PLS para IAD, com base em um teste t (de acordo com o recomendado pela norma ASTM E1655-05 [62]), nenhum dos valores de *bias* difere significativamente de zero, em um nível de confiança de 95%. A previsão das amostras de validação externa é apresentada nas Figuras 25b, 26b e 27b, para os modelos PLS dos parâmetros MON e RON e IAD, respectivamente. Os gráficos de T² de Hotelling vs. Q Residual (Figuras 25c, 26c e 27c) não detectaram nenhuma amostra com comportamento anômalo para esses modelos. As variáveis mais importantes nos modelos desenvolvidos estão mostradas nos gráficos das Figuras 25d, 26d e 27d. Como já foi observado durante a análise exploratória dos espectros do NIR de bancada, as regiões espectrais de primeiro sobretom de ligações C-H (~6000 cm⁻¹) e de combinações de ligações C-H (5800-4000 cm⁻¹) são as mais importantes para esses modelos. Com isso, a atualização atingiu uma confiabilidade bastante razoável para que seja possível provar o funcionamento do simulador para amostras de gasolina do Tipo 1. Porém, esta atualização foi realizada utilizando espectros adquiridos ao longo de 10 anos, em condições que não são exatamente as mesmas em que as amostras atuais foram analisadas. Portanto, para poder atribuir a confiabilidade adequada a nível de implementação em processo, é necessário que o modelo seja alimentado com amostras atuais que possuam uma faixa de octanagem mais ampla.

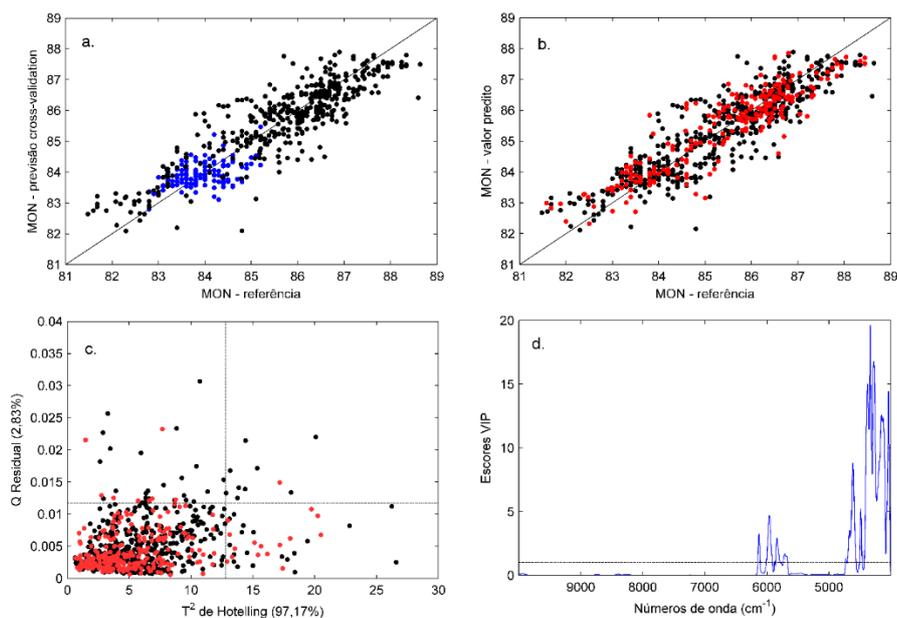


Figura 25. Modelo de regressão PLS para o parâmetro de octanagem MON, após atualização do modelo. (a.) Conjunto de calibração atual (azul) e do histórico (preto). (b.) Conjunto de calibração (preto) e previsão (vermelho) após atualização do modelo. (c.) Gráfico dos resíduos para amostras de calibração (preto) e previsão (vermelho). (d.) Gráfico dos escores para a importância das variáveis na projeção.

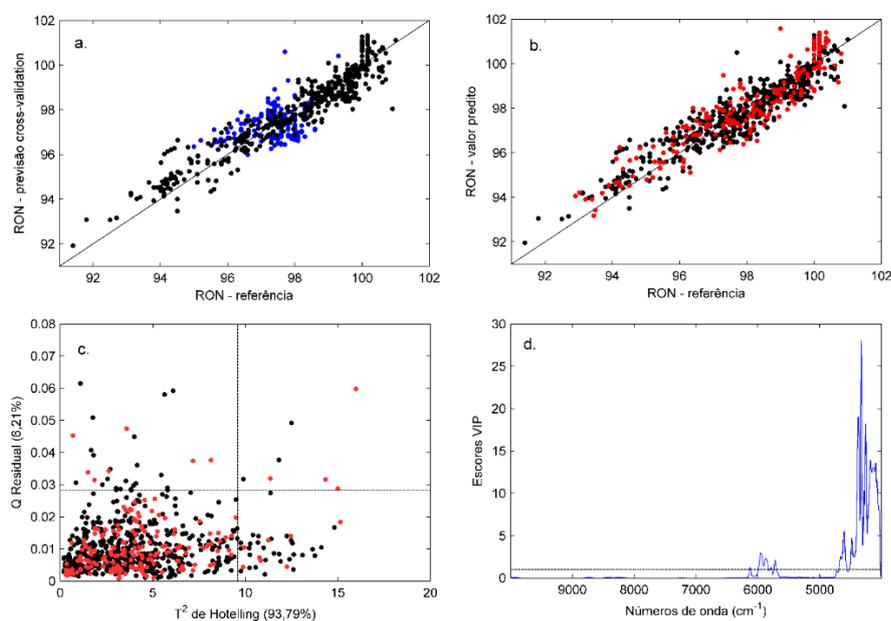


Figura 26. Modelo de regressão PLS para o parâmetro de octanagem RON, após atualização do modelo. (a.) Conjunto de calibração atual (azul) e do histórico (preto). (b.) Conjunto de calibração (preto) e previsão (vermelho) após atualização do modelo. (c.) Gráfico dos resíduos para amostras de calibração (preto) e previsão (vermelho). (d.) Gráfico dos escores para a importância das variáveis na projeção.

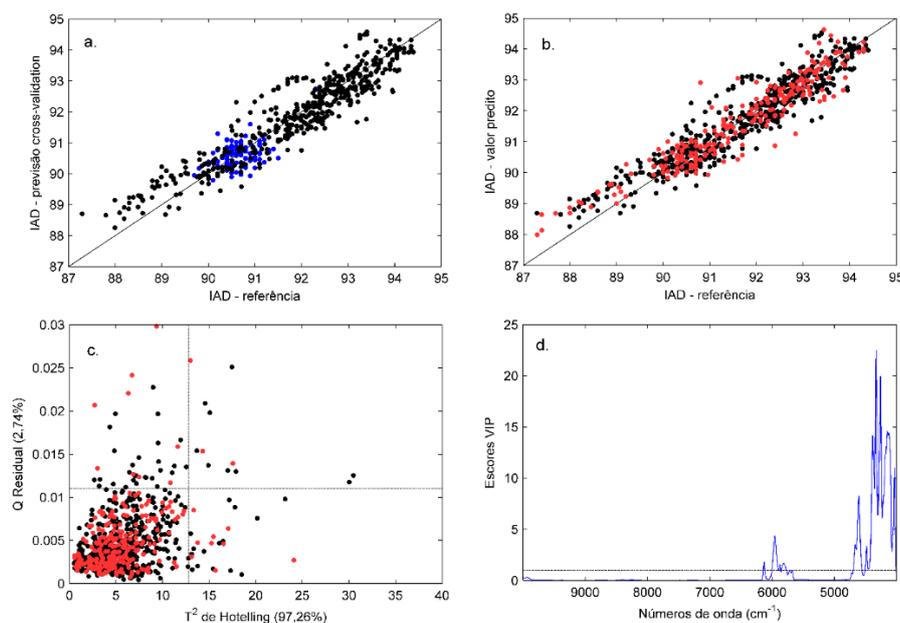


Figura 27. Modelo de regressão PLS para o parâmetro de octanagem IAD, após atualização do modelo. (a.) Conjunto de calibração atual (azul) e do histórico (preto). (b.) Conjunto de calibração (preto) e previsão (vermelho) após atualização do modelo. (c.) Gráfico dos resíduos para amostras de calibração (preto) e previsão (vermelho). (d.) Gráfico dos escores para a importância das variáveis na projeção.

4.2.4.2 MODELO PLS (GASOLINA DO TIPO 2)

Embora tenha sido produzida e analisada uma quantidade relativamente pequena de amostras de gasolinas do Tipo 2 (Tabela 6), os desempenhos dos modelos PLS para octanagem foram satisfatórios, demonstrando elevada capacidade preditiva.

Tabela 6. Comparação entre os modelos PLS para prever parâmetros de octanagem em gasolinas do Tipo 2 no NIR de bancada e ultracompacto.

		Validação cruzada					Validação externa			
		n° amostras	VL	R ²	RMSECV	bias	n° amostras	R ²	RMSEP	bias
Bancada	MON	32	3	0,89	0,20	0,0020	16	0,69	0,36	0,15
	RON	34	3	0,93	0,35	-0,0063	16	0,89	0,43	0,073
	IAD	33	3	0,89	0,33	-0,017	16	0,92	0,36	0,20
Ultra-compacto	MON	36	2	0,59	0,40	0,0011	18	0,64	0,34	0,037
	RON	35	3	0,60	0,90	0,0023	18	0,52	0,77	-0,087
	IAD	35	4	0,39	0,69	0,021	18	0,66	0,40	-0,044

Os coeficientes de determinação atingiram 0,9 nos modelos PLS desenvolvidos no NIR de bancada, porém se mantiveram em torno de 0,6 nos modelos do NIR ultracompacto. Os modelos desenvolvidos no NIR ultracompacto também apresentaram erros médios de previsão um pouco mais elevados em relação ao de bancada, mas ainda assim, esses valores representam em média 0,5% em relação a faixa de octanagem dos modelos PLS. Os modelos para MON, em ambos os instrumentos, e RON para o NIR de bancada, apresentam valores de RMSEP menores do que a reprodutibilidade dos métodos de referência para MON (2,00) e RON (0,70) (ASTM D2699 e D2700). O RMSEP para RON, no NIR ultracompacto se manteve em 0,77, um pouco acima da reprodutibilidade. De acordo com os gráficos de T^2 de Hotelling vs. Q Residual mostrados nas Figuras 28a, 28b e 28c, e também 29a, 29b e 29c, os modelos no NIR de bancada e ultracompacto, respectivamente, não apresentam amostras com comportamento anômalo. Os gráficos dos valores de referência vs. preditos para o NIR de bancada, nas Figuras 28d, 28e e 28f, mostram que os valores foram bem ajustados pelo modelo. No caso do ultracompacto (Figuras 29d, 29e e 29f) os valores se mostram mais dispersos, como já mostrado nos parâmetros de desempenho analítico. O comportamento dos gráficos mostrados nas Figuras 28g, 28h e 28i, é semelhante ao dos modelos desenvolvidos para gasolinas do Tipo 1, em que as variáveis mais importantes para os modelos estão, principalmente, na região de combinações de ligações C-H, no caso do NIR de bancada, com a região de primeiro sobretom contribuindo em uma proporção bem menor. Nos modelos do NIR ultracompacto, as variáveis mais importantes (Figuras 29g, 29h e 29i) se concentram na região de segundo sobretom de ligações C-H, em torno de 8500 cm^{-1} . De acordo com um teste t (ASTM E1655-05 [62]), apenas a validação externa do modelo PLS para IAD desenvolvido no NIR de bancada apresentou um *bias* significativamente diferente de zero, em um nível de confiança de 95%.

Do ponto de vista da utilização desses modelos em um processo industrial, a quantidade de amostras ainda é muito pequena, porém os modelos já apresentaram um alto potencial para prever os parâmetros de octanagem em gasolina. Dessa forma, o aumento da robustez e confiabilidade destes modelos, assim como o aumento na faixa de octanagem podem ser alcançados com a introdução de novas amostras na medida em que forem sendo produzidas novas gasolinas.

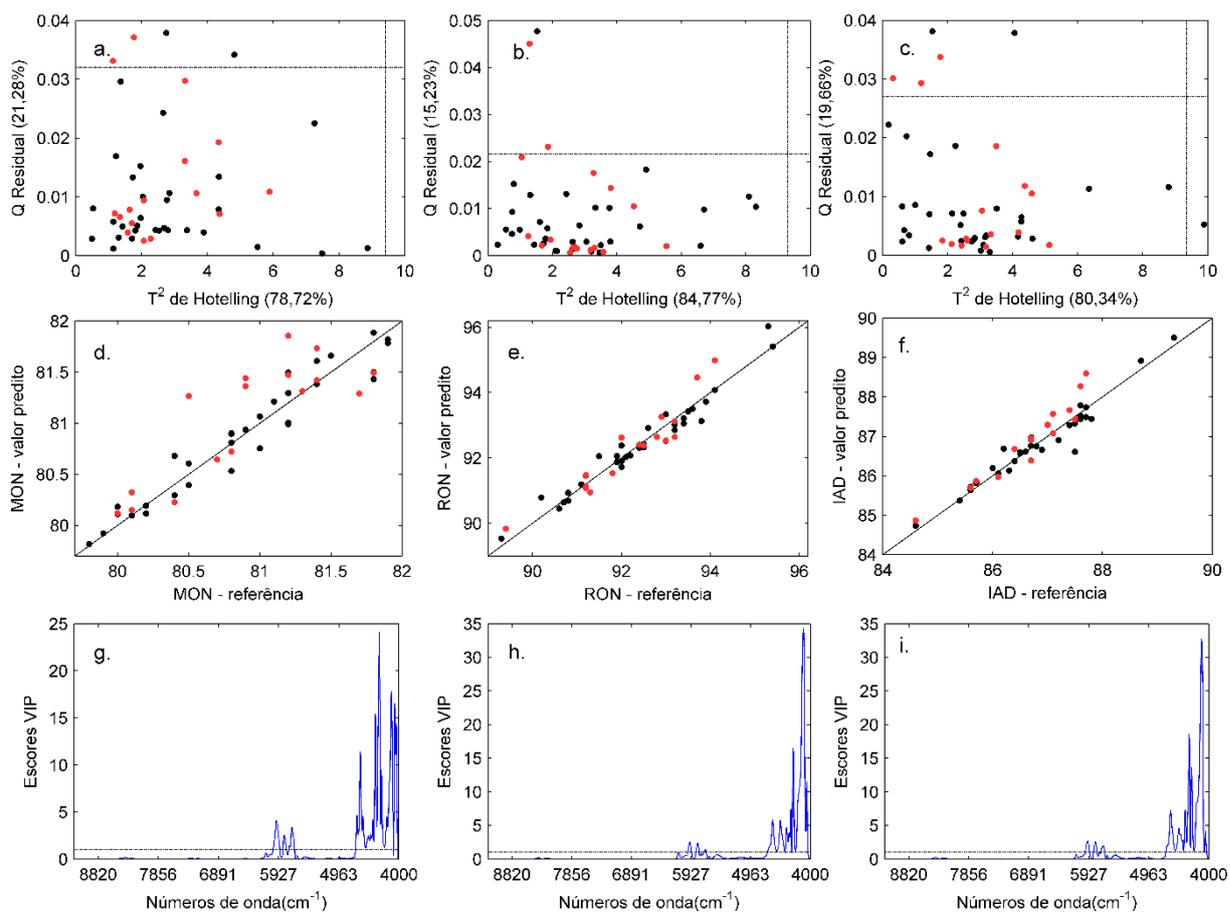


Figura 28. Modelos de regressão PLS desenvolvidos no NIR de bancada para os parâmetros de octanagem MON, RON e IAD. (a., b. e c.) Gráfico dos resíduos; (d., e. e f.) Gráfico dos valores preditos vs. referência; (g., h. e i.) Gráfico dos escores para a importância das variáveis na projeção; para MON, RON e IAD, respectivamente. Amostras de calibração (preto) e previsão (vermelho).

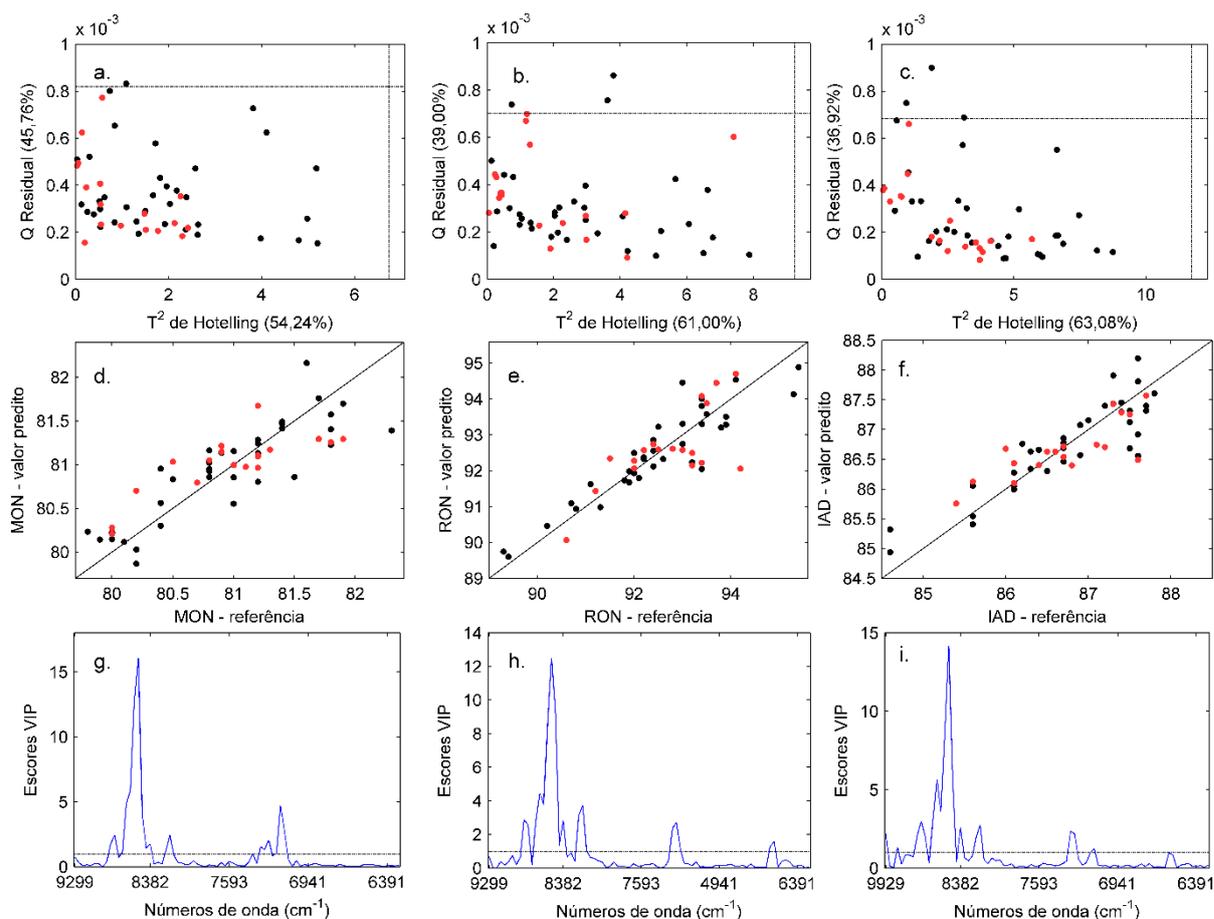


Figura 29. Modelos de regressão PLS desenvolvidos no NIR ultracompacto para parâmetros de octanagem MON (esquerda) e RON (direita). (a., b. e c.) Gráfico dos resíduos; (d., e. e f.) Gráfico dos valores preditos vs. referência; (g., h. e i.) Gráfico dos escores para a importância das variáveis na projeção; para MON, RON e IAD, respectivamente. Amostras de calibração (preto) e previsão (vermelho).

4.2.5 Efeito da variação na proporção das correntes de hidrocarbonetos na octanagem predita pelos modelos PLS

Após o desenvolvimento dos modelos PLS em ambos os espectrômetros, o próximo passo focou na validação desses modelos para prever os valores de octanagem em espectros simulados a partir dos espectros das correntes de hidrocarbonetos. No primeiro momento, foram feitas simulações de espectros de gasolinas do Tipo 1 e do Tipo 2 cujos espectros experimentais também foram obtidos. Como mostra a Tabela 7, os valores de octanagem obtidos a partir dos espectros simulados foram comparados com aqueles obtidos a partir dos espectros experimentais e também com o valor de octanagem do método de referência.

Tabela 7. Comparação entre os valores de octanagem do método de referência e previsões PLS a partir de espectros experimentais e simulados.

		NIR de bancada						NIR ultracompacto			
		Ref.	Exp.	Tipo 1		Tipo 2		Exp.	Tipo 2		
				Sim.	Erro	Exp.	Sim.		Erro	Sim.	Erro
Tempo 1	MON	80,4	-	-	-	-	-	-	80,3	80,3	0,0
	RON	91,8	-	-	-	-	-	-	91,7	91,8	0,1
	IAD	86,1	-	-	-	-	-	-	86,0	85,1	-0,9
Tempo 2	MON	81,4	-	-	-	-	-	-	81,4	81,3	-0,1
	RON	93,9	-	-	-	-	-	-	93,3	94,3	1,0
	IAD	87,7	-	-	-	-	-	-	87,3	87,0	-0,3
Tempo 3	MON	83,6	83,7	84,7	1,0	-	-	-	-	-	-
	RON	97,2	98,5	99,2	0,7	-	-	-	-	-	-
	IAD	90,4	90,9	91,7	0,8	-	-	-	-	-	-
Tempo 4	MON	84,6	84,4	84,6	0,2	-	-	-	-	-	-
	RON	97,9	96,4	96,5	0,1	-	-	-	-	-	-
	IAD	91,3	90,4	90,5	0,1	-	-	-	-	-	-
Tempo 5	MON	81,4	-	-	-	81,7	81,9	0,2	81,5	81,6	0,1
	RON	93,7	-	-	-	94,5	95,0	0,5	94,5	95,0	0,5
	IAD	87,6	-	-	-	88,3	88,5	0,2	87,8	87,9	0,1
Tempo 6	MON	81,2	-	-	-	-	-	-	81,7	81,9	0,2
	RON	94,1	-	-	-	-	-	-	94,7	96,8	2,1
	IAD	87,7	-	-	-	-	-	-	87,6	88,1	0,5
Tempo 7	MON	81,2	-	-	-	81,9	83,3	1,4	-	-	-
	RON	94,1	-	-	-	95,0	98,5	3,5	-	-	-
	IAD	87,7	-	-	-	88,6	90,7	2,1	-	-	-
Tempo 8	MON	81,7	-	-	-	-	-	-	81,7	82,0	0,3
	RON	93,4	-	-	-	-	-	-	93,3	97,4	4,1
	IAD	87,6	-	-	-	-	-	-	88,2	88,5	0,3
Tempo 9	MON	81,2	-	-	-	-	-	-	81,1	81,1	0,0
	RON	93,4	-	-	-	-	-	-	94,1	95,0	0,9
	IAD	87,3	-	-	-	-	-	-	87,4	87,3	-0,1
Tempo 10	MON	81,0	-	-	-	-	-	-	80,9	81,2	0,3
	RON	93,4	-	-	-	-	-	-	94,0	95,0	1,0
	IAD	87,2	-	-	-	-	-	-	87,4	87,3	-0,1
Tempo 11	MON	83,6	83,8	84,1	0,3	-	-	-	-	-	-
	RON	97,4	97,5	98,1	0,6	-	-	-	-	-	-
	IAD	90,5	90,6	91,0	0,4	-	-	-	-	-	-
Tempo 12	MON	84,0	83,6	84,1	0,5	-	-	-	-	-	-
	RON	97,6	98,0	99,3	1,3	-	-	-	-	-	-
	IAD	90,8	90,7	91,5	0,8	-	-	-	-	-	-
Tempo 13	MON	84,3	84,0	84,2	0,2	-	-	-	-	-	-
	RON	96,8	97,1	97,7	0,6	-	-	-	-	-	-
	IAD	90,6	90,6	90,8	0,2	-	-	-	-	-	-
Tempo 14	MON	83,2	83,9	84,0	0,1	-	-	-	-	-	-
	RON	97,6	97,0	97,3	0,3	-	-	-	-	-	-
	IAD	90,4	90,5	90,6	0,1	-	-	-	-	-	-
Tempo 15	MON	83,8	84,1	85,2	1,1	-	-	-	-	-	-
	RON	96,8	97,8	98,1	0,3	-	-	-	-	-	-
	IAD	90,3	90,9	91,6	0,7	-	-	-	-	-	-

Ref.: valor de referência; Exp. e Sim.: valores obtidos com o espectro experimental e simulado, respectivamente.

As previsões a partir dos espectros simulados se comportaram de forma semelhante a dos espectros experimentais, na maioria dos casos. Dessa forma, os desvios em relação aos valores de referência são da mesma ordem dos desvios observados na etapa de validação externa dos modelos PLS mostrados no tópico anterior.

Com os erros mostrados na Tabela 7, foi calculada a raiz do erro médio quadrático entre o valor experimental e simulado para cada parâmetro no NIR de bancada (utilizando os 9 tempos para os Tipos 1 e 2) e no ultracompacto. Para o NIR de bancada, os valores de RMSEP para MON, RON e IAD foram, respectivamente, 0,72, 1,32 e 0,85. No NIR ultracompacto, os valores de RMSEP foram 0,18, 1,86 e 0,43 para MON, RON e IAD, respectivamente. Em ambos os instrumentos, os valores de RMSEP para MON são menores que a reprodutibilidade do método ASTM (2,00). Porém, ocorreram duas situações anômalas na previsão dos valores de RON, no tempo 7 para o NIR de bancada e no tempo 8 para o NIR ultracompacto, em que o erro do valor previsto para o espectro simulado em relação ao espectro experimental foi de 3,5 e 4,1, respectivamente.

Esse erro elevado pode ter ocorrido devido ao fato de que nas simulações para os tempos 7 e 8, em particular, os espectros da corrente de hidrocarbonetos aromáticos, no caso do tempo 7, e pesados, no tempo 8, foram analisados 5 dias após a mudança da formulação. Embora a formulação tenha se mantido a mesma ao longo desses 5 dias, as composições dessas correntes podem ter sofrido mudanças que por sua vez, influenciaram os seus respectivos espectros. Isso pode ter prejudicado a representatividade dos espectros simulados, principalmente, no tempo 7, em que os parâmetros MON e IAD também apresentaram valores um pouco elevados, e no tempo 8. Vale ressaltar que os espectros experimentais foram adquiridos no mesmo dia em que ocorreu mudança na formulação. Então, após eliminar os tempos 7 e 8 no cálculo dos valores de RMSEP para RON na Tabela 7, ocorreu uma diminuição de 1,32 para 0,64, no NIR de bancada. Esse valor se mantém entre os valores de RMSEP para RON dos modelos PLS da gasolina Tipo 1 (0,78) e do Tipo 2 (0,43) e equivalente a reprodutibilidade do método ASTM (0,70). No caso do NIR ultracompacto, o RMSEP para RON na Tabela 7 cai de 1,86 para 1,12, ficando mais próximo do valor de RMSEP para o modelo PLS da gasolina Tipo 2 (0,77).

Ainda em relação ao parâmetro RON, na Tabela 7, a maioria dos valores apresentam um desvio positivo em relação aos valores de referência, indicando que pode haver um *bias* nas previsões. Previsões com *bias*, entretanto, podem ser facilmente corrigidas por meio de técnicas simples de transferência de calibração.

No tempo 5, foi observado que as previsões dos modelos PLS no NIR de bancada e ultracompacto para a mesma gasolina do Tipo 2 foram praticamente idênticas. Isso mostra que ambos os modelos estão processando as informações nos espectros NIR de forma coerente. Com estes resultados foi estabelecido que as simulações espectrais são capazes de fornecer previsões confiáveis tanto em relação aos espectros experimentais como aos valores de referência. Portanto, o procedimento seguinte consistiu em realizar algumas mudanças nas proporções das correntes de hidrocarbonetos dessas simulações para produzir diferentes gasolinas, atribuindo-as variações nos valores de octanagem. Levando em consideração o conhecimento sobre as características da composição das correntes e o seu efeito na octanagem, as variações causadas nos valores de octanagem devem manter uma coerência. Esse conhecimento, por sua vez, é útil para interpretar as previsões dos modelos PLS, uma vez que não há valores de referência nesse caso.

As mudanças realizadas foram conduzidas de forma sistemática e os valores previstos para as novas simulações, assim como a descrição das variações realizadas, são apresentados nas Tabelas 8 e 9. Nas simulações de gasolinas formuladas com 3 correntes de hidrocarbonetos, como nos tempos 3, 4, 5, 12 e 15 na Tabela 8, e os tempos 1, 5, 6 e 8 na Tabela 9, os valores de octanagem variaram de maneira diretamente proporcional com o aumento e diminuição na quantidade de hidrocarbonetos pesados, principalmente. O mesmo comportamento também acontece de maneira menos intensa com os hidrocarbonetos leves, nesse caso de forma inversamente proporcional. No caso das simulações para formulações com 4 correntes de hidrocarbonetos (isto é, nos tempos 7, 11, 13 e 14, na Tabela 8 e nos tempos 2, 9 e 10, na Tabela 9), é interessante notar que nos casos em que foram feitas mudanças de proporção com o aumento ou diminuição de hidrocarbonetos aromáticos sendo diretamente compensados com hidrocarbonetos pesados, as variações nos valores de octanagem foram mínimas ou nenhuma. Mas, por outro lado, foram observadas mudanças significativas quando as alterações na quantidade de hidrocarbonetos aromáticos foram compensadas com hidrocarbonetos leves ou intermediários.

As variações causadas nos valores de octanagem preditos pelos modelos PLS desenvolvidos tanto no NIR de bancada como no ultracompacto mostraram que os hidrocarbonetos pesados, de cadeias longas, assim como os hidrocarbonetos aromáticos exercem uma influência importante nos parâmetros de octanagem da gasolina automotiva. Adicionalmente, é importante ressaltar que algumas mudanças resultaram em valores de octanagem previstos fora da faixa de calibração nos modelos e, nesse caso, as extrapolações

não são confiáveis. Entretanto, esse padrão de comportamento observado ocorreu na grande maioria dos casos mostrados nas Tabelas 8 e 9. Os testes realizados até este ponto são bastante favoráveis para o desenvolvimento do simulador de octanagem para gasolina automotiva que será apresentado no próximo tópico.

Tabela 8. Previsões PLS (NIR de bancada) dos valores de octanagem nos espectros simulados para as variações nas proporções das correntes de hidrocarbonetos.

	MON	RON	IAD	Descrição da mudança realizada na composição
Gasolinas do Tipo 1				
Tempo 3	84,7	99,2	91,7	Proporção original
	83,4	95,2	89,3	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	86,5	104,8	94,9	Aumento de hidrocarbonetos pesados
	83,9	96,5	90,1	Aumento de hidrocarbonetos leves
	85,2	100,8	92,6	Diminuição de hidrocarbonetos leves
Tempo 4	84,6	96,5	90,5	Proporção original
	83,2	92,5	88,0	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	86,7	102,2	94,0	Aumento de hidrocarbonetos pesados
	83,6	94,3	89,1	Aumento de hidrocarbonetos leves
	85,3	98,0	91,5	Diminuição de hidrocarbonetos leves
Tempo 11	84,1	98,1	91,0	Proporção original
	84,2	98,1	91,0	Aumento de hidrocarbonetos aromáticos e diminuição de pesados
	84,1	98,0	90,9	Diminuição de hidrocarbonetos aromáticos e aumento de pesados
	84,9	99,8	92,1	Aumento de hidrocarbonetos aromáticos e diminuição de leves
	83,7	97,2	90,4	Diminuição de hidrocarbonetos aromáticos e aumento de leves
	84,9	100,0	92,2	Aumento de hidrocarbonetos aromáticos e diminuição de intermediários
Tempo 12	83,7	97,1	90,4	Diminuição de hidrocarbonetos aromáticos e aumento de intermediários
	84,1	99,3	91,5	Proporção original
	83,0	95,7	89,4	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	86,0	104,3	94,4	Aumento de hidrocarbonetos pesados
	83,0	97,6	90,3	Aumento de hidrocarbonetos leves
Tempo 13	84,9	100,5	92,3	Diminuição de hidrocarbonetos leves
	84,2	97,7	90,8	Proporção original
	84,2	97,7	90,8	Aumento de hidrocarbonetos aromáticos e diminuição de pesados
	84,2	97,7	90,8	Diminuição de hidrocarbonetos aromáticos e aumento de pesados
	84,8	99,3	91,8	Aumento de hidrocarbonetos aromáticos e diminuição de leves
	83,8	96,4	90,0	Diminuição de hidrocarbonetos aromáticos e aumento de leves
	84,8	99,6	92,0	Aumento de hidrocarbonetos aromáticos e diminuição de intermediários
Tempo 14	83,8	96,1	89,9	Diminuição de hidrocarbonetos aromáticos e aumento de intermediários
	84,0	97,3	90,6	Proporção original
	84,0	97,2	90,6	Aumento de hidrocarbonetos aromáticos e diminuição de pesados
	84,1	97,4	90,7	Diminuição de hidrocarbonetos aromáticos e aumento de pesados
	84,8	98,9	91,6	Aumento de hidrocarbonetos aromáticos e diminuição de leves
	83,3	95,7	89,6	Diminuição de hidrocarbonetos aromáticos e aumento de leves
	84,7	99,4	91,8	Aumento de hidrocarbonetos aromáticos e diminuição de intermediários
Tempo 15	83,4	95,2	89,4	Diminuição de hidrocarbonetos aromáticos e aumento de intermediários
	85,2	98,1	91,6	Proporção original
	83,4	93,6	88,7	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	87,8	104,3	95,6	Aumento de hidrocarbonetos pesados
	83,9	95,9	90,1	Aumento de hidrocarbonetos leves
	86,0	99,6	92,6	Diminuição de hidrocarbonetos leves

Tabela 8. Continuação

	MON	RON	IAD	Descrição da mudança realizada na composição
Gasolinas do Tipo 2				
Tempo 5	81,4	93,7	87,6	Proporção original
	80,1	88,7	84,7	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	83,8	102,1	92,7	Aumento de hidrocarbonetos pesados
	82,1	94,7	88,7	Aumento de hidrocarbonetos leves
	81,7	95,1	88,4	Diminuição de hidrocarbonetos leves
Tempo 7	81,2	94,1	87,7	Proporção original
	83,9	99,2	91,1	Aumento de hidrocarbonetos aromáticos e diminuição de pesados
	82,7	97,8	90,2	Diminuição de hidrocarbonetos aromáticos e aumento de pesados
	84,2	100,1	91,6	Aumento de hidrocarbonetos aromáticos e diminuição de leves
	82,4	96,9	89,7	Diminuição de hidrocarbonetos aromáticos e aumento de leves
	84,8	102,3	93,0	Aumento de hidrocarbonetos aromáticos e diminuição de intermediários
	81,8	94,8	88,4	Diminuição de hidrocarbonetos aromáticos e aumento de intermediários

Tabela 9. Previsões PLS (NIR ultracompacto) dos valores de octanagem nos espectros simulados para as variações nas proporções das correntes de hidrocarbonetos.

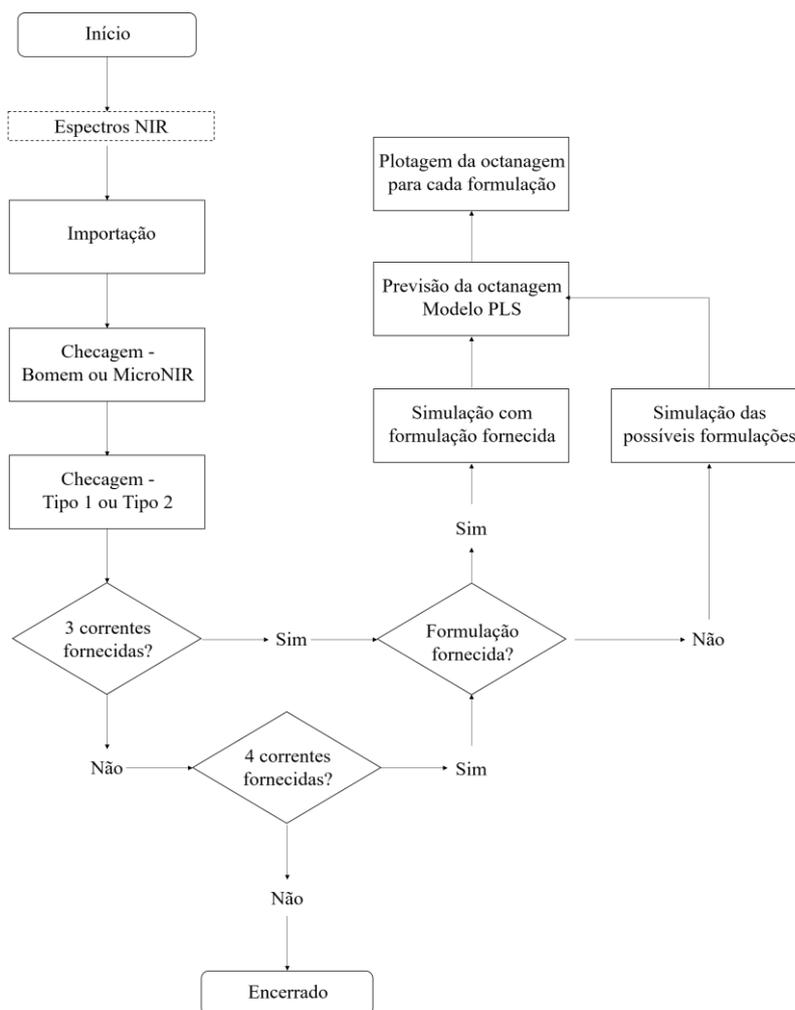
	MON	RON	IAD	Descrição da mudança realizada na composição
Gasolinas do Tipo 2				
Tempo 1	80,3	91,8	85,1	Proporção original
	79,8	87,2	83,7	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	80,3	95,4	86,1	Aumento de hidrocarbonetos pesados
	81,0	91,9	85,8	Aumento de hidrocarbonetos leves
	79,9	91,7	84,8	Diminuição de hidrocarbonetos leves
Tempo 2	81,3	94,3	87,0	Proporção original
	81,6	95,0	87,5	Aumento de hidrocarbonetos aromáticos e diminuição de pesados
	80,9	93,4	86,4	Diminuição de hidrocarbonetos aromáticos e aumento de pesados
	81,5	95,6	87,6	Aumento de hidrocarbonetos aromáticos e diminuição de leves
	81,2	93,1	86,5	Diminuição de hidrocarbonetos aromáticos e aumento de leves
	81,8	97,2	88,3	Aumento de hidrocarbonetos aromáticos e diminuição de intermediários
	80,8	91,4	85,8	Diminuição de hidrocarbonetos aromáticos e aumento de intermediários
Tempo 5	81,6	95,0	87,9	Proporção original
	80,4	87,8	85,1	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	82,3	100,0	89,9	Aumento de hidrocarbonetos pesados
	82,0	96,0	88,3	Aumento de hidrocarbonetos leves
	81,3	94,4	87,7	Diminuição de hidrocarbonetos leves
Tempo 6	81,9	96,8	88,1	Proporção original
	80,6	89,5	85,0	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	82,6	101,0	90,0	Aumento de hidrocarbonetos pesados
	82,4	99,3	88,8	Aumento de hidrocarbonetos leves
	81,6	95,1	87,6	Diminuição de hidrocarbonetos leves
Tempo 8	82,0	97,4	88,5	Proporção original
	80,6	89,3	85,3	Diminuição de hidrocarbonetos pesados e aumento de intermediários
	82,6	101,0	90,3	Aumento de hidrocarbonetos pesados
	82,5	100,4	89,1	Aumento de hidrocarbonetos leves
	81,5	94,6	87,9	Diminuição de hidrocarbonetos leves
Tempo 9	81,1	95,0	87,3	Proporção original
	81,1	95,0	87,2	Aumento de hidrocarbonetos aromáticos e diminuição de pesados
	81,1	94,9	87,3	Diminuição de hidrocarbonetos aromáticos e aumento de pesados
	81,0	95,3	87,3	Aumento de hidrocarbonetos aromáticos e diminuição de leves
	81,2	94,7	87,2	Diminuição de hidrocarbonetos aromáticos e aumento de leves
	81,4	97,3	88,1	Aumento de hidrocarbonetos aromáticos e diminuição de intermediários
80,8	92,6	86,4	Diminuição de hidrocarbonetos aromáticos e aumento de intermediários	

Tabela 9. Continuação

	MON	RON	IAD	Descrição da mudança realizada na composição
Tempo 10	81,2	95,0	87,3	Proporção original
	81,2	95,1	87,4	Aumento de hidrocarbonetos aromáticos e diminuição de pesados
	81,1	94,8	87,3	Diminuição de hidrocarbonetos aromáticos e aumento de pesados
	81,1	95,4	87,4	Aumento de hidrocarbonetos aromáticos e diminuição de leves
	81,2	94,5	87,2	Diminuição de hidrocarbonetos aromáticos e aumento de leves
	81,5	97,4	88,3	Aumento de hidrocarbonetos aromáticos e diminuição de intermediários
	80,9	92,5	86,4	Diminuição de hidrocarbonetos aromáticos e aumento de intermediários

4.2.6 Desenvolvimento do simulador de octanagem para gasolina automotiva

Após o conhecimento adquirido através das etapas de investigação dos dados, atualização do modelo PLS para gasolinas do Tipo 1 e, em seguida, validação dos modelos PLS, descritas nos tópicos anteriores, foi montada uma estratégia o mais simples possível para o funcionamento do simulador de octanagem NIRSIM, como mostra o diagrama de fluxo na Figura 30.

**Figura 30.** Diagrama de fluxo para o funcionamento do simulador de octanagem NIRSIM.

Todos os procedimentos seguidos pelo simulador, desde a importação dos espectros até a previsão da octanagem para as simulações realizadas, foram desenvolvidas em ambiente Matlab. Dessa forma, para facilitar a utilização dessa ferramenta, uma interface gráfica foi criada. A rotina escrita para a função “simugas” que é responsável pelos cálculos das simulações na interface pode ser verificada no Apêndice I. A interface está mostrada na Figura 31 e os passos para realizar uma simulação de octanagem para gasolina automotiva, descritos na Figura 30, serão apresentados em seguida.

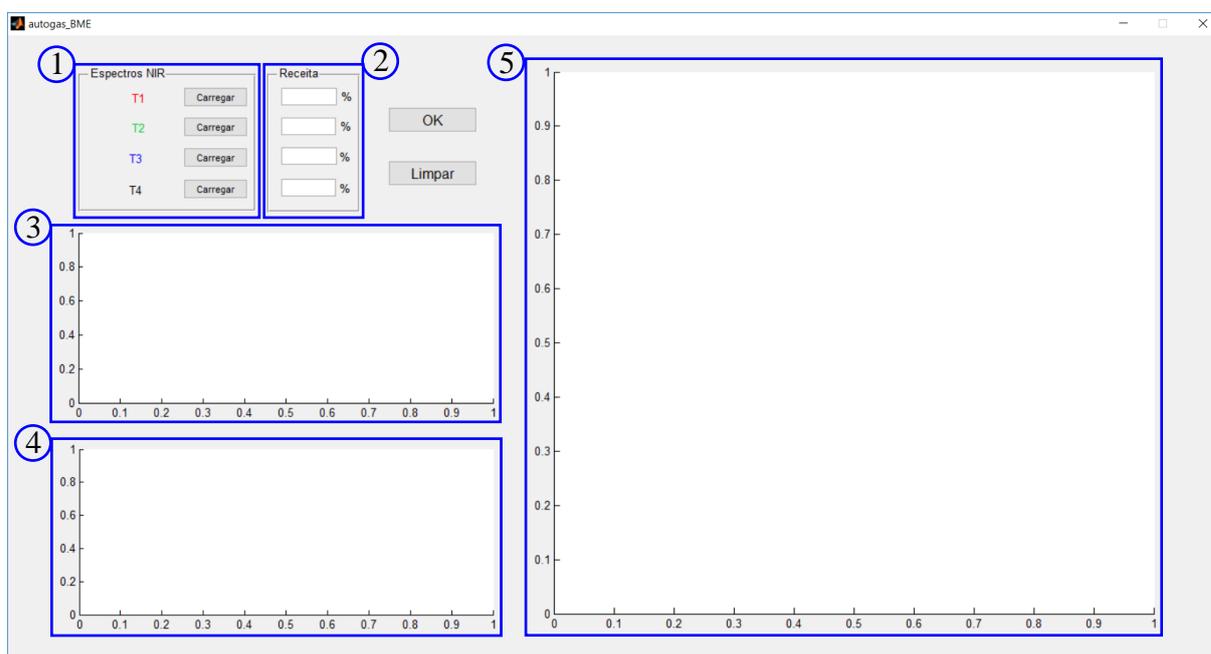


Figura 31. Interface gráfica do simulador de octanagem NIRSIM para gasolina automotiva.

Os dados de entrada para o simulador são fornecidos pelas plataformas mostradas nas janelas 1 e 2 da Figura 31. Na janela 1, os espectros das correntes de hidrocarbonetos são carregados e simultaneamente mostrados na janela 3, sendo que o operador pode escolher entre utilizar os 4 espectros/correntes ou apenas 3 deles. Enquanto que a janela 2 é uma janela opcional. Nesse caso, se a proporção entre as correntes for fornecida, o programa irá calcular o valor da octanagem apenas para a gasolina simulada com a proporção fornecida. Caso contrário, serão calculadas todas as possíveis formulações dentro das faixas de proporção observada no histórico da empresa para cada corrente de hidrocarbonetos. Os espectros simulados são, então, mostrados na janela 4. Esse procedimento torna possível a exploração de novas formulações que, porventura, possam apresentar um maior potencial em relação às formulações tradicionalmente utilizadas. Esse é um fator extremamente

importante uma vez que a composição das correntes de hidrocarbonetos não é constante e, portanto, a formulação ideal para fornecer a gasolina com a octanagem desejada não deve ser sempre a mesma.

Por fim, os valores de octanagem (MON vs. RON) para esses espectros são previstos pelos modelos PLS previamente validados, como mostraram os tópicos anteriores, e incorporados à interface gráfica e, em seguida, são apresentados na janela 5. A seguir serão mostradas algumas simulações utilizando a interface para 3 diferentes tempos da Tabela 7, dois deles para gasolinas do Tipo 1 e Tipo 2, com espectros de correntes adquiridos no NIR de bancada, e do Tipo 2, para espectros adquiridos no NIR ultracompacto.

Na Figura 32 foram utilizados os espectros do Tempo 3 (Tabela 7) para realizar uma simulação de octanagem de gasolinas automotivas do Tipo 1 utilizando espectros de 3 correntes de hidrocarbonetos adquiridos do espectrometro de bancada. Este é um caso em que a proporção entre as correntes não é fornecida.

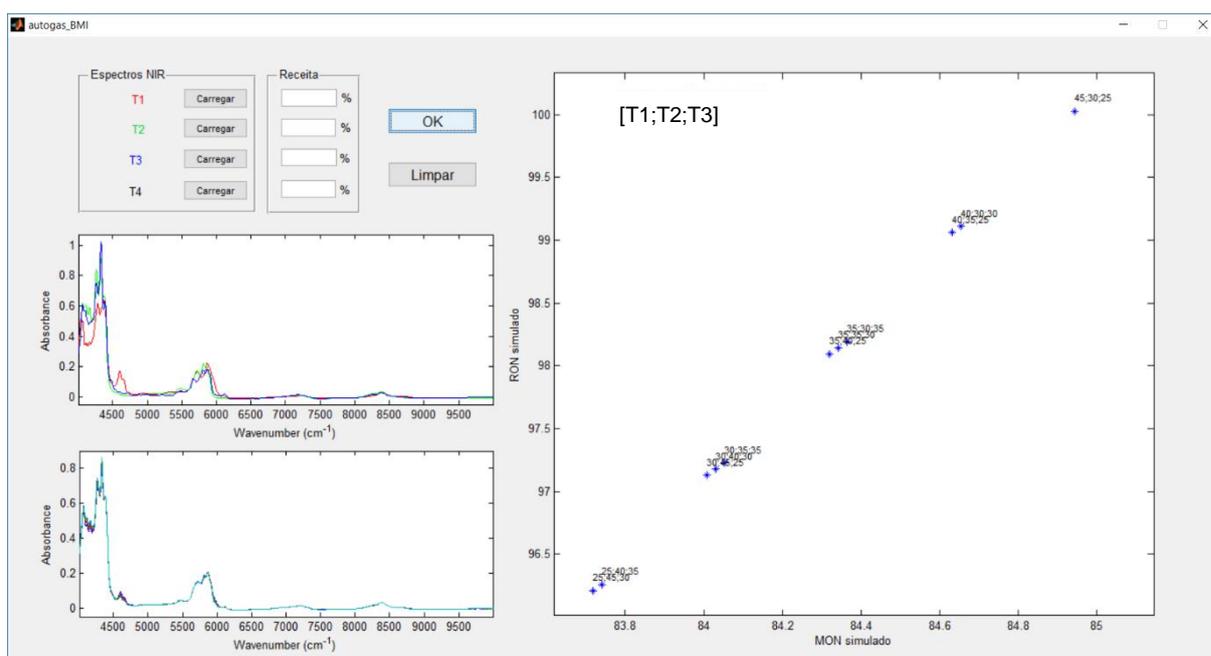


Figura 32. Simulação de octanagem para gasolinas do Tipo 1 utilizando espectros de correntes de hidrocarbonetos adquiridos no NIR de bancada.

No gráfico de MON vs. RON, os valores para estes parâmetros são apresentados para cada uma das 11 espectros simulados calculados e cada ponto é identificado com a sua respectiva formulação.

Na Figura 33, os espectros das correntes de hidrocarbonetos analisados com o NIR de bancada no Tempo 5 (Tabela 7), e a formulação original para a gasolina do Tipo 2 produzida nesse tempo são os dados de entrada para o simulador. Os valores simulados na interface correspondem, portanto, aos valores simulados apresentados na Tabela 7 para os parâmetros MON (81,9) e RON (95,0) e cujos respectivos valores de referência são 81,4 e 93,7.

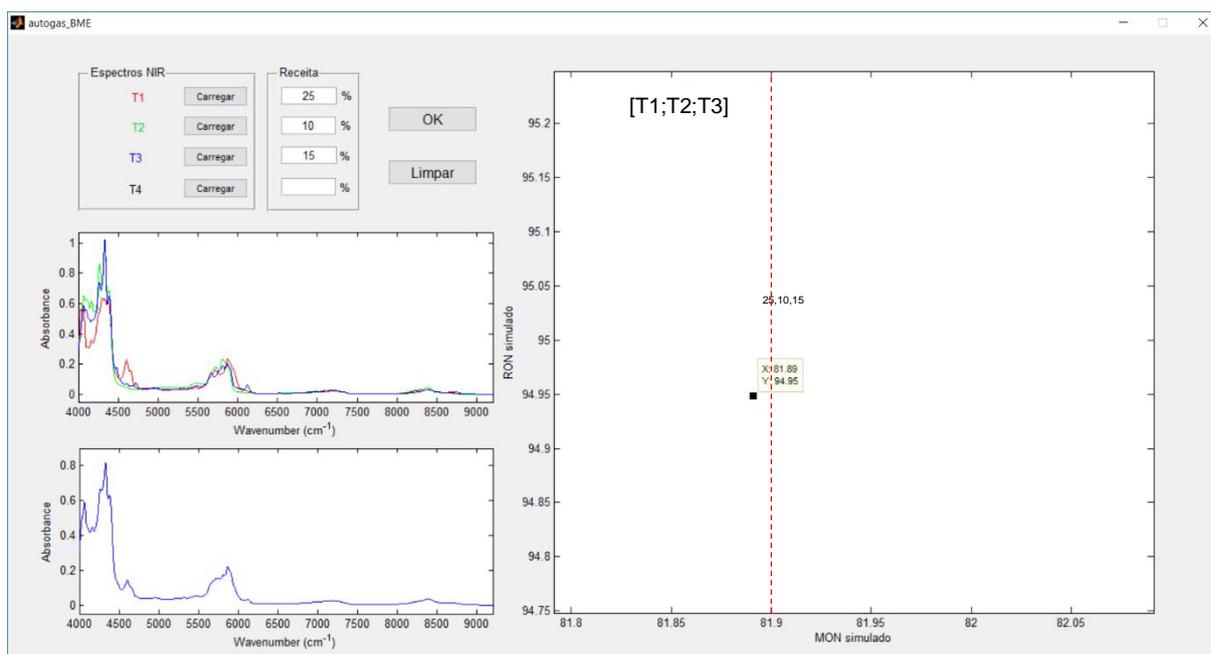


Figura 33. Simulação de octanagem para gasolinas do Tipo 2 utilizando espectros de correntes de hidrocarbonetos adquiridos no NIR de bancada. (--- Limite máximo para a faixa de calibração do modelo PLS para o parâmetro de octanagem MON).

A terceira simulação foi conduzida utilizando os espectros das 4 correntes de hidrocarbonetos adquiridos no NIR ultracompacto no Tempo 2 (Tabela 7) para simular a octanagem de gasolinas automotivas do Tipo 2. O resultado dessa simulação está mostrado na Figura 34.

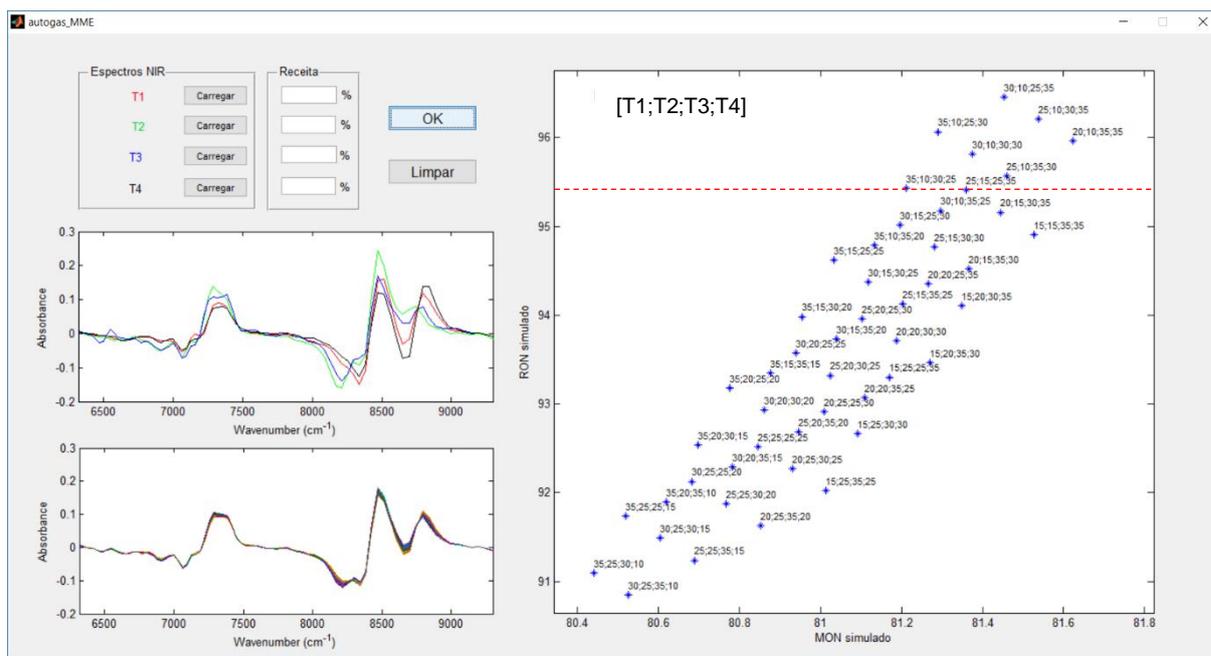


Figura 34. Simulação de octanagem para gasolinas do Tipo 2 utilizando espectros de correntes de hidrocarbonetos adquiridos no NIR ultracompacto. (--- Limite máximo para a faixa de calibração do modelo PLS para o parâmetro de octanagem RON).

Como foram utilizadas 4 correntes, o número de espectros simulados é maior nesse caso, 48 espectros. Observa-se no gráfico de MON vs. RON que alguns espectros possuem valores para o parâmetro RON fora da faixa de calibração do modelo PLS. Uma vez que são calculadas diferentes combinações com as proporções das correntes de hidrocarbonetos, é natural que alguns dos espectros apresentem um comportamento mais distinto ocasionando a previsão de valores fora da faixa de calibração dos modelos. Adicionalmente, é importante observar que os modelos PLS para gasolinas do Tipo 2, que inclui os modelos utilizados nas simulações das Figuras 33 e 34, foram desenvolvidos com amostras de gasolinas recentes. Como já foi discutido, a faixa de octanagem desses modelos ainda é bastante estreita comparando com os valores históricos. Portanto, para evitar este comportamento, a calibração desses modelos deve ser constantemente atualizada com a introdução de amostras recentes ao longo da produção. Portanto, quanto a esses valores, eles devem ser desprezados pois não apresentam a confiabilidade adequada.

O estudo mostrou que os espectros NIR, em ambos os instrumentos de bancada e ultracompacto, são dados de entrada adequados para o simulador off-line de octanagem para gasolina automotiva desenvolvido neste trabalho. Isso torna o simulador proposto uma ferramenta com alto potencial para implementação no processo de formulação da gasolina.

5 CONCLUSÕES

5.1 Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais

Este trabalho avaliou o desempenho de um espectrômetro NIR ultracompacto para determinar parâmetros de qualidade de amostras de gasolina e misturas diesel/biodiesel e também demonstrou a viabilidade de uma transferência de modelos de calibração entre um instrumento de bancada de alta resolução e um instrumento ultracompacto utilizando duas abordagens, uma com padrões de transferência experimentais e outra com padrões matematicamente construídos (padrões virtuais). Para a gasolina, os modelos PLS desenvolvidos no instrumento de bancada e no ultracompacto atingiram valores de RMSEP menores do que a reprodutibilidade dos métodos de referência, para os parâmetros de qualidade estudados. Para as misturas diesel/biodiesel, no entanto, não foi possível fazer essa comparação devido às limitações encontradas no método de referência.

A padronização reversa foi utilizada para transferir conjuntos de calibração de um espectrômetro de alta resolução para um ultracompacto utilizando padrões experimentais ou padrões virtuais obtidos a partir de espectros dos componentes puros majoritariamente presentes na composição dos combustíveis. Os parâmetros densidade, T90 e ponto final de ebulição (PFE) mostraram valores de RMSEP semelhantes para as duas abordagens do método de transferência de calibração, utilizando padrões experimentais ou virtuais. Para os outros parâmetros (T10, T50 e %Bio) os valores de RMSEP foram significativamente diferentes entre essas duas abordagens, sendo que a padronização com padrões experimentais apresentou valores de RMSEP menores. Os modelos PLS construídos após a transferência de calibração com padrões virtuais apresentaram um *bias* estatisticamente significativo para três parâmetros (T10, T50 e densidade). Além disso, diferenças estatisticamente significativas foram observadas para os valores de *bias* de T10 e T50 comparando-os com os obtidos após padronização com padrões experimentais. Contudo, após realizar uma correção de *slope* e *bias* bastante simples nos valores preditos, os valores de *bias* corrigidos para esses três parâmetros não apresentaram significância estatística. Para os parâmetros T10 e T50, os valores de *bias* corrigidos não apresentaram diferenças significativas quando comparados com a padronização utilizando padrões experimentais.

Outro aspecto bastante importante é que a transferência de calibração de modelos PLS de um instrumento de alta resolução para um ultracompacto é vantajosa, seja utilizando padrões de transferência experimentais ou padrões matematicamente construídos a partir de espectros dos componentes puros. A principal vantagem de utilizar os padrões virtuais é que neste caso não é necessário transportar as amostras de transferência nos casos em que os instrumentos se encontram distantes um do outro, o que representa um problema no caso de amostras de combustíveis, devido às leis de segurança envolvidas. Além disso, o uso de padrões virtuais dispensa a necessidade de armazenar as amostras de transferência, evitando problemas com a instabilidade dessas amostras em longos períodos de armazenamento. Portanto, o conjunto de calibração de um instrumento primário pode ser corrigido e transferido para um instrumento secundário, por meio da padronização reversa, utilizando apenas os espectros de componentes puros adquiridos comercialmente.

5.2 Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR

Durante os testes preliminares realizados no NIR ultracompacto se observou que apesar da alta volatilidade nas correntes de hidrocarbonetos, a pequena variabilidade presente nos espectros das 6 repetições das gasolinas formuladas em laboratório demonstravam que essas misturas físicas poderiam ser reproduzidas. Além disso, também foram obtidos resultados promissores na comparação entre espectros experimentais de gasolinas formuladas em laboratório e os espectros simulados. Com isso, os espectros NIR já se mostravam adequados no início dos testes. Na etapa de exploração dos dados se verificou que, de fato, existem diferenças significativas entre os espectros das gasolinas do Tipo 1 e do Tipo 2, apesar de as duas faixas de octanagem serem próximas.

A atualização do modelo PLS para gasolinas do Tipo 1 utilizando espectros do histórico de dados da empresa adquiridos no espectrômetro de bancada conseguiu atingir uma confiabilidade adequada, mesmo se tratando de espectros de até 10 anos. Mas, a introdução de amostras recentes é indispensável para uma possível implementação na linha de produção. Os modelos desenvolvidos no NIR de bancada para gasolinas do Tipo 2 apresentaram um desempenho melhor que os modelos desenvolvidos no NIR ultracompacto, porém este último ainda apresentou um desempenho satisfatório. Entretanto, apesar do

desempenho adequado, a faixa de calibração para os parâmetros de octanagem ainda é muito pequena, por isso, esses modelos devem ser alimentados constantemente na medida em que são produzidas novas amostras.

A comparação entre as previsões dos valores de octanagem a partir dos espectros simulados, os espectros experimentais e os valores de referência, em diferentes tempos ao longo de 2 anos de produção, mostrou que as simulações são capazes de fornecer valores preditos confiáveis, consequentemente, validando o desempenho dos modelos PLS utilizados posteriormente no simulador de octanagem. Os valores de RMSEP para os erros das previsões dos espectros simulados em relação aos experimentais foram de 0,72 (MON) e 0,64 (RON) no NIR de bancada, e 0,18 (MON) e 1,12 (RON) no NIR ultracompacto. Além disso, as variações nos valores de octanagem causadas pelas mudanças propositais na formulação dos espectros simulados em diferentes tempos ocorreram de forma mais intensa com a proporção de hidrocarbonetos pesados e aromáticos.

Vários procedimentos realizados ao longo do desenvolvimento deste trabalho foram incorporados em uma função “simugas” e implementados na criação de uma interface de uso muito simples. Mais uma vez, em função das pequenas faixas de calibração ainda observadas nos modelos PLS para gasolinas do Tipo 2, algumas dos espectros calculados pelo simulador forneceram valores extrapolados e que não podem ser levados em consideração. Porém, com este estudo foi possível demonstrar que os espectros NIR, adquiridos tanto no NIR de bancada como no ultracompacto, forneceram resultados semelhantes mostrando o potencial para utilização de instrumentos portáteis como mais uma ferramenta de otimização. A comercialização do MicroNIR como ferramenta acoplada em linhas de produção vem se estabelecendo cada vez mais. A portabilidade e robustez do MicroNIR são essenciais para uma possível utilização do NIRSIM como um simulador in-line. Isso atribui ao método desenvolvido um alto potencial para implementação em um processo de formulação de gasolina automotiva.

6 PERSPECTIVAS

6.1 Desenvolvimento de uma metodologia de transferência de calibração entre um instrumento de alta resolução e um ultracompacto utilizando padrões virtuais

- Testar novos parâmetros de qualidade da gasolina e misturas diesel/biodiesel
- Otimizar o número e o tipo dos componentes puros
- Testar modelos PLS em toda a região do MicroNIR, utilizando um percurso ótico menor

6.2 Desenvolvimento de um simulador de octanagem para formulações de gasolina automotiva utilizando espectroscopia NIR

- Melhorar os modelos PLS alimentando-os com amostras atuais
- Implementação do método na linha de produção
- Implementar o NIRSIM em uma interface independente do Matlab

REFERÊNCIAS

- [1] **Process Analytical Technology - Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries**. Edição de Katherine A. Bakeev, 1. ed. Nova Iorque: Blackwell Publishing Ltd, 2005.
- [2] ANUÁRIO ESTATÍSTICO BRASILEIRO DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS: 2016 / **Agência Nacional do Petróleo, Gás Natural e Biocombustíveis**. - Rio de Janeiro: ANP. Disponível em: http://www.anp.gov.br/wwwanp/images/publicacoes/Anuario_Estatistico_ANP_2016.pdf. Acessado em: 20 de janeiro de 2017.
- [3] FERNANDES, H. L.; RAIMUNDO JR, I. M.; PASQUINI, C.; ROHWEDDER, J. J. R. Simultaneous determination of methanol and ethanol in gasoline using NIR spectroscopy: Effect of gasoline composition. **Talanta**, v. 75, p. 807-810, 2008.
- [4] OLIVEIRA, F. S.; TEIXEIRA, L. S. G.; ARAUJO, M. C. U.; KORN, M. Screening analysis to detect adulterations in Brazilian gasoline samples using distillation curves. **Fuel**. v. 83, p. 917-923, 2004.
- [5] PORTARIA MAPA Nº 75, de 05/03/2015. **MAPA - Ministério da Agricultura, Pecuária e Abastecimento**, 06 de mar. 2015. Resolução Nº 1, de 04/03/2015. **CIMA - Conselho Interministerial do Açúcar e do Alcool**, 06 de mar. 2015. Disponível em: <www.anp.gov.br/?dw=74502>. Acessado em: 02 de fevereiro de 2017.
- [6] LIMA, A. S.; ASSIS, C. S.; RALDENES, E.; PEREIRA, J. P. M. L. O Programa de Monitoramento da Qualidade dos Combustíveis – PMQC. **Séries Temáticas ANP**, nº 5. Rio de Janeiro, 2012. Disponível em: <www.anp.gov.br/?dw=63073> Acessado em: 02 de fevereiro de 2017.
- [7] RESOLUÇÃO ANP Nº 40, de 28.10.2013 - Dispõe sobre as especificações das gasolinas de uso automotivo. **ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis**, 25 de out. de 2013. Disponível em: <http://nxt.anp.gov.br/NXT/gateway.dll/leg/resolucoes_anp/2013/outubro/ranp%2040%20-%202013.xml> Acessado em: 02 de fevereiro de 2017.

- [8] BALABIN, R. M.; SAFIEVA, R. Z.; LOMAKINA, E. I. Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. **Chem. Intell. Lab. Syst.** v. 88, p. 183-188, 2007.
- [9] BALABIN, R. M.; SAFIEVA, R. Z.; LOMAKINA, E. I. Wavelet neural network (WNN) approach for calibration model building based on gasoline near infrared (NIR) spectra. **Chem. Intell. Lab. Syst.** v. 93, p. 58-62, 2008.
- [10] HE, K.; QIAN, F.; CHENG, H.; DU, W. A novel adaptive algorithm with near-infrared spectroscopy and its application in online gasoline blending processes. **Chem. Intell. Lab. Syst.** v. 140, p. 117-125, 2015.
- [11] LEE, S.; CHOI, H.; CHA, K.; CHUNG, H. Random forest as a potential multivariate method for near-infrared (NIR) spectroscopic analysis of complex mixture samples: Gasoline and naphtha. **Microchem. J.** v. 110, p. 739-748, 2013.
- [12] REBOUÇAS, M. V.; SANTOS, J. B.; PIMENTEL, M. F.; TEIXEIRA, L. S. G. A novel approach for development of a multivariate calibration model using a Doehlert experimental design: Application for prediction of key gasoline properties by Near-infrared Spectroscopy. **Chem. Intell. Lab. Syst.** v. 107, p. 185-193, 2011.
- [13] HONORATO, F. A.; NETO, B. B.; PIMENTEL, M. F.; STRAGEVITCH, L.; GALVÃO, R. K. H. Using principal component analysis to find the best calibration settings for simultaneous spectroscopic determination of several gasoline properties. **Fuel.** v. 87, p. 3706-3709, 2008.
- [14] ALVES, J. C. L.; POPPI, R. J. Quantification of conventional and advanced biofuels contents in diesel fuel blends using near-infrared spectroscopy and multivariate calibration. **Fuel.** v. 165, p. 379-388, 2016.
- [15] FENG, F.; WU, Q.; ZENG, L. Rapid analysis of diesel fuel properties by near infrared reflectance spectra. **Spectrochim Acta A.** v. 149, p. 271-278, 2015.
- [16] ALVES, J. C. L.; POPPI, R. J. Biodiesel content determination in diesel fuel blends using near infrared (NIR) spectroscopy and support vector machines (SVM). **Talanta.** v. 104, p. 155-161, 2013.

- [17] GONZAGA, F. B.; PASQUINI, C. A low cost short wave near infrared spectrophotometer: Application for determination of quality parameters of diesel fuel. **Anal. Chim. Acta.** v. 670, p. 92-97, 2010.
- [18] 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.
- [19] COOPER, J. B.; LARKIN, C. M.; ABDELKADER, M. F. Calibration transfer of near-IR partial least squares property models of fuels using virtual standards. **J. Chemometrics.** v. 25, p. 496-505, 2011.
- [20] COOPER, J. B.; LARKIN, C. M.; ABDELKADER, M. F. Virtual standard *slope* and *bias* calibration transfer of partial least squares jet fuel property models to multiple near infrared spectroscopy instruments. **J. Near Infrared Spectrosc.** v. 19, p. 139-150, 2011.
- [21] ABDELKADER, M. F.; COOPER, J. B.; LARKIN, C. M. Calibration transfer of partial least squares jet fuel property models using a segmented virtual standards *slope-bias* correction method. **Chem. Intell. Lab. Syst.** v. 110, p. 64-73, 2012.
- [22] DE PEINDER, P.; VISSER, T.; PETRAUSKAS, D. D.; SALVATORI, F.; SOULIMANI, F.; WECKHUUSEN, B. M. Prediction of Long-Residue Properties of Potential Blends from Mathematically Mixed Infrared Spectra of Pure Crude Oils by Partial Least-Squares Regression Models. **Energy & Fuels.** v. 23, p. 2164–2168, 2009.
- [23] AMAT-TOSELLO, S.; DUPUY, N.; KISTER, J. Contribution of external parameter orthogonalisation for calibration transfer in short waves—Near infrared spectroscopy application to gasoline quality. **Anal. Chim. Acta.** v. 642, p. 6-11, 2009.
- [24] LUTZ, O. M. D.; BONN, G. K.; RODE, B. M.; HUCK, C. W. Reproducible quantification of ethanol in gasoline via a customized mobile near-infrared spectrometer. **Anal. Chim. Acta.** v. 826, p. 61-68, 2014.
- [25] PAIVA, E. M.; ROHWEDDER, J. J. R.; PASQUINI, C.; PIMENTEL, M. F.; PEREIRA, C. F. Quantification of biodiesel and adulteration with vegetable oils in

diesel/biodiesel blends using portable near-infrared spectrometer, **Fuel**. v. 160, p. 57-63, 2015.

[26] PEREIRA, C. F.; PIMENTEL, M. F.; GALVÃO, R. K. H.; HONORATO, F. A.; STRAGEVITCH, L.; MARTINS, M. N. A comparative study of calibration transfer methods for determination of gasoline quality parameters in three different near infrared spectrometers. **Anal. Chim. Acta**. v. 611, p. 41-47, 2008.

[27] FEUDALE, R. N.; WOODY, N. A.; TAN, H.; MYLES, A. J.; BROWN, S. D.; FERRÉ, J. Transfer of multivariate calibration models: a review. **Chem. Intell. Lab. Syst.** v. 64, p. 181-192, 2002.

[28] BOIRET, M.; MEUNIER, L.; GINOT, Y. Tablet potency of Tianeptine in coated tablets by near infrared spectroscopy: Model optimisation, calibration transfer and confidence intervals. **J. Pharm. Biom. Anal.** v. 54, p. 510-516, 2011.

[29] SULUB, Y.; LOBRUTTO, R.; VIVILECCHIA, R.; WABUYELE, B. W. Content uniformity determination of pharmaceutical tablets using five near-infrared reflectance spectrometers: A process analytical technology (PAT) approach using robust multivariate calibration transfer algorithms. **Anal. Chim. Acta**. v. 611, p. 143-150, 2008.

[30] PEREIRA, L. S. A.; CARNEIRO, M. F.; BOTELHO, B. G.; SENA, M. M. Calibration transfer from powder mixtures to intact tablets: A new use in pharmaceutical analysis for a known tool. **Talanta**. v. 147, p. 351-357, 2016.

[31] PENG, J; PENG, S.; JIANG, A.; TAN, J. Near-infrared calibration transfer based on spectral regression. **Spectrochim. Acta A**. v. 78, p. 1315-1320, 2011.

[32] LIU, X.; HAN, L. J.; YANG, Z. L. Transfer of near infrared spectrometric models for silage crude protein detection between different instruments. **J. Dairy Sci.** v. 94, p. 5599-5610, 2011.

[33] FERNANDEZ-AHUMADA, E.; GARRIDO-VARO, A.; GUERRERO-GINEL, J. E.; PEREZ-MARIN, D.; FEARN, T. Taking NIR Calibrations of Feed Compounds from the Laboratory to the Process: Calibration Transfer between Predispersive and Post dispersive Instruments. **J. Agric. Food Chem.** v. 56, p. 10135-10141, 2008.

- [34] IGNE, B.; ROGER, J.; ROUSSEL, S.; BELLON-MAUREL, V.; HURBURGH, C.R. Improving the transfer of near infrared prediction models by orthogonal methods. **Chem. Intell. Lab. Syst.** v. 99, p. 57-65, 2009.
- [35] FAN, W.; LIANG, Y.; YUAN, D.; WANG, J. Calibration model transfer for near-infrared spectra based on canonical correlation analysis. **Anal. Chim. Acta.** v. 623, p. 22-29, 2008.
- [36] GALVÃO, R. K. H.; SOARES, S.F.C.; MARTINS, M.N.; PIMENTEL, M.F.; ARAÚJO, M.C.U. Calibration transfer employing univariate correction and robust regression. **Anal. Chim. Acta.** v. 864, p. 1-8, 2015.
- [37] DE LIRA, L. F. B.; DE VASCONCELOS, F. V. C.; PEREIRA, C. F.; PAIM, A. P. S.; STRAGEVITCH, L.; PIMENTEL, M. F. Prediction of properties of diesel/biodiesel blends by infrared spectroscopy and multivariate calibration. **Fuel.** v. 89, p. 405-409, 2010.
- [38] KRAMER, K. E.; MORRIS, R. E.; ROSE-PEHRSSON, S. L. Comparison of two multiplicative signal correction strategies for calibration transfer without standards. **Chem. Intell. Lab. Syst.** v. 92, p. 33-43, 2008.
- [39] KRAPF, L. C.; NAST, D.; GRONAUER, A.; SCHMIDHALTER, U.; HEUWINKEL, H. Transfer of a near infrared spectroscopy laboratory application to an online process analyser for in situ monitoring of anaerobic digestion. **Biores. Tech.** v. 129, p. 39-50, 2013.
- [40] SALGUERO-CHAPARRO, L.; PALAGOS, B. ; PEÑA-RODRÍGUEZ, F.; ROGER, J. M. Calibration transfer of intact olive NIR spectra between a pre-dispersive instrument and a portable spectrometer. **Computers and Electronics in Agriculture.** v. 96, p. 202-208, 2013.
- [41] BARREIRO, P.; HERRERO, D.; HERNÁNDEZ, N.; GRACIA, A.; LEÓN, L. Calibration Transfer Between Portable And Laboratory NIR Spectrophotometers. **Acta Hort. (ISHS).** v. 802, p. 373, 2008.
- [42] PIERNA, J. A. F.; VERMEULEN, P.; LECLER, B.; BAETEN, V.; DARDENNE, P. Calibration Transfer from Dispersive Instruments to Handheld Spectrometers. **Appl. Spectrosc.** v. 64, p. 644-648, 2010.

- [43] ZAMORA-ROJAS, E.; PÉREZ-MARÍN, D.; DE PEDRO-SANZ, E.; GUERRERO-GINEL, J. E.; GARRIDO-VARO, A. Handheld NIRS Analysis For Routine Meat Quality Control: Database Transfer From At-Line Instruments. **Chem. Intell. Lab. Syst.** v. 114, p. 30-35, 2012.
- [44] HOFFMANN, U.; PFEIFER, F.; HSUING, C.; SIESLER, H.W. Spectra Transfer Between a Fourier Transform Near-Infrared Laboratory and a Miniaturized Handheld Near-Infrared Spectrometer. **Appl. Spectrosc.** v. 70, p. 852-860, 2016.
- [45] DA SILVA, V. H.; DA SILVA, J. J.; PEREIRA, C. F. Portable near-infrared instruments: Application for quality control of polymorphs in pharmaceutical raw materials and calibration transfer. **J. Pharm. Biom. Anal.** v. 134, p. 287-294, 2017.
- [46] OHLWEILER, O. A. **Fundamentos de Análise Instrumental**. Rio de Janeiro: Livros Técnicos e Científicos, 1981.
- [47] PASQUINI, C. Near-Infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **J. Braz. Chem. Soc.** v. 14, n. 2, p. 198-219, 2003.
- [48] HOLLER, F. J.; SKOOG, D. A.; CROUCH, S. R. **Princípios de Análise Instrumental**. Tradução de Celio Pasquini, Jarbas José Rodrigues Rohwedder, *et al.* 6. ed. Porto Alegre: Bookman, 2009.
- [49] a. KINGTON, L. R.; JONES, T. M. Applications of Near Analysis of Beverages, p. 535-542; b. BUCHANAN, B. Near Analysis of Petrochemicals, p. 647-658; c. KRADJEL, C. Near Analysis of Polymers, p. 659-702; em **Handbook of Near-Infrared Analysis**. Edição de Donald A. Burns, Emil W. Ciurczak, 2. ed. Nova York: Marcel Dekker, 2001.
- [50] a. BOKOBZA, L. Origin of Near-Infrared Absorption Bands, 11 -39; b. SIESLER, H. W. Introduction, 1 -10; em **Near-Infrared Spectroscopy: Principles, Instruments, Applications**. Edição de H. W. Siesler, Y. Ozaki, *et al.* Weinheim: WILEY-VCH, 2002.
- [51] SARRAGUÇA, M. C.; LOPES, J. A. Quality control of pharmaceuticals with NIR: From lab to process line. **Vib. Spectrosc.** v. 49, p. 204-210, 2009.
- [52] BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. **Chemometrics: A Pratical Guide**, New York: Wiley-Interscience, 1998.

- [53] DHANOA, M. S. *et al.* The link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) transformations of NIR spectra. **J. Near Infrared Spectrosc.** v. 2, p. 43-47, 1994.
- [54] RINNAN, Å.; VAN DEN BERG, F.; ENGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. **Trends Anal. Chem.** v. 28, n. 10, p. 1201-1222, 2009.
- [55] NAES, T.; ISAKSSON, T.; FEARN, T.; DAVIES, T. **A user-friendly guide to multivariate calibration and classification**, Chichester: NIR Publications, 2002.
- [56] BRERETON, R. G. **Chemometrics: Data Analysis for the Laboratory and Chemical Plant**, Chichester: Wiley, 2003.
- [57] WESTAD, F.; MARTENS, H. Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression. **J. Near Infrared Spectrosc.** v. 8, p. 117-124, 2000.
- [58] NORGAARD, L. *et al.* Interval Partial Least-Squares Regression (*i*PLS): A comparative chemometric study with an example from Near-Infrared Spectroscopy. **Appl. Spectrosc.** v. 54, n. 3, p. 413-419, 2000.
- [59] GALVÃO, R. K. H. *et al.* A method for calibration and validation subset partitioning. **Talanta.** v. 67, p. 736-740, 2005.
- [60] INMETRO, Instituto Nacional de Metrologia, Normalização e qualidade industrial. **Orientação sobre validação de métodos de ensaios químicos**. QOQCGCRE 088, Revisão 03 fevereiro/2010.
- [61] BRUNS, R. E.; SCARMINIO, I. S.; NETO, B. B. **Statistical design - Chemometrics**, The Netherlands: Elsevier, 2006.
- [62] **STANDARD PRACTICES FOR INFRARED MULTIVARIATE QUANTITATIVE ANALYSIS**. ASTM - American Society for Testing and Materials. ASTM E1655-05.

- [63] MILLER, J.N.; MILLER, J.C. **Statistics and Chemometrics for Analytical Chemistry**, 6. ed., England: Pearson Education, 2010.
- [64] WANG, Y.; VELTKAMP, D. J.; KOWALSKI, B. R. Multivariate instrument standardization, **Anal. Chem.** v. 63, p. 2750-2756, 1991.
- [65] FEUDALE, R. N.; WOODY, N. A.; TAN, H.; MYLES, A. J.; BROWN, S. D.; FERRÉ, J. Transfer of multivariate calibration models: a review, **Chem. Int. Lab. Syst.** v. 64, p. 181, 2002.
- [66] HONORATO, F. A.; NETO, B.B.; MARTINS, M.N.; GALVÃO, R.K.H.; PIMENTEL, M.F. Transferência de calibração em métodos multivariados, **Quim. Nova.** v. 30, p. 1301, 2007.
- [67] BOLETIM DE MONITORAMENTO DA QUALIDADE DOS COMBUSTÍVEIS. **ANP – Agência Nacional de Petróleo, Gás Natural e Biocombustíveis**. Disponível em:<
http://www.anp.gov.br/wwwanp/images/publicacoes/boletins-anp/Boletim_Monitoramento_Qualidade_Combustiveis/BQ_COMBUSTIVEIS_201612_rev.pdf>. Acessado em: 28 de janeiro de 2017.
- [68] **SIGNIFICANCE OF TESTS FOR PETROLEUM PRODUCTS**. Edição de Rand, S. J. (8ª edição). Newburyport: ASTM INTERNATIONAL, 2010.
- [69] WORKMAN Jr., J.; WEYER, L. **Practical Guide to Interpretive Near-infrared spectroscopy**, Boca Raton: CRC Press Taylor & Francis Group, 2007.
- [70] WEYER, L. G. Near-Infrared Spectroscopy of Organic Substances. **Appl. Spectrosc. Rev.** v. 21, p. 1-43, 1985.

APÊNDICE A –
Rotina em Matlab da função “simugas” do simulador NIRSIM.

```

function [form,MON_simu,RON_simu,Espec_simu] = simugas(T1,T2,T3,T4,Formulacao)

% Esta etapa é para pré-processar os espectros das correntes;
if isempty(T4)
    T1pre = baseline(savgol(T1,25,2,0), [2800 3800]);
    T2pre = baseline(savgol(T2,25,2,0), [2800 3800]);
    T3pre = baseline(savgol(T3,25,2,0), [2800 3800]);
else if ~isempty(T4)
    T1pre = baseline(savgol(T1,25,2,0), [2800 3800]);
    T2pre = baseline(savgol(T2,25,2,0), [2800 3800]);
    T3pre = baseline(savgol(T3,25,2,0), [2800 3800]);
    T4pre = baseline(savgol(T4,25,2,0), [2800 3800]);
end
end

% Esta etapa realiza as simulações possíveis;
clear cartProd form sets3c sets4c

if isempty(Formulacao)
    if isempty(T4)
        sets3c = {40:5:60, 10:5:25, 25:5:35};
        [x,y,z] = ndgrid(sets3c{:});
        cartProd = [x(:) y(:) z(:)];
        for i = 1:size(cartProd,1);
            if (x(i) + y(i) + z(i)) == 100;
                form(i,:) = [x(i) y(i) z(i)]/100;
            end
        end
        end

        form(all(form==0,2),:)=[];
        Simulacoes = zeros(length(form),4148);
        for i = 1:size(form,1)
            Simulacoes(i,:)=(T1pre*form(i,1)+T2pre*form(i,2)+T3pre*form(i,3));
        end

    else if length(T4) == 4148
        sets4c = {15:5:35, 10:5:25, 25:5:35, 10:5:35};
        [x,y,z,w] = ndgrid(sets4c{:});
        cartProd = [x(:) y(:) z(:) w(:)];
        for i = 1:size(cartProd,1);
            if (x(i) + y(i) + z(i) + w(i)) == 100;
                form(i,:) = [x(i) y(i) z(i) w(i)]/100;
            end
        end
        end

        form(all(form==0,2),:)=[];
        Simulacoes = zeros(length(form),4148);
        for i = 1:size(form,1)
            Simulacoes(i,:)=(T1pre*form(i,1)+T2pre*form(i,2)+T3pre*form(i,3)+T4pre*form(i,4));
        end
    else
        error('As dimensões de FB963B estão incorretas')
    end
end

else if length(Formulacao) == 3;
    if (Formulacao(1,1) + Formulacao(1,2) + Formulacao(1,3)) == 100;
        Simulacoes = (T1pre*Formulacao(1)/100 + T2pre*Formulacao(2)/100 +
            T3pre*Formulacao(3)/100);
    end
end

```

APÊNDICE A – Continuação.

```

        form = Formulacao;
    else
        warningMessage = sprintf('As percentagens não somam 100%');
        uiwait(warndlg(warningMessage));
        return;
    end
else if length(Formulacao) == 4;
    if (Formulacao(1,1) + Formulacao(1,2) + Formulacao(1,3) +
        Formulacao(1,4)) == 100;
        Simulacoes = (T1pre*Formulacao(1)/100 +
            T2pre*Formulacao(2)/100 + T3pre*Formulacao(3)/100 +
            T4pre*Formulacao(4)/100);
        form = Formulacao;
    else
        warningMessage = sprintf('As percentagens não somam 100%');
        uiwait(warndlg(warningMessage));
        return;
    end
end
end
end

% Essa etapa carrega os modelos PLS para MON e RON
load('ModelosPLS.mat','PLS_Model_ME_MON','PLS_Model_ME_RON');
coefreg_MON = PLS_Model_ME_MON.content.reg;
meanspec_MON = PLS_Model_ME_MON.content.detail.preprocessing{1, 1}.out{1,
1}(1,1:2700);
mean_MON = PLS_Model_ME_MON.content.detail.preprocessing{1, 2}.out{1,1};
std_MON = PLS_Model_ME_MON.content.detail.preprocessing{1, 2}.out{1,2};

coefreg_RON = PLS_Model_ME_RON.content.reg;
meanspec_RON = PLS_Model_ME_RON.content.detail.preprocessing{1, 1}.out{1,
1}(1,1:2700);
mean_RON = PLS_Model_ME_RON.content.detail.preprocessing{1, 2}.out{1,1};
std_RON = PLS_Model_ME_RON.content.detail.preprocessing{1, 2}.out{1,2};

% Essa etapa realiza as previsões para MON
Simulacoes_mc_MON = Simulacoes(:,1:2700) -
repmat(meanspec_MON,size(Simulacoes,1),1);
Est1_MON = Simulacoes_mc_MON*coefreg_MON;
MON_simu = (Est1_MON*std_MON)+mean_MON;

% Essa etapa realiza as previsões para RON
Simulacoes_mc_RON = Simulacoes(:,1:2700) -
repmat(meanspec_RON,size(Simulacoes,1),1);
Est1_RON = Simulacoes_mc_RON*coefreg_RON;
RON_simu = (Est1_RON*std_RON)+mean_RON;

% Essa etapa reúne e plota as formulações possíveis
Espec_simu = Simulacoes(:,1:2700);
plot(MON_simu,RON_simu,'*');

```

ANEXO A –
Tabela dos parâmetros de qualidade da gasolina automotiva comum.

Característica	Unidade	Método de Referência	Especificação	
			Tipo A	Tipo C
Aspecto	-	NBR14954 / D4176	límpido e isento de impurezas	
Cor	-	visual	(1)	(2)
Massa Específica a 20°C	kg/m ³	NBR7148, NBR14065 / D1298, D4052	Anotar	
Teor de Álcool Etílico Anidro Combustível (AEAC)	% vol.	NBR13992 / D5501	(3)	(4)
10% evaporado, máx.	°C	NBR9619 / D86	65,0	
50% evaporado, máx.	°C	NBR9619 / D86	120,0	80,0
90% evaporado, máx.	°C	NBR9619 / D86	190,0	
TPF, máx.	°C	NBR9619 / D86	215,0	
Enxofre	mg/kg	D2622	-	50
Infravermelho (Teor de benzeno)	% vol.	NBR15289, NBR15441 / D3606, D5443, D6277	-	1,0
Infravermelho (MON, RON, IAD, e composição - olefinas, saturados e aromáticos)	-	D6277	-	-
Marcador de solvente	-	Conforme metodologia de identificação do fornecedor do marcador	-	-

(1) De incolor a amarelada, isenta de solvente. (2) De incolor a amarelada, se isenta de corante - teor máximo 50ppm, exceto cor azul. (3) Proibida a adição. (4) Teor de EAC conforme legislação vigente [6].

ANEXO B –
Tabela dos parâmetros de qualidade do óleo diesel BX a B30.

Característica	Unidade	Método de Referência	Especificação	
			S10	S500
Aspecto	-	Visual	límpido e isento de impurezas	
Cor	-	visual	(1)	Vermelho (2)
Cor ASTM, máx.	-	NBR14483 / D6045, D1500		3,0
Massa Específica a 20°C	kg/m ³	NBR7148, NBR14065, NBR15983 / D1298, D4052, D7042	817,8 a 865,0	817,8 a 875,5
Viscosidade a 40°C	mm ² /s	NBR10441 / D445	1,9 a 4,1	2,0 a 5,0
Teor de Biodiesel, % vol.	% volume	NBR15568 / EN14078 / D7861, D7371		(3)
Enxofre, máx.	mg/kg	D5453, D2622, D7039, D7220	10	-
10% evaporado, máx.	°C	D5453, D2622, D7039, NBR9619 / D86, D1160	180	-
50% evaporado, máx.	°C	NBR9619 / D86, D1160	245,0 a 295,0	245,0 a 310,0
85% evaporado, máx.	°C	NBR9619 / D86, D1160		Anotar
90% evaporado, máx.	°C	NBR9619 / D86, D1160		Anotar

ANEXO 2 – Continuação.

Característica	Unidade	Método de Referência	Especificação	
			S10	S500
95% evaporado, máx.	°C	NBR9619 / D86, D1160	Anotar	
Ponto de fulgor, °C, mín.	°C	NBR7974, NBR14598 / D56, D93, D3828	38	
Ponto de entupimento de filtro a frio, máx.	°C	NBR14747 / D6371	Limites de acordo com estado	
Resíduo de carbono – 10% amostra, máx.	%massa	NBR14318 / D524	0,25	
Cinzas, máx.	%massa	NBR9842 / D482	0,01	
Corrosividade ao cobre	-	NBR14359 / D130	1	
Teor de água, máx.	mg/kg	D6304 / EN ISO 12937	200	500
Água e sedimentos, máx.	% vol	D2709	-	0,05
Estabilidade á oxidação, mín.	h	EN15751	20	
Contaminação total, máx.	mg/kg	EN12662	24	Anotar
Número de Acidez, máx.	mg KOH/g	NBR14248 / D664	0,3	
Condutividade elétrica, mín.	pS/m	D2624, D4308	25	
Número de cetano ou número de cetano derivado (NCD), mín.	-	D613, D6890, D7170, D7668	48	42

(1) De incolor a amarelada. (2) Corante vermelho especificado pela Resolução ANP n. 50/2013, ou outra que venha substituí-la. (3) Admitida variação de $\pm 0,5\%$ vol. para óleo diesel abaixo de B20 e $\pm 1,0\%$ vol. de B20 a B30.