

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Instituto de Ciências Exatas

Programa de Pós-graduação em Química

Francisca Gabriela Lopes Rosado

**APLICAÇÃO DE TÉCNICAS
QUIMIOMÉTRICAS NA
DETERMINAÇÃO DE PARÂMETROS
FÍSICO-QUÍMICOS PARA
COMBUSTÍVEIS**

Belo Horizonte

2025

UFMG/ICEX/DQ. 1.654
T. 751

Francisca Gabriela Lopes Rosado

**APLICAÇÃO DE TÉCNICAS
QUIMIOMÉTRICAS NA DETERMINAÇÃO DE
PARÂMETROS FÍSICO-QUÍMICOS PARA
COMBUSTÍVEIS**

Tese apresentada ao Departamento de Química do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de Doutora em Ciências - Química.

Orientador: Prof. Dr. Paulo Jorge Sanches Barbeira

Belo Horizonte

2025

Ficha Catalográfica

R789a
2025
T

Rosado, Francisca Gabriela Lopes.

Aplicação de técnicas quimiométricas na determinação de parâmetros físico-químicos para combustíveis [manuscrito] / Francisca Gabriela Lopes Rosado. 2025. 84 f. : il., gráfs., tabs.

Orientador: Paulo Jorge Sanches Barbeira.

Tese (doutorado) – Universidade Federal de Minas Gerais – Departamento de Química.

Bibliografia: f. 78-84.

1. Química analítica – Teses. 2. Quimiometria – Teses. 3. Aviões – Combustíveis – Teses. 4. Biodiesel – Teses. 5. Combustíveis – Análise – Teses. 6. Espectroscopia de fluorescência – Teses. 7. Análise discriminante – Teses. I. Barbeira, Paulo Jorge Sanches, Orientador. II. Título.

CDU 043

Elaborada por Márcia Cavalcanti Moreira – Bibliotecária – CRB6/3847.



UNIVERSIDADE FEDERAL DE MINAS GERAIS

UFMG

Programa de Pós-Graduação em Química
Departamento de Química - ICEX



"Aplicação de Técnicas Quimiométricas Na Determinação de Parâmetros Físico- químicos Para Combustíveis"

Francisca Gabriela Lopes Rosado

Tese aprovada pela banca examinadora constituída pelos Professores:

Prof. Paulo Jorge Sanches Barbeira - Orientador UFMG

Profª. Mariana Ramos de Almeida UFMG

Prof. Helvécio Costa Menezes UFMG

Prof. João Flávio da Silveira Petrucci UFU

Prof. Habdias de Araujo Silva Neto UFSC

Belo Horizonte, 28 de maio de 2025.



Documento assinado eletronicamente por **Paulo Jorge Sanches Barbeira, Professor do Magistério Superior**, em 28/05/2025, às 19:23, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Habdias de Araujo Silva Neto, Usuário Externo**, em 05/05/2025, às 20:01, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **João Flávio da Silveira Petrucci, Usuário Externo**, em 06/06/2025, às 16:20, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Mariana Ramos de Almeida, Professora do Magistério Superior**, em 03/06/2025, às 23:14, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Helvecio Costa Menezes, Professor do Magistério Superior**, em 06/06/2025, às 15:49, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [/sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0](http://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4244273** e o código CRC **49BE0297**.

Referência: Processo nº 23072.232647/2025-17 SEI nº 4244273

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus por me guiar e proteger ao longo de toda a minha trajetória.

Aos meus pais e demais familiares, minha eterna gratidão por todo amor, compreensão e apoio incondicional. Sem vocês, essa caminhada não teria sido possível.

Ao meu orientador, professor Dr. Paulo Jorge Sanches Barbeira, agradeço profundamente pela orientação dedicada, confiança, paciência, disponibilidade e por me conduzir com sabedoria e rigor científico ao longo de toda a pesquisa. Sua orientação foi fundamental para o amadurecimento deste trabalho e para minha formação como pesquisadora.

Aos colegas do grupo de pesquisa PRH 1.1, meu sincero agradecimento: à Natalia, por tornar os dias na UFMG mais leves; ao Henrique, por ouvir minhas reclamações e me orientar nas vezes em que deixei de ler as orientações do grupo; e ao Gabriel, por ser um parceiro de trabalho exemplar. Agradeço também à coordenadora do programa, professora Vânia Pasa, pelo constante incentivo e confiança.

Aos colegas do LEC, sou grata pelo apoio, pela disponibilidade, pelas contribuições valiosas e pelo ambiente intelectual enriquecedor, que tanto favoreceu meu crescimento acadêmico e pessoal.

Aos amigos e colegas que, de diferentes formas, estiveram presentes nos momentos difíceis, oferecendo apoio, escuta e palavras de encorajamento, meu profundo agradecimento. Em especial, à Fabi e à Naty, por lidarem com meu mau humor nos dias mais desafiadores.

Por fim, agradeço o apoio financeiro do Programa de Recursos Humanos da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – PRH-ANP, viabilizado com recursos provenientes do investimento de empresas petrolíferas qualificadas na Cláusula de P, D&I da Resolução nº 918/2023 ANP, bem como à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo nº 2024/12519-5.

A todas e todos, meus mais sinceros agradecimentos.

RESUMO

Combustíveis derivados de petróleo, os chamados combustíveis não renováveis, são a principal fonte de energia mundial. O querosene de aviação (QAV) é um combustível utilizado em turbinas aeronáuticas, obtido por meio da destilação direta do petróleo com faixa de temperatura de 150 a 300 °C, composto predominantemente por hidrocarbonetos parafínicos contendo de 9 a 15 átomos de carbono. Dentro da classe dos combustíveis renováveis temos o biodiesel (B100), este biocombustível vem sendo misturado ao diesel mineral e é comercializado para o funcionamento dos motores a diesel. Os combustíveis seja ele renovável ou não, seguem as especificações estabelecidas pela ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis) que visa estabelecer os padrões de qualidade. Um dos maiores problemas envolvendo estes combustíveis é a não conformidade de acordo com as regulamentações de qualidade. Para o biodiesel esse problema está relacionado com a formação de sólidos provenientes da sua oxidação, que é cada vez mais comum à medida que o teor de biodiesel aumenta na mistura diesel/biodiesel. Já para o QAV, os problemas estão mais relacionados aos parâmetros físico-químicos como o Ponto de Congelamento e a contaminação por diferentes solventes. Atualmente, existem métodos padronizados de análise de combustíveis, tais métodos são, em geral, dispendiosos, consomem grandes quantidades de amostra e utilizam equipamentos específicos. Neste contexto, o objetivo do trabalho foi desenvolver modelos de classificação supervisionada e não supervisionada e modelos de calibração para identificar e prever as não conformidades relacionadas a estes combustíveis. O uso da Espectrofluorimetria Sincronizada em conjunto com os métodos quimiométrico PLS e PLS-DA possibilitou a obtenção de modelos capaz de prever a Estabilidade Oxidativa do B100 com R^2_{val} de 0,822 e quantificar o teor de contaminantes no QAV com R^2_{val} acima de 0,9 para as amostras contaminadas com gasolina (GA), gasolina de aviação (GAV) e querosene iluminante (QL) e para o contaminante querosene comercial (QC) o R^2_{val} foi de 0,741. Já os modelos de classificação tanto para os tipos de contaminantes quanto para o Ponto de congelamento do QAV obtiveram uma taxa de eficiência acima de 80% para o conjunto de teste dos modelos. Por fim, a utilização da Espectrofluorimetria Sincronizada associada as técnicas quimiométricas possibilitou a análise de dois tipos de combustíveis se a necessidade de nenhum preparo de amostra de maneira rápida e eficaz quando comparada aos métodos estabelecidos pela ANP.

Palavras-chave: querosene de aviação; biodiesel; espectrofluorimetria; classificação supervisionada; espectroscopia de fluorescência síncrona; PLS-DA.

ABSTRACT

Petroleum-derived fuels, known as non-renewable fuels, constitute the primary global energy source. Aviation kerosene (Jet A-1), depending on the standard) is a fuel utilized in aircraft turbines, obtained through the direct distillation of petroleum within a temperature range of 150 to 300 °C, and is predominantly composed of paraffinic hydrocarbons containing 9 to 15 carbon atoms. Within the class of renewable fuels, we have biodiesel (B100). This biofuel has been increasingly blended with mineral diesel and is commercially available for the operation of diesel engines. Fuels, whether renewable or non-renewable, adhere to specifications established by the ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis), which aims to set quality standards. A major issue involving these fuels is non-compliance with quality regulations. For biodiesel, this problem is related to the formation of solids resulting from its oxidation, a phenomenon increasingly common as the biodiesel content increases in diesel/biodiesel blends. Regarding Jet A-1, the problems are more associated with physicochemical parameters such as the Freezing Point and contamination by different solvents. Currently, standardized fuel analysis methods exist; however, these methods are generally expensive, consume large quantities of sample, and require specific equipment. In this context, the objective of this work was to develop supervised and unsupervised classification models and calibration models to identify and predict non-conformities related to these fuels. The use of Synchronous Fluorescence Spectroscopy (SFS) in conjunction with chemometric methods such as PLS and PLS-DA enabled the development of models capable of predicting the Oxidative Stability of B100 with a validation R^2 of 0.822 and quantifying the content of contaminants in Jet A-1 with a validation R^2 above 0.9 for samples contaminated with gasoline (GA), aviation gasoline (GAV), and illuminating kerosene (QL). For the contaminant commercial kerosene (QC), the validation R^2 was 0.741. Furthermore, the classification models for both the types of contaminants and the Freezing Point of Jet A-1 achieved an efficiency rate above 80% for the test set of the models. Finally, the utilization of Synchronous Fluorescence Spectroscopy associated with chemometric techniques enabled the analysis of two types of fuels without the need for any sample preparation, in a rapid and effective manner when compared to the methods established by the ANP.

Keywords: aviation kerosene; biodiesel; spectrofluorimetry; supervised classification; synchronous fluorescence spectroscopy; PLS-DA.

LISTA DE ABREVIATURAS

ANP - Agência Nacional de Petróleo, Gás Natural e Biocombustíveis

ASTM - American Society for Testing and Materials

CVCE - Erro de classificação de validação cruzada (*Cross-Validation Classification Error*)

FTIR - Infravermelho com transformada de Fourier (*Fourier Transform Infrared*)

IR - Infravermelho (*Infrared*)

LDA - Análise Discriminante Linear (*Linear Discriminant Analysis*)

LEC - UFMG - Laboratório de Ensaio de Combustíveis da Universidade Federal de Minas Gerais

MCC - Coeficiente de correlação de Matthews (*Matthews Correlation Coefficient*)

MIR - Infravermelho médio (*Mid-Infrared*)

MLR - Regressão linear múltipla (*Multiple-linear regression*)

MS - Espectrometria de massas (*Mass Spectrometry*)

MSC - Correção do espalhamento multiplicativo (*Multiplicative Scatter Correction*)

NIR - Infravermelho próximo (*Near-Infrared*)

PC - Componente principal (*Principal Component*)

PCA - Análise de componentes principais (*Principal Component Analysis*)

PCR - *Principal Component Regression*

PLS - Mínimos quadrados parciais (*Partial Least Squares*)

PLS-DA - Análise discriminante por mínimos quadrados parciais (*Partial Least Square Discriminant Analysis*)

R^2 - Coeficiente de Determinação

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

RMSEP - Raiz quadrada do erro quadrático médio de predição (*Root mean square error of prediction*)

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

SNV - Variação normal padrão (*Standard Normal Variate*)

VL - Variáveis latentes

EFC - Espectrofluorimetria de Fluorescência Convencional

EFS – Espectrofluorimetria de Fluorescência Sincronizada

LISTA DE FIGURAS

Figura 1- Efeito da simplificação espectral da espectrofluorimetria sincronizada.	17
Figura 2- (a) Espectros de excitação e emissão do tetraceno. (b) Sinal de fluorescência sincronizada do tetraceno.	17
Figura 3- Esquema de uma PCA.	20
Figura 4- PCA para dados espectrais.	21
Figura 5- Planejamento fatorial com ponto central (A), planejamento composto central 22 em estrela (B) e planejamento composto central 22 de face centrada (C).	27
Figura 6- Espectros de fluorescência síncrona para todas as amostras B100 em diferentes valores de $\Delta\lambda$. a) 10; b) 30; c) 70; d) 100 nm.	36
Figura 7- Diagrama de contorno para erros de classificação de validação cruzada, R^2_{cal} , R^2_{val} e razão entre erros de calibração e validação para espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 30 nm.	38
Figura 8 - Valores de $RMSEC$ e $RMSEP$ para diferentes valores de $\Delta\lambda$.	39
Figura 9 - Gráfico das variáveis significativas da regressão (VIP scores) para os modelos com diferentes valores de $\Delta\lambda$. a) 30; b) 40 nm.	40
Figura 10- Gráfico dos coeficientes de regressão para os modelos com diferentes valores de $\Delta\lambda$. a) 30; b) 40 nm.	41
Figura 11- Gráfico do Y predito pelo Y medido para as amostras de B100 para diferentes modelos normalizados para $\Delta\lambda$ de 30 nm. (●) Calibração; (▼) Validação.	44
Figura 12- Espectros de fluorescência sincronizada ($\Delta\lambda$ 10 nm) das matrizes de QAV (M1, M2, M3 e M4) e os contaminantes (GA, QL, QC, GAV).	49
Figura 13- Cromatograma dos cinco continentes: Azul-QL, Rosa-QC, Perto-QAV, Vermelho-D, Verde-GA, e Azul escuro-GAV.	50
Figura 14- Espectros de fluorescência sincronizada para matrizes de QAV dopadas com GA, GAV, QC e QL obtidos em diferentes valores de $\Delta\lambda$: a) 10; b) 40 nm.	52
Figura 15- Gráfico de superfície de contorno do planejamento de face centrada $\Delta\lambda$ de 20 nm.	53
Figura 16- PLS2-DA para discriminação de QAV-NC e QAV contaminados com: GA, QL, QC e GAV. Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.	55
Figura 17- PLS-DA para discriminação de QAV contaminados com: GA+GAV (Classe 1) e QL+QC (Classe 2). Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.	56
Figura 18- PLS-DA para discriminação de QAV contaminados com GA (Classe1) e GAV (Classe 2). Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.	58
Figura 19- PLS-DA para discriminação de QAV contaminados com QL e QC. Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.	60
Figura 20- Gráficos VIP scores dos modelos PLS-DA e PLS2-DA de fluorescência síncrona para classificação das amostras contaminadas com: a) QAV(NC) x QAV (GA, GAV, QL e QC), b) QAV(NC) x QAV(C), c) QAV(GA) x QAV(GAV) e d) QL x QC.	61
Figura 21 - Gráficos dos Y preditos pelo Y medidos para as amostras de QAV.	64
Figura 22 - Estrutura química da 1,4 dialquilamino antraquinona.	65

- Figura 23 - Gráfico das variáveis significativas da regressão (VIP scores) para os modelos de quantificação de contaminantes em QAV. 66
- Figura 24 - PLS2-DA para discriminação de QAV em diferentes faixas de Ponto de congelamento. Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite. 73
- Figura 25 - Gráficos VIP escores dos modelos PLS2-DA de fluorescência síncrona para classificação das amostras em diferentes faixas de Ponto de congelamento. (a) classe das amostras Reprovadas, (b) classe das amostras com valores Críticos e (c) classe das amostras Aprovadas. 75

LISTA DE TABELAS

Tabela 1 - Relação de estudos usando espectrofluorimetria sincronizada e quimiometria para otimização de ensaios para biodiesel e diesel.	19
Tabela 2 - Intervalo de valores de estabilidade oxidativa para as 65 amostras B100 usadas nos modelos.	33
Tabela 3 - Tamanho da janela, grau polinomial e erros de classificação nos conjuntos de treinamento, validação cruzada e teste para o melhor modelo em cada $\Delta\lambda$ (nm).	38
Tabela 4 - Tamanho da janela (J), grau polinomial (P), coeficientes de determinação, erro quadrático médio da validação cruzada e razão entre os erros quadráticos médios da predição e classificação com e sem normalizações para $\Delta\lambda$ de 30 e 40 nm.	42
Tabela 5 - Tabela com trabalhos encontrados na literatura para Estabilidade Oxidativa de B100.	43
Tabela 6 - Tamanho de Janela e grau de polinômio para os melhores modelos em cada $\Delta\lambda$ (nm).	53
Tabela 7- Parâmetros de desempenho calculados para o modelo PLS2-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 10 nm para classificar QAV contaminado com quatro classes de contaminantes.	55
Tabela 8- Parâmetros de desempenho calculados para o modelo PLS-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 10 nm para classificar QAV contaminado com duas classes de contaminantes: GA+GAV (classe 1) versus QL+QC (classe 2).	56
Tabela 9 - Parâmetros de desempenho calculados para o modelo PLS2-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ 40 nm para classificar QAV-NC e QAV contaminados com GA e GAV.	58
Tabela 10 - Parâmetros de desempenho calculados para o modelo PLS-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 20 nm para classificar QAV contaminado com: QL (Classe 1) e QC (Classe 2).	59
Tabela 11 - Valores dos comprimentos de onda (nm) presentes nos gráficos de importância variável na projeção VIP escores de todos os modelos PLS-DA e PLS2-DA para classificação de amostras de QAV contaminados.	62
Tabela 12 - Parâmetros de desempenho dos modelos PLS calculados para os modelos de quantificação de contaminantes em amostras de QAV.	63
Tabela 13 - Valores dos comprimentos de onda (nm) presentes nos gráficos de importância variável na projeção VIP escores de todos os modelos PLS para quantificação de amostras de QAV contaminados.	65
Tabela 14 - Tabela com trabalhos encontrados na literatura.	67
Tabela 15 - Distribuição das 48 amostras de QAV, em diferentes faixas de valores do Ponto de Congelamento para os modelos PLS-DA.	70
Tabela 16 - Parâmetros de desempenho calculados para o modelo PLS2-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 10 nm para classificar QAV em diferentes faixas de PC.	72
Tabela 17 - Valores dos comprimentos de onda (nm) presentes nos gráficos de importância variável na projeção VIP escores para os modelos PLS2-DA para os faixas de PC.	74
Tabela 18 - Tabela com trabalhos encontrados na literatura que avaliaram o ponto de congelamento de combustível de aviação, utilizando diferentes técnicas.	76

SUMÁRIO

1.0	INTRODUÇÃO	14
1.1	<i>Combustíveis</i>	14
1.2.1	<i>Espectrofluorimetria</i>	16
1.2.2	<i>Espectrofluorimetria Síncrona</i>	16
1.3	<i>Quimiometria</i>	20
1.3.1	<i>Métodos de Classificação</i>	20
1.3.2.	<i>Métodos não supervisionados</i>	20
1.3.4	<i>Análise Discriminante pelo Método dos Quadrados Mínimos Parciais PLS-DA</i> ...	
1.3.5.	<i>Modelos de Calibração</i>	24
1.4	Otimização de variáveis através do Planejamento de Face Centrada	26
2.0	OBJETIVOS	28
3.0	MODELO QUIMIOMETRICO PARA PREVISÃO DA ESTABILIDADE OXIDATIVA DO BIODIESEL	29
3.1.	<i>Introdução</i>	29
3.2	<i>Materiais e métodos</i>	33
3.3	<i>Resultados e discussões</i>	35
3.4	<i>Conclusão</i>	44
4.0	CLASSIFICAÇÃO E QUANTIFICAÇÃO DE CONTAMINANTES EM QUEROSENE DE AVIAÇÃO	45
4.1	<i>Introdução</i>	45
4.2	<i>Materiais e métodos</i>	47
4.3	<i>Resultados e discussões</i>	49
4.4	<i>Conclusão</i>	68
5.0	MODELO QUIMIOMETRICO PARA PREVISÃO DO PONTO DE CONGELAMENTO EM QUEROSENE DE AVIAÇÃO	69
5.1	<i>Introdução</i>	69
5.2	<i>Materiais emétodos</i>	70
5.3	<i>Resultados e discussão.</i>	71
5.4	<i>Conclusão</i>	76
6.0	CONCLUSÃO GERAL	77
	REFERÊNCIAS BIBLIOGRÁFICAS	78

1. INTRODUÇÃO

1.1. Combustíveis

Os combustíveis desempenham um papel importante para diversas atividades humanas seja na área industrial, doméstica ou nos transportes. Na área de transportes, diferentes combustíveis são utilizados no Brasil para atender as diferentes necessidades dos motores a combustão e turbinas aeronáuticas, tendo origem no processamento do petróleo (como o gás natural, o óleo diesel, o querosene de aviação e as gasolinas automotivas e de aviação) ou sintetizados a partir de matérias primas vegetais e animais (como o etanol e o biodiesel).

No Brasil, desde a publicação da Lei nº 9.478, de 6 de agosto de 1997, cabe à Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), entre outras competências, promover a regulação, a contratação e a fiscalização das atividades econômicas integrantes da indústria do petróleo, do gás natural e dos biocombustíveis. Em decorrência dessas atribuições, a agência tem empreendido esforços contínuos para melhorar a qualidade dos combustíveis por meio da especificação de características físico-químicas que garantam que os combustíveis tenham a qualidade mínima necessária para atender aos anseios da sociedade quanto à eficiência energética e ao meio ambiente [1].

Dentre os combustíveis o querosene de aviação e o biodiesel se destacam pelo maior número de parâmetros de qualidade regulados pelas portarias da ANP em relação aos demais combustíveis, que abrangem diferentes propriedades físico-químicas e teores de constituintes e contaminantes específicos. Além dos limites estabelecidos para os diferentes parâmetros, as portarias da ANP [2-3] designam os métodos de ensaio a serem utilizados, que são normatizados e estabelecidos por diferentes organismos como a ASTM e a ABNT, dentre outros.

Diante do grande número de parâmetros a serem analisados pelos laboratórios responsáveis por esses monitoramentos, da complexidade dos ensaios, do custo dos equipamentos e da elevada quantidade de amostra necessária, em alguns casos, para se obter resultados satisfatórios, tem-se buscado desenvolver metodologias alternativas para a determinação dos parâmetros de qualidade dos combustíveis que demandem menos tempo e custo, obtendo-se resultados confiáveis para um elevado número de amostras a serem analisadas diariamente.

Uma das formas para a otimização dos trabalhos de fiscalização de combustíveis é a obtenção de modelos quimiométricos aliado às técnicas rápidas, sem preparo de amostra, que permita análise “*in-situ*” e que sirvam como “*screening*”, ou seja, uma ferramenta ou metodologia utilizada para identificar, de forma rápida e eficiente, grupos que apresentam características específicas, auxiliando assim os ensaios padronizados na determinação das propriedades físico-

químicas dos combustíveis. Os modelos quimiométricos, tanto os de classificação como a análise de componentes principais (*Principal Component Analysis*, PCA) e a análise discriminante pelo método dos quadrados mínimos parciais (*Partial Least Squares Discriminant Analysis*, PLS-DA) quanto os de calibração (PLS), são na maioria das vezes, usados juntamente com as técnicas analíticas de espectroscopia no infravermelho (FTIR- NIR / MIR) [4], métodos cromatográficos [5], ressonância magnética nuclear (RMN)[6], espectrofluorimetria convencional e síncrona (EFC, EFS) [7-8] dentre outras.

Dentre as técnicas analíticas citadas, a EFS tem como vantagens quando comparada com a espectroscopia eletrônica maiores sensibilidade e seletividade e a característica multidimensional [8-9], mesmo sendo uma técnica que não permite a obtenção de um “*fingerprint*” das amostras. Dessa forma o presente trabalho teve como objetivo o uso da espectrofluorimetria síncrona e das análises quimiométricas PLS e PLS-DA para prever algumas propriedades físico-químicas em biodiesel (B100) e querosene de aviação (QAV) e também classificar amostras contaminadas de QAV.

1.2. Espectrofluorimetria

A espectrofluorimetria é uma técnica analítica que vem crescendo nos últimos anos, essa técnica consiste na obtenção de um espectro de emissão molecular, obtidos por moléculas com características particulares, relacionando uma intensidade de sinal com o comprimento de onda de emissão, a partir da excitação da amostra em comprimentos de onda específicos, podendo ser realizada de duas formas: convencional ou sincronizada.

Na espectrofluorimetria convencional, um espectro de emissão é obtido fazendo-se a varredura do monocromador de emissão enquanto a amostra é excitada a um comprimento de onda de excitação fixo [10-11]. Mesmo sendo bastante aplicada em várias situações analíticas, a espectrofluorimetria convencional tem algumas limitações quando se trata de amostras complexas. As várias espécies presentes na amostra podem causar interferências como auto absorção e sobreposição de sinais [12], tornando-se então, pouco viável para caracterizar tais amostras. Uma alternativa para a análise destas amostras é utilizar a espectrofluorimetria sincronizada.

1.2.1. Espectrofluorimetria Síncrona

A espectrofluorimetria de varredura sincronizada consiste em se efetuar uma varredura simultânea nos monocromadores de emissão e excitação, mantendo-se uma diferença de comprimento de onda fixa entre eles ($\Delta\lambda$), registrando-se o espectro obtido da intensidade de sinal em função do comprimento de onda de emissão, relativa à diferença aplicada [12].

Dependendo da taxa de varredura dos monocromadores, três tipos de espectrofluorimetria de varredura sincronizada são possíveis [13]. A primeira é a espectrofluorimetria de varredura sincronizada com comprimento de onda constante, onde a taxa de varredura é constante para ambos os monocromadores e uma diferença entre λ_{ex} e λ_{em} ($\Delta\lambda$) é mantida constante entre eles. Esta é a técnica mais usada e referida simplesmente como EVS. A segunda é a EVS com energia constante ao qual os λ_{ex} e λ_{em} são variados simultaneamente mantendo a diferença de frequência constante entre eles. A terceira e última é a EVS com ângulo variável, onde os λ_{ex} e λ_{em} podem ser variados simultaneamente, mas a diferentes taxas de varredura. Essas diferentes taxas permitem a construção de planos com ângulos entre 45 e 90° para o eixo x de excitação ao longo de todo o espectro.

As principais vantagens da EVS são a formação de bandas mais estreitas e definidas, a contração da faixa espectral e a possibilidade de se selecionar um $\Delta\lambda$ que aumente apenas a intensidade de emissão da espécie de interesse, o que não é possível no método convencional. Para entender melhor esta técnica, é necessário apresentar as equações relacionadas com o fenômeno de fluorescência.

Um sinal sincronizado é observado apenas quando o $\Delta\lambda$ combina o intervalo entre uma banda de emissão e uma banda de absorção, como apresentado na Figura 1. O $\Delta\lambda$ selecionado só permite a combinação de uma banda de excitação com uma banda de emissão, gerando apenas um pico sincronizado.

Na Figura 2 são apresentados os espectros de emissão e excitação do tetraceno. Há uma pequena diferença entre os comprimentos de onda (3 nm) das bandas correspondentes às transições π - π^* , chamado deslocamento de Stokes. Quando se utiliza um $\Delta\lambda$ correspondente a 3 nm para traçar o espectro sincronizado, o resultado será um único pico, correspondente à sobreposição destas bandas de emissão e excitação. Nenhum outro pico sincronizado é observado, pois não é possível outra combinação entre as bandas considerando o intervalo de 3 nm.

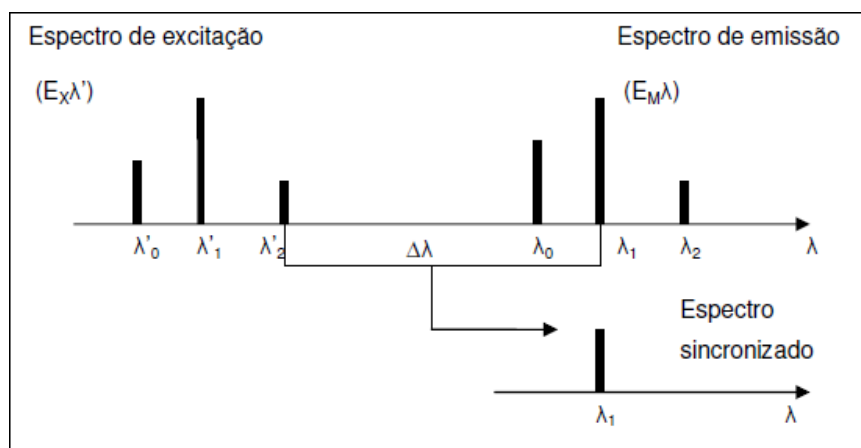


Figura 1. Efeito de simplificação espectral da espectrofluorimetria sincronizada.

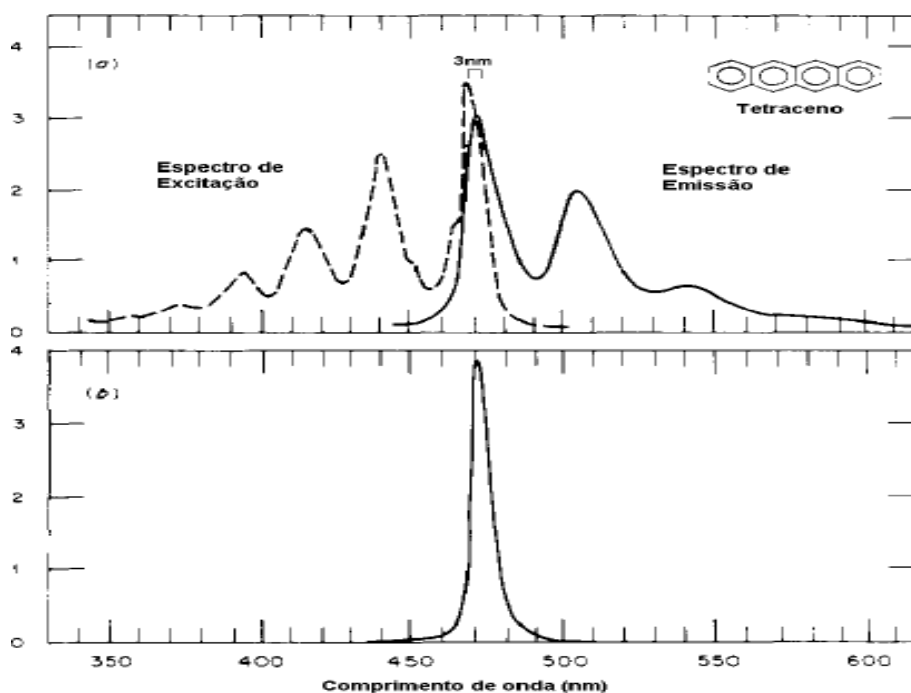


Figura 2. (a) Espectros de excitação e emissão do tetraceno. (b) Sinal de fluorescência sincronizada do tetraceno. *Fonte [14].*

A aplicabilidade da espectrofluorimetria à análise de derivados de petróleo se deve à presença, nesses produtos, de HPAs que são altamente fluorescentes. Esta técnica analítica pode ser útil não só na caracterização de misturas, mas também na detecção de agentes indesejáveis. Dos espectros de emissão molecular são extraídas informações qualitativas e quantitativas, sendo que a intensidade da radiação emitida pela amostra é proporcional à concentração de uma determinada espécie.

O método da fluorescência sincronizada, vem sendo empregado em análises de compostos como diesel, querosene e demais derivados de petróleo [15], sua aplicação visa detectar e quantificar a adulteração de diesel com querosene [16], além de permitir a determinação de hidrocarbonetos aromáticos policíclicos em amostras de água e óleos lubrificantes [17–18]. Técnicas espectrofluorimétricas são amplamente empregadas na caracterização e na detecção de adulterações em derivados de petróleo, como evidenciado por diversos estudos disponíveis na literatura (Tabela 1).

Tabela 1. Relação de estudos usando espectrofluorimetria sincronizada e quimiometria para otimização de ensaios para biodiesel e diesel.

Técnica Quimiométrica	Técnica Analítica	Amostra	Tipo de Ensaio	Referência
PARAFAC-PLSR (MCR-ALS)-PLSR TBF1-PLSR	EEM	Diesel	Contaminação de água do mar por óleo	19
PLS e PCA	EFS	Biodiesel	Estabilidade Oxidativa	20
--	EEMF e EEMSS	Diesel	Adulteração em combustível	15
--	SFS, EEMF e EEMSS	Gasolina, diesel, querosene	Contaminação de gasolina e diesel por querosene	16
--	SFS	Gasolina e diesel	Adulteração de Gasolina e Diesel por Querosene	17
PCR PLSR	SFS	Gasolina	Calibração multivariada	21
PACs	SFS	Diesel, gasolina e querosene	Quantificação de óleos em motores e o efeito de solventes nesses óleos	22
SUVF MCR-ALS PLS	SFS	Óleos de petróleo	Previsão da origem do petróleo bruto (tipo de rocha)	23
PCA	SFC	Óleos brutos e refinados	Medição da similaridade de dois espectros	24
PCA e PLS	EFC	Diesel	Contaminação de diesel por óleo residual	25

1.3. Quimiometria

Segundo a Sociedade Internacional de Quimiometria (ICS, *International Chemometrics Society*), esta é definida como a “disciplina química que utiliza métodos matemáticos e estatísticos, entre outros, empregando uma lógica formal para: definir ou selecionar condições ótimas para procedimentos de análise e experimentos; obter o máximo de informação química relevante a partir da análise de dados químicos”.

Genericamente, pode-se dividir a quimiometria em três áreas principais, planejamento e otimização de experimentos, reconhecimento de padrões e calibração multivariada. Os métodos de reconhecimento de padrões podem ser utilizados para análise exploratória de dados, classificação de amostras e resolução de curvas [26].

1.3.1. Métodos de Classificação

Existem vários modelos quimiométricos para a discriminação ou classificação supervisionados de amostras, onde as características das amostras são informadas ao modelo, e não supervisionados onde o modelo não contém nenhuma informação sobre as características da amostra.

1.3.1.1 Métodos não supervisionados

Dentre os métodos não supervisionados a análise de componentes principais (PCA) tem como objetivo a redução da dimensão dos dados através de combinações lineares de suas variáveis originais. A ideia da PCA é exemplificada na Figura 3.

$$\begin{array}{c} m \\ \square \\ X \\ n \end{array} = \begin{array}{c} a \\ \square \\ T \\ n \end{array} \begin{array}{c} a \\ \square \\ P' \\ m \end{array} + \begin{array}{c} m \\ \square \\ E \\ n \end{array}$$

Figura 3. Esquema de uma PCA. Fonte [27]

Os dados instrumentais são representados em X com n linhas (amostras) e m colunas (variáveis). Esta matriz é decomposta em uma multiplicação entre as matrizes T , contendo os escores, e a matriz P , contendo os pesos. A matriz E contém os resíduos não explicados pela PCA e deve ser somada à $T \times P'$ para que a igualdade se mantenha.

A matriz X é decomposta em uma multiplicação de vetores de dimensionalidade 1, com t contendo o escore e p contendo os pesos, segundo a Equação 1.

$$X = t_1 * p_1 + t_2 * p_2 + \dots + t_m * p_m \quad \text{Equação (1)}$$

A decomposição de X ocorre em a componentes principais (PC), sendo a componente principal uma nova variável gerada a partir da combinação linear de variáveis originais com alta correlação e cada componente principal é construída de modo a explicar a máxima variabilidade dos dados. O vetor peso contém a contribuição de cada variável para as componentes principais e geometricamente representa os cossenos dos ângulos entre a respectiva componente principal e os eixos definidos pelas variáveis originais. O vetor escore contém a projeção da variável original na nova componente principal e o conjunto destes vetores formam as matrizes T e P [28].

Na maioria das vezes a PCA é capaz de representar o conjunto de dados em poucas componentes principais, permitindo a visualização de similaridades e diferenças entre as amostras por meio da análise dos gráficos de escores. As variáveis responsáveis pela similaridade ou diferença entre as amostras podem ser observadas no gráfico dos pesos versus variáveis.

A Figura 4 mostra um exemplo da PCA para dados espectrais (dados contínuos). A matriz X (contendo os espectros de n amostras e m variáveis) é decomposta nas matrizes T (n, A), contendo os escores, e na matriz P (A, m) em a PCs. A relação entre as amostras e os possíveis agrupamentos são observados em um gráfico dos escores das componentes principais um contra o outro.

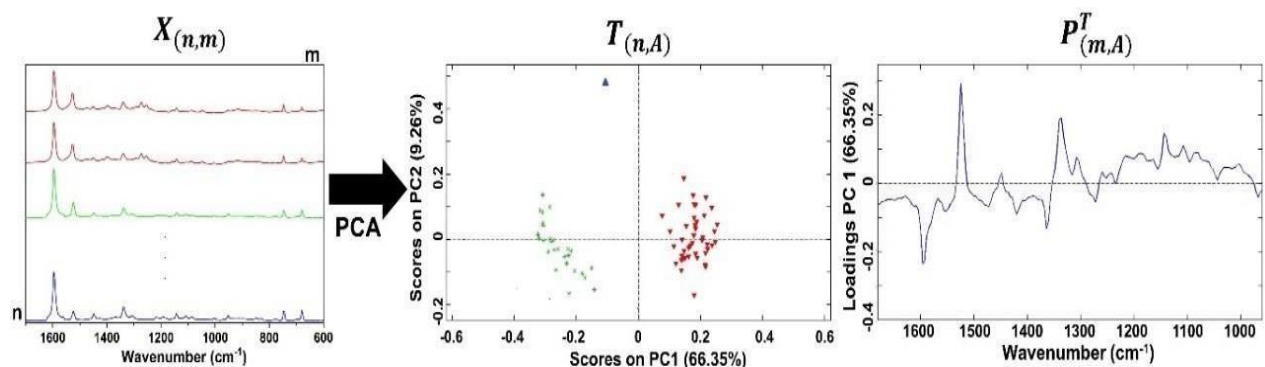


Figura 4. PCA para dados espectrais adaptada de [27].

Na Figura 4 observa-se um grupo de amostras no lado positivo da *PC1* (vermelho) e outro grupo no lado negativo da *PC1* (verde), indicando que a primeira componente é formada pela combinação de números de onda nos quais estes grupos de amostras apresentam diferenças. Ainda no gráfico de escores observa-se uma amostra com maior valor de escore na *PC2* (azul), distante das demais amostras. Neste caso, a *PC2* traz informação das características dessa amostra, pois são essas características que a diferem dos dois grupos descritos por *PC1*.

Para o entendimento das diferenças observadas no gráfico de escores é necessário a análise dos pesos (ou *loadings*) das *PCs*. No gráfico de pesos de uma *PC* é mostrado quais variáveis contribuem para caracterizar os agrupamentos observados no gráfico de escores. Observe na figura que as bandas espectrais que caracterizam as amostras no lado positivo da *PC1* são as bandas próximas a 1500 e 1350 cm^{-1} , enquanto as amostras no lado negativo da *PC1*, a região espectral com maior peso é em 1600 cm^{-1} [29].

1.3.1.2 Análise Discriminante pelo Método dos Quadrados Mínimos Parciais PLS-DA

A análise discriminante pelo método dos quadrados mínimos parciais (*Partial Least Squares Discriminant Analysis*, PLS-DA) é um método multivariado de classificação não a uma determinada classe, respectivamente. Se \mathbf{y} é um vetor, temos um modelo PLS1-DA e se \mathbf{Y} é uma matriz, temos a PLS2-DA, prevendo mais de duas classes simultaneamente. Neste caso, o número de colunas da matriz \mathbf{Y} é igual ao número de classes. Quando são avaliadas mais de duas classes, não obrigatoriamente deve-se usar o modelo PLS2-DA, pode-se também usar o PLS1-DA, prevendo uma classe de cada vez.

Os valores previstos pelo modelo PLS-DA não são idealmente iguais a 1 ou 0, mas valores reais que devem ser aproximados desses números. Como não são exatos, um limite (*threshold*) é estabelecido para delimitar as classes e se o valor previsto está acima deste limite, a amostra é considerada pertencente à classe em questão. Este “*threshold*” pode ser calculado baseado na estatística bayesiana, estimado de modo a minimizar os erros de classificação, assumindo que a variância de \mathbf{y} segue uma distribuição semelhante à que será observada para as futuras amostras.

Quando a composição das amostras não é controlada, é necessário algum método sistemático para separar as amostras mais representativas do conjunto para a construção do modelo. Esse objetivo é alcançado utilizando diferentes algoritmos tais como o de Kennard- Stone, que realiza a seleção de amostras com maior variabilidade do conjunto total. Para se obter boa representatividade de amostras de todas as classes, o algoritmo deve ser aplicado para cada classe

individualmente.

Os blocos X e Y são decompostos simultaneamente em uma soma de variáveis latentes (VL). As VL são combinações lineares das variáveis originais, que descrevem simultaneamente maior parte das informações nos dois blocos de dados. Para a construção de um bom modelo preditivo é essencial a escolha do número correto de VL , que deve ser realizada por meio da validação cruzada. A estratégia mais comum de validação cruzada para conjuntos de amostras pequenos é o “*leave-one-out*”, no qual treina-se o modelo com $n-1$ dados, e o avalia utilizando a observação restante. O procedimento é repetido n vezes, excluindo em cada momento uma observação diferente. Quando o conjunto amostral é maior, são utilizados os blocos contínuos, os subconjuntos aleatórios ou “*venetian blinds*”.

Em geral, no processo de validação cruzada, vários modelos prévios são construídos, nos quais um bloco de amostras do treinamento é retirado e, com as restantes, constrói-se o modelo e estima-se a classe do bloco de amostras retirado, e isso é repetido para todas as amostras. A média dos erros de previsão é calculada para cada número de VL e o que apresentar o menor erro será o número de VL sugerido para o modelo. Este erro é chamado de erro de classificação de validação cruzada (*cross validation classification error, CVCE*).

A detecção de amostras anômalas é realizada a partir da análise do gráfico dos valores de Hotelling (T^2) e os resíduos (Q) para todas as amostras. A estatística T^2 é dada pela soma dos escores normalizados ao quadrado, e é uma medida da variação de cada amostra no modelo. São excluídas aquelas amostras que apresentam altos valores de T^2 e Q , simultaneamente, com 95% de confiança, respeitando o limite de remoção de, no máximo, $2/9$ do número total de amostras.

Depois de construído e validado com um conjunto de amostras externas, ou seja, amostras que não foram utilizadas em seu treinamento, o modelo de classificação supervisionada deve ser avaliado quanto aos seus parâmetros de qualidade. A sensibilidade é a habilidade do método em detectar amostras verdadeiramente positivas como positivas enquanto a taxa de sensibilidade (SEN) corresponde à probabilidade de um método classificar como positiva uma amostra sabidamente positiva. De maneira análoga, a especificidade ou seletividade corresponde à habilidade do método em detectar amostras verdadeiramente negativas como negativas, sendo a taxa de especificidade (SPE) a probabilidade de o método classificar como negativa uma amostra sabidamente negativa [30]. A taxa de falso-positivos (TFP) é definida como a razão entre o número de resultados falso-positivos (FP) e a soma deste com o número de resultados verdadeiro-negativos (VN) e a taxa de falso-negativos (TFN) é a razão entre o número de resultados falso-negativos (FN) e a soma de FN com o número de resultados verdadeiro-positivos (TP). Eficiência (EFF), também denominada de acurácia, é um valor global único para medir o desempenho do modelo de

classificação [31].

Esses parâmetros são calculados separadamente para os conjuntos de treinamento e teste, conforme as Equações 2 a 6.

$$SEN = \frac{VP}{(VP+FN)} \quad (\text{Equação 2})$$

$$SPE = \frac{VN}{(VN+FP)} \quad (\text{Equação 3})$$

$$EFE = \frac{(VN+VP)}{(VN+FP+VP+FN)} \quad (\text{Equação 4})$$

$$TFP = \frac{FP}{(FP+VN)} \quad (\text{Equação 5})$$

$$TFN = \frac{FN}{(FN+VP)} \quad (\text{Equação 6})$$

1.3.2. Modelos de Calibração

A análise da regressão ou modelos de calibração é a metodologia estatística usada para prever valores de uma ou mais variáveis dependentes a partir de valores de um conjunto de variáveis independentes.

Um dos objetivos da análise da regressão é o desenvolvimento de uma equação que deve permitir ao investigador prever a resposta para os valores das variáveis preditoras. Para isso, é necessário ajustar um modelo aos dados [32]. Um método bastante utilizado é a regressão por mínimos quadrados parciais (PLS, *Partial Least Squares*).

Dentre os métodos de calibração multivariada a Regressão por Mínimos Quadrados Parciais (PLS) se destaca como sendo a mais usada, que é realizada através de um pequeno conjunto de fatores ortogonais, obtidos das variáveis preditoras. O modelo PLS é desenvolvido a partir de um conjunto de treinamento de n amostras, com k variáveis x , representadas por x_k ($k= 1, 2, \dots, n$) e m variáveis y , denotadas por y_m ($m = 1, 2, \dots, m$). Este conjunto de treinamento forma as duas matrizes X e Y , de dimensões $(n * k)$ e $(n * m)$, respectivamente. A matriz X é formada pelas variáveis independentes, ou preditoras, e a matriz Y é formada pelas variáveis dependentes.

Antes de se começar a fazer as análises quimiométricas, as variáveis são transformadas de forma que sua distribuição seja simétrica e todas passem a ter a mesma importância. Uma forma de se fazer isso é através do autoescalamento, entretanto esse pré-processamento nem sempre é conveniente. Espectros de absorvância, por exemplo, não devem ser autoescalados, pois com este tratamento os ruídos adquirem importância equivalente à das bandas mais intensas.

As variáveis x_k são transformadas em novas variáveis, os X -escores (t_a), por combinações lineares com os coeficientes ponderados w_{ka} , de acordo com a Equação 7.

$$t_{ia} = \sum w_{ka} x_{ik} \quad (\text{Equação 7})$$

Os X -escores são ortogonais e são bons preditores de Y , quando combinados com os pesos c_{ma} , gerando pequenos resíduos f_{im} , como apresentado na Equação 8.

$$y_{im} = \sum c_{ma} t_{ia} + f_{im} \quad (\text{Equação 8})$$

Por outro lado, Y também pode ser descrito como uma combinação de Y -escores u_{ia} e pesos c_{ma} , gerando pequenos resíduos g_{im} , como na Equação 9.

$$y_{im} = \sum u_{ia} c_{ma} + g_{im} \quad (\text{Equação 9})$$

A Equação 8 pode ser reescrita como um modelo de regressão múltipla, de acordo com a Equação 10:

$$y_{im} = \sum c_{ma} \sum w_{ka} x_{ik} + f_{im} = \sum b_{mk} x_{ik} + f_{im} \quad (\text{Equação 10})$$

onde os parâmetros “ b ” são os coeficientes da regressão.

Pode-se concluir, então, que no modelo de regressão PLS novas variáveis (variáveis latentes, ta) são formadas como combinações lineares das antigas variáveis x e utilizadas para prever os valores de Y .

Os valores dos escores e pesos podem ser calculados pelo algoritmo *NIPALS* e contêm informações sobre as amostras e as variáveis, respectivamente.

A parte dos dados que não é explicada pelo modelo, os resíduos, também trazem informação sobre o sistema. Por exemplo, grandes valores de resíduos em Y indicam que o modelo não está bem ajustado e os resíduos em X podem fornecer informações como identificação de amostras anômalas (*outliers*).

Após a calibração, um procedimento de validação deve ser utilizado para se verificar a validade do modelo. Normalmente, validação cruzada “*leave-one-out*” é utilizada. Este método também é utilizado para a definição do número de variáveis latentes (ou fatores) que descrevem o modelo.

O número de fatores a ser utilizado para a construção do modelo é aquele que minimiza a raiz quadrada dos erros médios quadrados da validação cruzada (*RMSECV*, *Root Mean Square Error of Cross Validation*”), definido pela Equação 11.

$$RMSECV = \sqrt{\frac{\sum (y_p - y_r)^2}{n}} \quad (\text{Equação 11})$$

onde:

y_p = valor de y previsto

y_r = valor real de y

n = número de amostras

A qualidade do modelo é avaliada pelos erros de calibração (*RMSEC*) e de previsão (*RMSEP*), obtidos a partir da Equação 11 para as amostras dos conjuntos de treinamento e de validação externa, respectivamente. O *RMSEC* representa a dispersão dos pontos no modelo de calibração e, quanto menor o valor de *RMSEP*, maior a exatidão obtida no resultado de previsão fornecido pelo modelo de calibração [21].

Outros parâmetros a serem analisados são o coeficiente de correlação (*r*) entre os valores reais e previstos e o erro relativo (*ER*) para a previsão, dado pela Equação 12.

$$ER(\%) = \frac{y_p - y_r}{y_r} \cdot 100 \quad \text{Equação (12)}$$

onde:

y_r = valor real de y

y_p = valor de y previsto

1.4. Otimização de variáveis através do Planejamento de Face Centrada

O conhecimento sobre quais as variáveis (fatores) afeta um sistema em estudo, bem como a extensão desta influência, são importantes para a melhoria do processo de uma maneira geral. Na abordagem multivariada, são conduzidos o menor número de experimentos necessários para extrair o máximo de informação dos dados coletados para avaliar e/ou otimizar um produto ou processo. Para isto, todos os fatores são alterados simultaneamente em um conjunto de experimentos pré-determinados e as respostas são utilizadas para a construção de modelos matemáticos que descrevem o seu comportamento dentro do domínio experimental investigado [33].

O planejamento composto central (CCD, *Central Composite Design*) é uma ampliação natural do planejamento fatorial com ponto central (Figura 5A) no qual são adicionados pontos axiais que fornecem os níveis adicionais para o cálculo dos coeficientes do modelo quadrático.

A distância dos pontos axiais ao centro do planejamento é igual $a \pm \alpha$, sendo que α pode variar de 1 até \sqrt{k} , onde k é o número de fatores. O valor de α depende das propriedades desejáveis para o planejamento e da possibilidade de realização de experimentos ao longo do domínio experimental.

Quando $\alpha = \sqrt{k}$, temos um planejamento em estrela, com os pontos em uma distribuição octogonal. Quando $\alpha = 1$, os pontos axiais estão localizados sobre as arestas do quadrado para um planejamento fatorial 22 (Figura 5B) e sobre as faces de um cubo para um planejamento 23 e é chamado planejamento composto central de face centrada.

Quando $\alpha = \sqrt{k}$, temos um planejamento em estrela, com os pontos em uma distribuição octogonal. Quando $\alpha = 1$, os pontos axiais estão localizados sobre as arestas do quadrado para um planejamento fatorial 22 (Figura 5C) e sobre as faces de um cubo para um planejamento 23 e é chamado planejamento composto central de face centrada [34].

Levando em consideração a importância da otimização dos ensaios de controle de qualidade para o querosene de aviação e o biodiesel, bem como a diminuição dos custos dessas análises, o uso de técnicas multivariadas em conjunto com a calibração multivariada *PLS* possibilitará que, com um único ensaio, sejam previstos os valores de vários parâmetros físico-químicos previstos pelas Portarias nº 5 e 37 da ANP, reduzindo custos e tempo de análise.

Os resultados obtidos por esses modelos podem ser usados como “*screening*”, selecionando amostras com resultados próximos aos limites das Portarias e Resoluções, para a realização de ensaios nos equipamentos específicos, agilizando o processo analítico, além de permitir a identificação rápida de possíveis contaminações nos combustíveis, o que é de extrema importância para evitar futuros acidentes causados por problemas nos motores e turbinas das aeronaves.

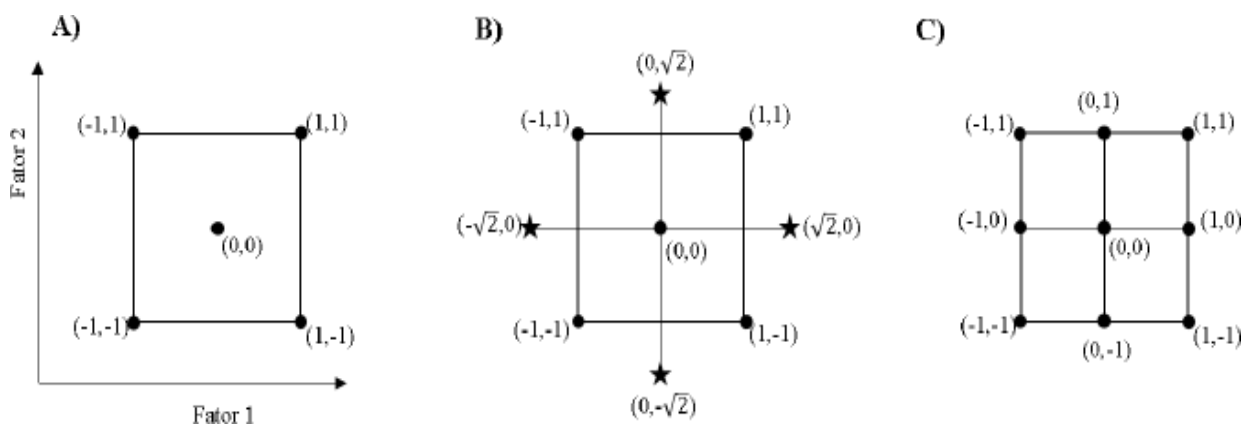


Figura 5. Planejamento fatorial com ponto central (A), planejamento composto central 22 em estrela (B) e planejamento composto central 22 de face centrada (C) fonte [34].

2. OBJETIVOS

Desenvolver novas metodologias analíticas com o auxílio de ferramentas quimiométricas para a caracterização e determinação de parâmetros físico-químicos relacionados às Portarias da ANP. Para os seguintes objetivos específicos:

- Construir um modelo quimiométrico de regressão, baseado na espectrofluorimetria sincronizada, para a previsão da Estabilidade Oxidativa do biodiesel (B100) comercial;

- Construir modelos quimiométricos de classificação, baseados na espectrofluorimetria sincronizada para a previsão do Ponto de Congelamento do querosene de aviação;

- Construir modelos quimiométricos de classificação e regressão, baseados na espectrofluorimetria, para a identificação de contaminações do querosene de aviação por outros derivados de petróleo como gasolinas automotiva e de aviação e querosenes iluminante e mineral.

3.0 MODELO QUIMIOMETRICO PARA PREVISÃO DA ESTABILIDADE OXIDATIVA DO BIODIESEL

3.1 Introdução

De acordo com o relatório da GlobeNewswire, em 2021, o mercado global de biodiesel foi avaliado em US\$ 32,1 bilhões e espera-se que cresça a uma taxa de crescimento anual composta de 10,0% de 2022 a 2030. Entre os biocombustíveis, o biodiesel surge como uma alternativa ao uso do diesel, pois emite menos gases de efeito estufa e pode ser utilizado em motores a diesel existentes sem a necessidade de modificações [35].

O biodiesel é um combustível que possui uma queima limpa, proveniente de ácidos graxos de cadeia longa, podendo ser derivado de óleos vegetais e/ou gorduras animais. O principal componente da matéria-prima utilizada para produzir biodiesel é composto principalmente por triglicerídeos e componentes em proporções menores, como água, esteróis, ácidos graxos livres, fosfolipídios e impurezas, quando se trata de óleo residual [36].

Devido às características similares, o biodiesel e o diesel convencional são misturados e utilizados como combustível comercial. Estudos mostram que essas misturas podem ser realizadas até certo nível sem modificações ou danos aos equipamentos originalmente planejados para utilizar diesel convencional [37]. As misturas são chamadas de *BX*, onde *X* é a fração de biodiesel presente na mistura.

Para atender às demandas globais de redução das emissões de gases de efeito estufa, como a Rio 92, o Protocolo de Kyoto, a Rio+20, entre outros, o Brasil estabeleceu a Lei de Resolução nº 11.097, que iniciou a introdução do biodiesel na cadeia de combustíveis do Brasil, exigindo inicialmente a adição de 2% de biodiesel ao diesel fóssil em 2005 [39-40], chegando a 14% em 2024 com expectativa de chegar a 20% em 2030 [38].

A produção de biodiesel no território brasileiro acompanhou a crescente adição e saltou de pouco menos de 3 milhões de metros cúbicos em 2013 [41] para mais de 7,5 milhões de metros cúbicos em 2023, tendo o óleo de soja como matéria prima com maior participação na fabricação, atingindo quase 70% do volume de matéria prima processada, seguido por uma mistura de outros materiais graxos [42].

O biodiesel se mostra mais suscetível a degradação principalmente por processos oxidativos, quando comparado ao diesel [43], isto se deve ao grau de insaturação dos alquilésteres presentes nas moléculas de biodiesel [44], e com a posição das duplas ligações na cadeia carbônica. Quanto maior o número de insaturações, mais susceptível está a molécula à degradação tanto térmica quanto oxidativa, formando produtos insolúveis que ocasionam problemas de formação

de depósitos e entupimento do sistema de injeção de combustível do motor [45]. O biodiesel quando oxidado apresenta características indesejáveis, como elevada acidez, aumento da viscosidade e tendência à formação de sólidos. A extensão destes efeitos depende das concentrações destes compostos, que podem variar de acordo com a matéria prima utilizada na sua produção, além das condições de armazenamento (umidade, calor e exposição à luz).

Desta forma, o ensaio de estabilidade à oxidação ou estabilidade oxidativa, é um dos mais importantes para a qualidade do biodiesel, pois os produtos de oxidação do biodiesel não prejudicam apenas as características do combustível, mas também comprometem a durabilidade do motor [46-47].

Portanto, para que o biodiesel seja usado como combustível, é necessário garantir sua qualidade por meio da adequação de seus parâmetros físico-químicos, como viscosidade e estabilidade oxidativa, entre outros. A regulamentação, supervisão e monitoramento dos combustíveis e biocombustíveis no Brasil são responsabilidade da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), que, por meio da resolução nº 920/2023, especifica os testes, metodologias e valores limites que o biodiesel comercializado deve atender [3].

As metodologias de referência para a determinação da estabilidade oxidativa são especificadas pelas normas do Comitê Europeu de Normalização EN 14112 e EN 15751, [48-49] sendo que a primeira apresenta os procedimentos a serem realizados com biodiesel puro, chamado de B100, enquanto a segunda apresenta os procedimentos para misturas de diesel fóssil com biodiesel, chamadas de blends. O procedimento de referência para análise das amostras do presente trabalho é aquele encontrado na norma EN 14112 [48].

A EN 14112 [48] descreve o método convencional usado para determinar a estabilidade oxidativa baseado na aceleração do processo de envelhecimento da amostra pelo aumento da temperatura e passagem de corrente de ar contínuo sobre ela. No ensaio uma amostra de biodiesel é aquecida a 110 °C e misturada numa corrente de ar filtrado, que leva os produtos da oxidação para uma câmara com água destilada, que tem sua condutividade elétrica medida continuamente até um súbito aumento de condutividade da solução absorvedora dos produtos de degradação. O tempo decorrido até este súbito aumento da condutividade é chamado de Período de Indução. Acima deste ponto é constatado um rápido aumento da taxa de oxidação, da absorção de oxigênio e de formação de voláteis, o Período de Indução pode variar entre 8 e 48 horas [47].

No Brasil, a RANP 920/2023 estabelece que 13 horas é o tempo mínimo para que o biodiesel seja considerado conforme. No entanto, a EN 14112-2016 admite que o ensaio possa durar até 48 horas, a depender da amostra [50].

A literatura científica traz, alguns estudos utilizando diferentes técnicas analíticas juntamente com análise quimiométrica para obtenção de métodos alternativos que auxiliam os ensaios padronizados na determinação das propriedades físico-químicas dos combustíveis. A maioria dos trabalhos publicados para previsão da Estabilidade Oxidativa em biodiesel faz uso da Espectroscopia no Infravermelho próximo (FT-NIR) e médio (FT-MIR) [51 – 54].

Velvarská et al. [51] desenvolveram um modelo quimiométrico baseado na FT-NIR para determinação da estabilidade oxidativa de 133 amostras preparadas em laboratório, entre diesel, biodiesel e misturas de diesel/biodiesel. O biodiesel utilizado continha ésteres metílicos de óleo de colza, girassol ou óleo de cozinha usado e os valores de referência de estabilidade oxidativa foram determinados pelo aparelho PetroOxy (Petrotest Instruments, Alemanha), variando na faixa de 20,66 a 126,43 min. O coeficiente de correlação entre a medida obtida pelo método de referência e aquela prevista por modelos de PLS utilizando validação cruzada foi de 0,96.

Canha et al. [52] obtiveram modelos de PLS baseados em espectros de FT-NIR para determinação da estabilidade oxidativa de 247 amostras de biodiesel, sintetizadas e comerciais, contendo ou não antioxidantes, envelhecidas artificialmente em diferentes condições de luminosidade e temperatura. Para as amostras sem antioxidantes o modelo apresentou um *RMSEP* de 0.6 h, em uma faixa de períodos de indução de 0,66 a 17,7 h, e o modelo para amostras contendo antioxidantes apresentou um *RMSEP* de 1,26 h para uma faixa de 0,43 a 22,1 h.

Já Varghese, Saeed e Rutt [53] utilizaram FT-NIR para obtenção de modelos PLS com 11 amostras de biodiesel sintetizado em laboratório a partir de diferentes matérias primas como azeite e óleos de colza, soja, girassol e milho, dentre outros, previamente analisadas pelo método Rancimat para previsão da estabilidade oxidativa, na faixa de 0,67 a 3,97 h.

De Lira et al. [52] compararam o uso de FT-NIR e FT-MIR na construção de modelos baseados em mínimos quadrados parciais e regressão linear múltipla, com e sem seleção de variáveis e diferentes processamentos de espectros, no monitoramento da estabilidade oxidativa, número de acidez e teor de água do biodiesel. No caso dos modelos para estabilidade oxidativa, foram utilizados para construção do modelo um conjunto de 72 amostras provenientes de várias regiões, que foram oxidadas de forma acelerada por seis semanas conforme a ASTM D4625 a 43 °C em ar atmosférico, com um período de indução variando de 1,5 a 8,7 h. Foram obtidos valores de *RMSEP* de 0,48 e 0,73 para os modelos baseados no NIR e MIR, respectivamente.

Além da espectroscopia no infravermelho, outras técnicas analíticas de simples operação e de baixa complexidade como a Espectroscopia de Fluorescência convencional (EFC) e a Espectroscopia de Fluorescência Sincronizada (EFS), onde para a EFC os espectros de emissão são obtidos fazendo-se a varredura do monocromador de emissão (λ_{em}) enquanto a amostra é

excitada a um comprimento de onda de excitação (λ_{ex}) fixo, já na EFS o comprimento de onda do monocromador de excitação (λ_{ex}) e o comprimento de onda do monocromador de emissão (λ_{em}) são varridos simultaneamente mantendo-se uma diferença de comprimento de onda fixa entre eles, chamada de $\Delta\lambda$ [55-56].

A EFC e a EFS têm sido utilizadas para criação de modelos para classificação do biodiesel por tipo de matéria prima [57], e para estimar a porcentagem de biodiesel em blends de diesel/biodiesel [25], bem como outras propriedades como estabilidade oxidativa, índice de cetano, valor calorífico bruto e cor [58].

Meira et al. [20] determinaram a estabilidade oxidativa usando um modelo quimiométrico PLS juntamente com a espectrofluorimetria convencional a partir de um conjunto de dez amostras de biodiesel de soja preparadas a partir de uma mesma matriz do combustível, as amostras de biodiesel foram oxidadas de forma induzida e os valores das estabilidades oxidativas foram obtidos através do método Rancimat abrangendo uma faixa de período de indução de 0,0 a 5,90 h. Vinte e quatro espectros de emissão foram obtidos para cada amostra e o modelo PLS foi obtido a partir de uma matriz 3D de excitação-emissão, resultando em um coeficiente de determinação de 0,98, e *RMSECV* igual a 0,6h para predição do período de indução do biodiesel.

Insausti et al. [58] construíram modelos quimiométricos com base em espectros de fluorescência sincronizada e dados parâmetros de qualidade (porcentagem de biodiesel, índice de cetano, valor calorífico bruto e cor) obtidos por um analisador espectroscópico de infravermelho próximo e médio (PetroSpec) em misturas de diesel/biodiesel.

Em um outro trabalho, Insausti et al. [57] utilizaram dados de fluorescência sincronizada em diferenças constantes de comprimento de onda ($\Delta\lambda = \lambda_{em} - \lambda_{ex}$), que variou de 5 a 50 nm, juntamente com UV-Vis e FT-NIR, para obtenção de modelo de análise discriminante linear com algoritmo de projeções sucessivas (SPA-LDA) para classificação de biodiesel por matéria prima, utilizando amostras sintetizadas de biodiesel de milho, girassol e soja. Os espectros para cada $\Delta\lambda$ foram avaliados de acordo com o número de erros de treinamento e validação obtidos pelo SPA-LD, e os melhores valores de $\Delta\lambda$ foram 10 e 15 nm, para as quais todas as amostras de treinamento e validação foram classificadas corretamente.

Corgozinho e Barbeira [25] desenvolveram um método analítico simples e de baixo custo para a quantificação de biodiesel em misturas biodiesel-diesel, 68 amostras de óleo diesel puro e misturas com 2 % (v/v) de biodiesel, de diferentes origens e tipos, foram utilizadas combinando calibração multivariada à espectrofluorimetria sincronizada para um $\Delta\lambda$ de 40 nm, a determinação de biodiesel foi obtida em uma faixa que variou de 0,4 a 19,0% (v/v) com erros de predição inferiores a 0,35% (v/v).

Kumar e Mishra [59] utilizaram dados de fluorescência sincronizada para obtenção de modelos de PLS para determinação do teor de biodiesel de óleos de “*karanja*”, utilizando 21 blendas preparadas em laboratório a partir de biodiesel e diesel comercial, com teor variando de zero a 100%. O modelo de PLS obtido pelos autores apresentou coeficiente de determinação igual a 0,99, além de *RMSEC* igual a 0,91% e *RMSEP* igual a 0,86%.

Levando em consideração a importância do ensaio da estabilidade oxidativa para garantir a qualidade do B100, um modelo quimiométrico simples baseado em análise de Mínimos Quadrados Parciais (PLS) foi proposto para determinar a estabilidade oxidativa de amostras comerciais naturalmente degradadas de biodiesel, neste estudo, foram utilizados espectros de emissão de fluorescência sincronizada, uma vez que esta técnica possui alta sensibilidade para análise de amostras contendo múltiplos componentes [60] não necessita de nenhum preparo de amostra e a aquisição dos dados é extremamente rápida.

3.2 Materiais e Métodos

3.2.1. Amostras

Foram analisadas 65 amostras de biodiesel (B100) pelo Laboratório de Ensaio de Combustíveis do Departamento de Química da UFMG (LEC-UFMG). Essas amostras são provenientes de empresas e produtores de diferentes regiões do Brasil, e tiveram sua estabilidade oxidativa determinada pelo método Rancimat (EN14112), na faixa de 6 a 18 horas.

A Tabela 2 mostra a distribuição dos valores de Estabilidade Oxidativa para as amostras utilizadas na construção do modelo quimiométrico a partir dos espectros de fluorescência sincronizada, com o objetivo de ter uma distribuição uniforme de valores em toda a faixa estudada.

Tabela 2. Intervalo de valores de estabilidade oxidativa para as 65 amostras de B100 usadas nos modelos.

Faixa	No. de amostras	Faixa	No. de amostras
6,00 to 7,00	6	12,01 a 13,00	5
7,01 to 8,00	5	13,01 a 14,00	6
8,01 to 9,00	6	14,01 a 15,00	6
9,01 to 10,00	5	15,01 a 16,00	5
10,01 to 11,00	6	16,01 a 17,00	5
11,01 to 12,00	5	17,01 a 18,00	5

3.2.2. Instrumentação

As medidas espectrofluorimétricas foram realizadas usando um equipamento Shimadzu RF-5301PC equipado com uma lâmpada de xenônio. As amostras foram analisadas em uma cubeta de quartzo com quatro lados polidos e dimensões de 10 mm x 10 mm x 45 mm. O equipamento foi configurado com uma largura de fenda de excitação de 5 nm e largura de fenda de emissão de 10 nm, e resolução espectral de 1 nm. Os espectros foram coletados através da varredura simultânea dos monocromadores de excitação e emissão na faixa de 220-800 nm, com $\Delta\lambda$ constante entre eles, de 10 a 100 nm, em intervalos de 10 nm. As medidas de fluorescência foram realizadas em triplicata, e o espectrofluorímetro foi verificado diariamente usando uma amostra de controle.

3.2.3. Software e processamentos

O software MATLAB versão 7.10.0 - R2010A (The MathWorks, Natick, MA, EUA) e a ferramenta PLS Toolbox versão 5.2.2 (Eigenvector Technologies, Manson, WA, EUA) foram utilizados para conduzir as análises multivariadas.

Para a construção do modelo PLS, os dados foram organizados em uma matriz bidimensional \mathbf{X} (amostras $\times \lambda_{em}$) sendo que, para cada $\Delta\lambda$, os tratamentos de dados foram realizados separadamente. Os espectros de fluorescência foram pré-processados usando a primeira derivada com suavização usando o filtro Savitzky-Golay, e os dados foram centralizados na média. Parâmetros como grau polinomial e largura do filtro foram selecionados com base em uma análise do desenho composto central.

A escolha para determinar as melhores condições foi feita através dos erros de classificação de resposta nos conjuntos de treinamento, validação cruzada e teste nos modelos PLS. As amostras foram separadas em conjuntos de calibração e validação usando o algoritmo de Kennard-Stone, com 45 amostras no conjunto de calibração e 20 amostras no conjunto de validação, totalizando 65 amostras de biodiesel.

O bloco de variáveis independentes (\mathbf{X}) contendo intensidades de fluorescência, e o bloco de variáveis dependentes (\mathbf{y}) contendo os valores do período de indução obtidos no método de referência, foram decompostos simultaneamente em variáveis latentes para cada $\Delta\lambda$, escolhidos para minimizar o erro quadrático médio de calibração ($RMSEC$) e o erro quadrático médio de validação cruzada ($RMSECV$), estimados usando validação cruzada “*contiguous block*”. Amostras que simultaneamente apresentavam valores residuais Q e valores de T^2 de *Hotelling* acima do

intervalo de confiança de 95% foram consideradas outliers e removidas, respeitando um limite máximo de remoção de 22% das amostras por conjunto.

3.2.4. Estabilidade oxidativa

A Estabilidade Oxidativa das amostras foi determinada através do método convencional, EN14112, também chamado de Método Rancimat, no qual três gramas de biodiesel são colocados num vaso reacional selado aquecido à 110 °C, pelo qual passa um fluxo de ar purificado de 10L/h. Produtos de oxidação voláteis como aldeídos, cetonas, ésteres dentre outros, são carregados pela corrente de ar através de um duto e borbulhados num segundo frasco contendo água desmineralizada ou destilada, no qual os produtos de oxidação dissociam-se ionicamente, além de um condutivímetro acoplado a um dispositivo eletrônico. Durante a fase conhecida como período de indução, a quantidade de produtos de oxidação formada é pequena e, uma vez atingida a formação acelerada de produtos de oxidação, que são carregados pela corrente de ar, ocorre um aumento repentino da condutância da água no frasco. O dispositivo registra e indica quando houver este aumento, que sinaliza o fim do período de indução. O tempo, em horas, decorrido do início do ensaio até esse instante é anotado como o período de indução do biodiesel.

3.3. RESULTADOS E DISCUSSÃO

Para todos os $\Delta\lambda$ utilizados, diferentes bandas de emissão podem ser observadas entre as regiões de 400 a 700 nm, cujas intensidades e posições dependem do $\Delta\lambda$ utilizado, como mostra a Figura 6. O perfil do espectro de fluorescência das amostras se modifica conforme o $\Delta\lambda$ utilizado na aquisição dos dados, devido às diferentes combinações de comprimento de onda de excitação e emissão possíveis. Os espectros de fluorescência sincronizada do biodiesel são o resultado da sobreposição da fluorescência dos produtos de oxidação, como hidroperóxidos, aldeídos, cetonas, álcoois e alcanos [61].

Na utilização da fluorescência sincronizada, a emissão de um fluoróforo em um comprimento de onda λ_{em} apenas será detectada se obedecer à relação $\lambda_{em} = \lambda_{ex} + \Delta\lambda$, resultando em perfis de fluorescência diferentes em cada delta, com picos de diferentes intensidades e posições, contendo ora mais, ora menos informações químicas sobre as amostras. Isso mostra a necessidade de obtenção e avaliação de modelos quimiométricos individuais para cada delta, uma vez que as informações de cada delta são independentes entre si.

Durante a construção de modelos multivariados, o pré-processamento é uma etapa importante e deve ser otimizada. Combinar suavização com a primeira derivada tem como objetivo atenuar desvios de linha de base e ruídos espectrais e aumentar a variância da banda espectral, com um constante equilíbrio entre redução de ruído e distorção de sinal [62-63].

Neste estudo, os espectros de fluorescência síncrona das amostras de biodiesel foram pré-processados por meio da primeira derivada com alisamento usando o algoritmo de Savitzky-Golay que, após suavização e derivatização, foram centrados na média.

A escolha do alisamento com a primeira derivada baseia-se na premissa de equilibrar as necessidades de amenizar ruído e informação não sistemática sem dispensar informações relevantes para os modelos quimiométricos.

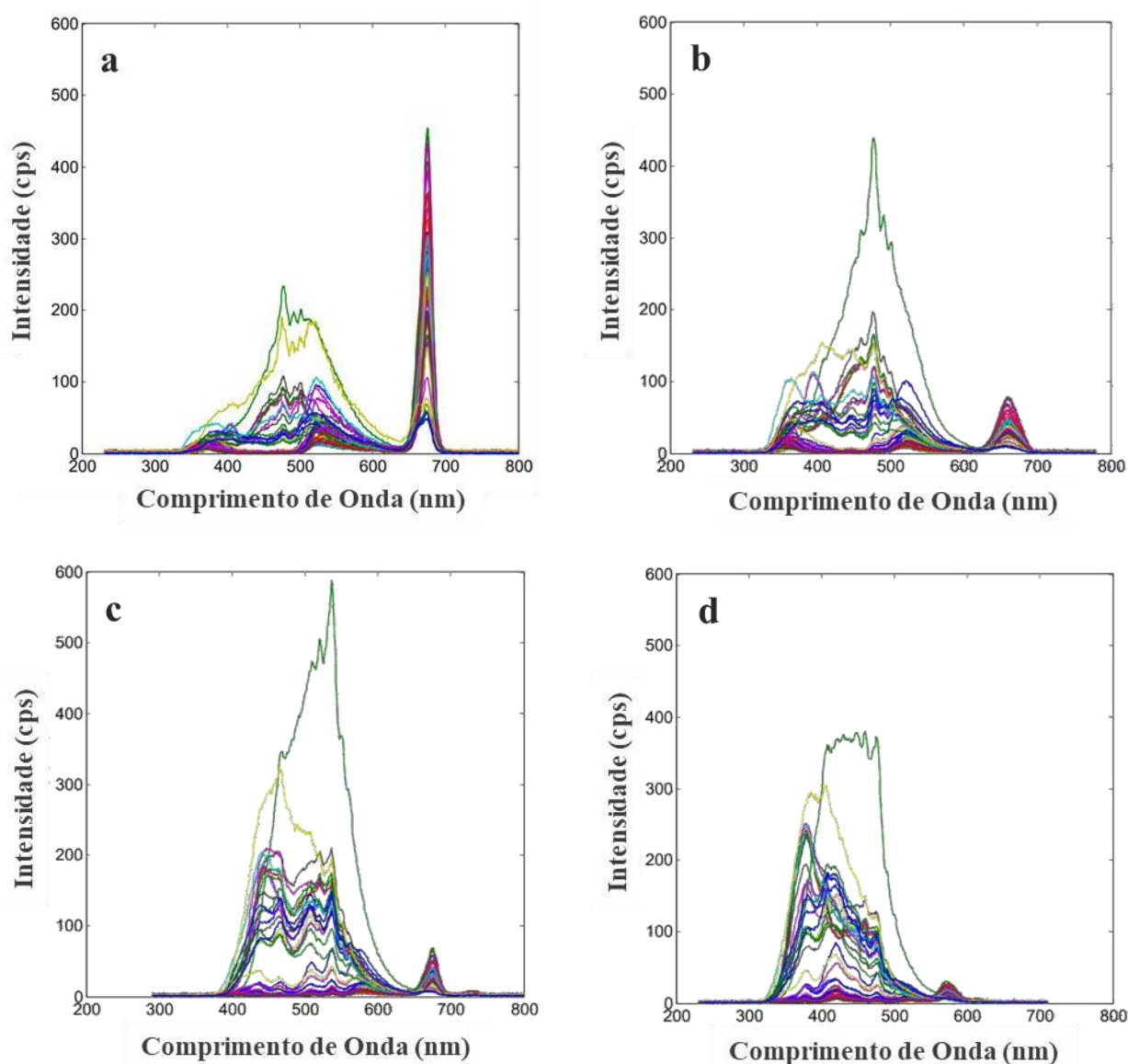


Figure 6. Espectros de fluorescência síncrona para todas as amostras B100 em diferentes valores de $\Delta\lambda$. a) 10; b) 30; c) 70; d) 100 nm.

Para determinar os parâmetros de suavização mais adequados de Savitzky-Golay (tamanho da janela e grau polinomial) para cada $\Delta\lambda$, foram realizados planejamentos de face centrada, visando os erros mínimos de classificação nos conjuntos de treinamento, validação cruzada e teste, além dos valores da razão entre erros em conjuntos de teste e validação. Esses planejamentos de face centrada foram utilizados para construir superfícies de resposta devido as limitações experimentais no fator de grau polinomial, no qual os níveis só podiam ser variados de 1 a 6 usando o software PLS ToolBox. Já o tamanho de janela pode ser variado sem limitações no software, no entanto todo pré-processamento pelo filtro de Savitsky-Golay deve considerar que, o uso de janelas muito grandes leva a um sinal processado muito ruidoso e o uso de janelas pequenas leva a um sinal com formato distorcido em relação ao sinal original [64]. Como ambas as condições são indesejadas, a escolha dos tamanhos de janela 11, 19 e 25 nm neste trabalho baseou-se na análise visual dos sinais processados.

A resposta desse tipo de planejamento é dada por uma técnica de otimização simultânea que consiste em encontrar os valores dos fatores que otimizem uma ou mais respostas, ou no mínimo as mantenha em faixas desejáveis, chamada de desejabilidade (Equação 13). Neste caso as variáveis otimizadas tiveram como critérios os maiores valores para o coeficiente de determinação (R^2) e menores valores para os erros $RMSECV$, $RMSEC$ e $RMSEP$. Cada resposta é então associada a uma função de desejabilidade parcial, que pode ter valores de 0 a 1, de acordo com o valor alvo para aquela resposta, dessa forma as funções de desejabilidade parciais são combinadas em um valor de desejabilidade global, D , obtida pela média geométrica das desejabilidades parciais, tornando-se um problema de otimização de apenas uma resposta.

$$D = (d_1 \times d_2 \times \dots \times d_n)^{1/n} \quad \text{Equação 13}$$

Desta forma, foram construídos e avaliados nove modelos para cada $\Delta\lambda$ estudado (10 a 100 nm), totalizando 90 modelos. Para exemplificar, na Tabela 3 estão apresentados os valores das figuras de méritos do melhor modelo para cada delta, obtidos como resposta para o planejamento experimental e a Figura 7 apresenta a superfície de resposta para o planejamento de face centrada para o $\Delta\lambda$ de 30 nm, indicando que a melhor combinação para este delta foi para uma janela de 19 pontos e polinômio de ordem 4.

O mesmo procedimento foi realizado para os demais $\Delta\lambda$ s e os melhores modelos, para cada $\Delta\lambda$, foram construídos utilizando o tamanho da janela e grau do polinômio indicados nas superfícies de resposta dos planejamentos de face centrada como mostrado na Tabela 3, o que deixa claro a importância do uso desta ferramenta para a escolha da melhor condição para a elaboração do melhor modelo.

Tabela 3. Tamanho da janela, grau polinomial e erros de classificação nos conjuntos de treinamento, validação cruzada e teste para o melhor modelo em cada $\Delta\lambda$ (nm).

$\Delta\lambda$	J	P	$RMSEC$	$RMSECV$	$RMSEP$	R^2_{CAL}	R^2_{VAL}	$RMSEP/RMSEC$
10	25	2	1,669	2,102	2,040	0,715	0,760	1,222
20	11	6	2,034	2,362	2,249	0,607	0,573	1,105
30	19	4	1,495	2,722	1,871	0,808	0,749	1,251
40	11	2	1,086	2,850	1,953	0,894	0,699	1,797
50	19	2	1,866	2,616	2,357	0,699	0,603	1,263
60	19	6	1,493	2,436	2,646	0,791	0,488	1,772
70	19	4	1,014	2,340	2,400	0,894	0,707	2,365
80	19	2	1,168	1,773	2,280	0,865	0,627	1,952
90	19	4	1,441	2,545	2,360	0,789	0,625	1,636
100	11	4	1,849	2,795	2,347	0,693	0,578	1,269

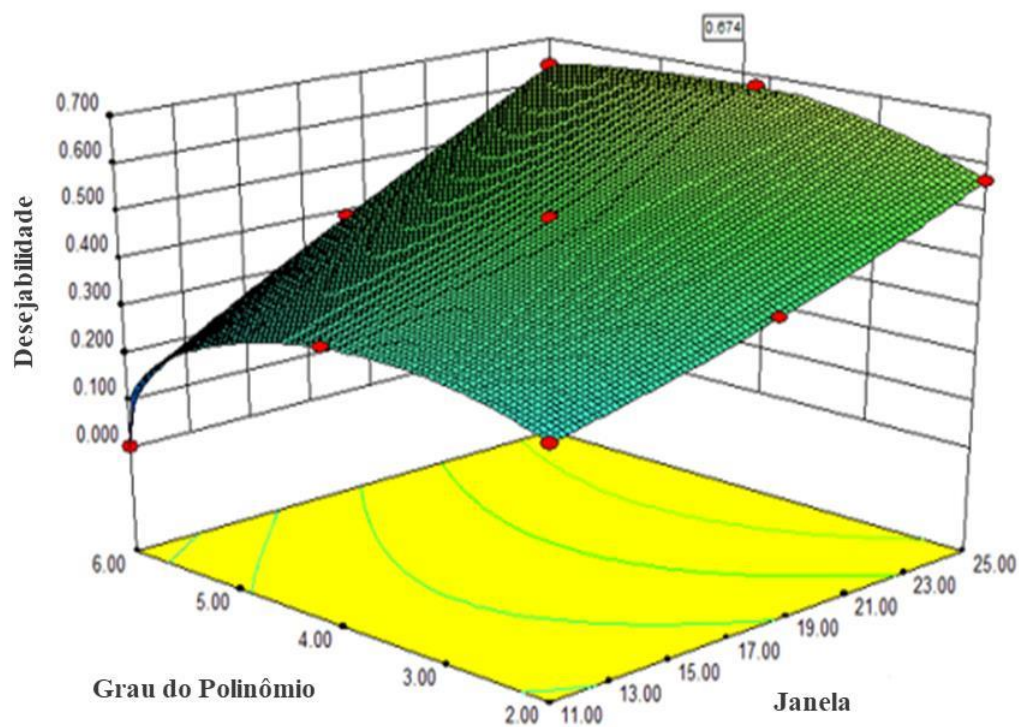


Figura 7. Diagrama de contorno para erros de classificação de validação cruzada, R^2_{cal} , R^2_{val} e razão entre erros de calibração e validação para espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 30 nm.

A avaliação dos modelos foi realizada determinando e analisando o coeficiente de determinação (R^2) da curva obtida entre os valores previstos versus medidos e os erros de validação cruzada e predição, cujos valores indicam a proporção de variação na variável de resposta y explicada pelas variáveis independentes X no modelo de regressão linear. Assim, quanto maior seu valor, maior a variabilidade explicada pelo modelo de regressão linear.

A avaliação do desempenho do modelo preditivo também considerou outros critérios estatísticos, como $RMSECV$ e a relação entre $RMSEC$ e $RMSEP$, com uma métrica para aceitar o ajuste do modelo adotado como aqueles com uma relação menor ou igual a 2,4, ou seja, garantindo que os modelos não estão sobre ajustados.

Como pode ser observado, cada $\Delta\lambda$ exibiu uma combinação ótima de tamanho de janela e grau polinomial, refletindo as diferentes bandas de emissão obtidas para os vários compostos fluorescentes presentes em cada condição estudada. A partir dos dados da Tabela 3, um gráfico de superfície foi construído com o objetivo de avaliar dentre os de modelos mostrados, qual seria o melhor modelo quimiométrico para a previsão da estabilidade oxidativa das amostras de B100. A Figura 8 mostra que os modelos com os menores erros de calibração e predição são notadamente os obtidos com os $\Delta\lambda$ de 30 e 40 nm, em relação aos demais modelos.

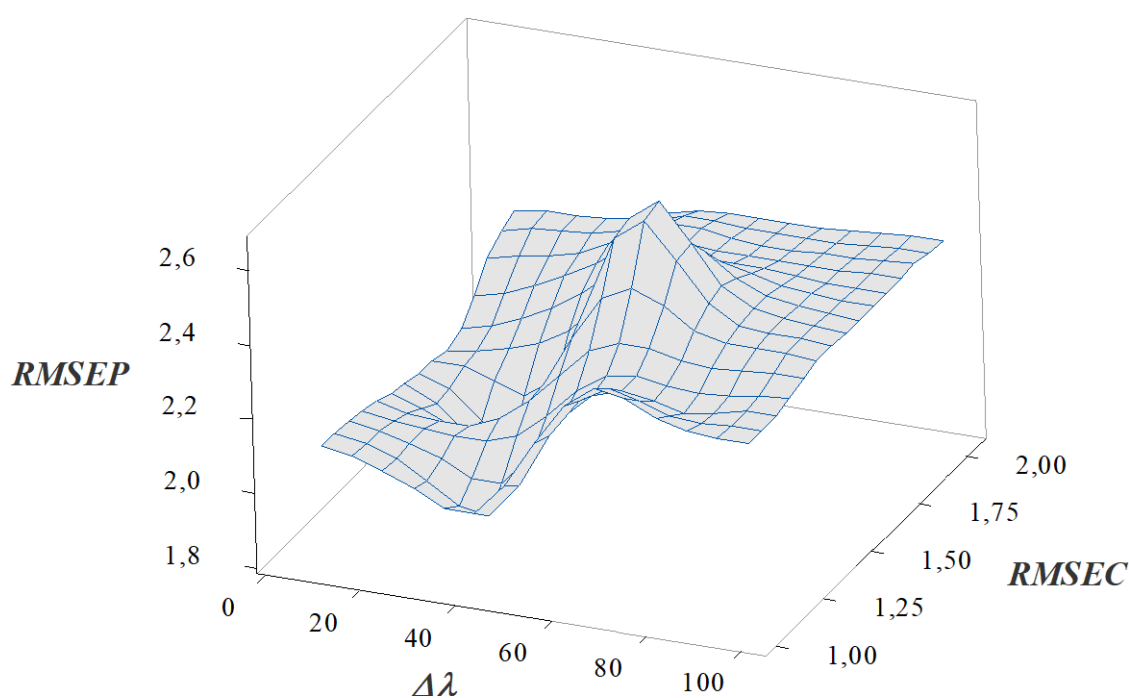


Figura 8. Valores de $RMSEC$ e $RMSEP$ para diferentes valores de $\Delta\lambda$.

Os gráficos de importância das variáveis na projeção (*Variable Importance in Projection*, VIP escores) para os melhores modelos com $\Delta\lambda$ de 30 e 40 nm (Figura 9) mostram perfis parecidos indicando que dois dos comprimentos de onda mais relevantes são semelhantes para os dois modelos, destacando-se os comprimentos de onda de próximos a 360, 470, 480, 640, 650 e 680 nm. A atribuição desses comprimentos de onda específicos é difícil, pois a SFS não é uma técnica comumente usada para impressão digital.

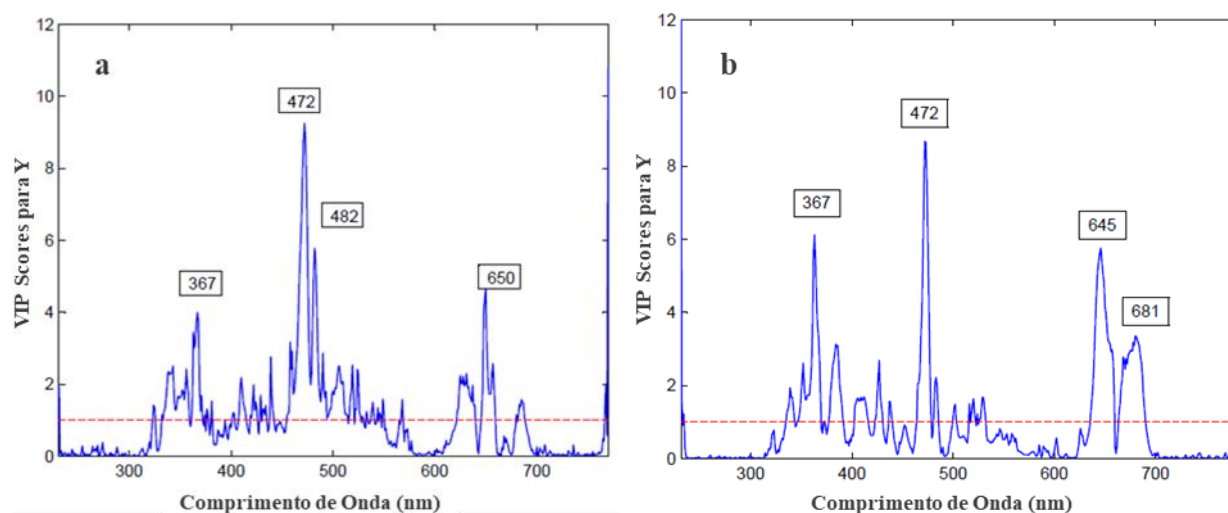


Figura 9. Gráfico das variáveis significativas da regressão (VIP scores) para os modelos com diferentes valores de $\Delta\lambda$. a) 30; b) 40 nm.

Com o objetivo de aumentar a quantidade de informação disponível para a regressão e minimizar os erros de calibração e validação, uma nova estratégia de pré-processamento dos dados foi utilizada, a partir da normalização dos dados pela intensidade de picos com pouca importância para a regressão. Para tal, foi feita uma análise dos gráficos de importância das variáveis na projeção e do gráfico do vetor de regressão dos modelos intermediários para os modelos com $\Delta\lambda$ de 30 e 40 nm, que proporcionaram modelos destacadamente melhores que os demais deltas, onde foram observados e identificados picos com baixa significância para a obtenção desses modelos (Figura 10).

Para o modelo de $\Delta\lambda$ de 30 nm foram escolhidos para a normalização os comprimentos de onda de 351, 385, 427 e 501 nm (Figura 10a) e os de 384, 427, 501, 560 nm para o modelo com $\Delta\lambda$ de 40 nm (Figura 10b). Estes picos apresentaram simultaneamente, nos modelos intermediários, altos coeficientes negativos no gráfico do vetor de regressão e baixos valores de importância da variável. Desta forma, as intensidades de fluorescência registradas nesses comprimentos de onda foram usadas para a normalização da matriz X para os $\Delta\lambda$ de 30 e 40 nm.

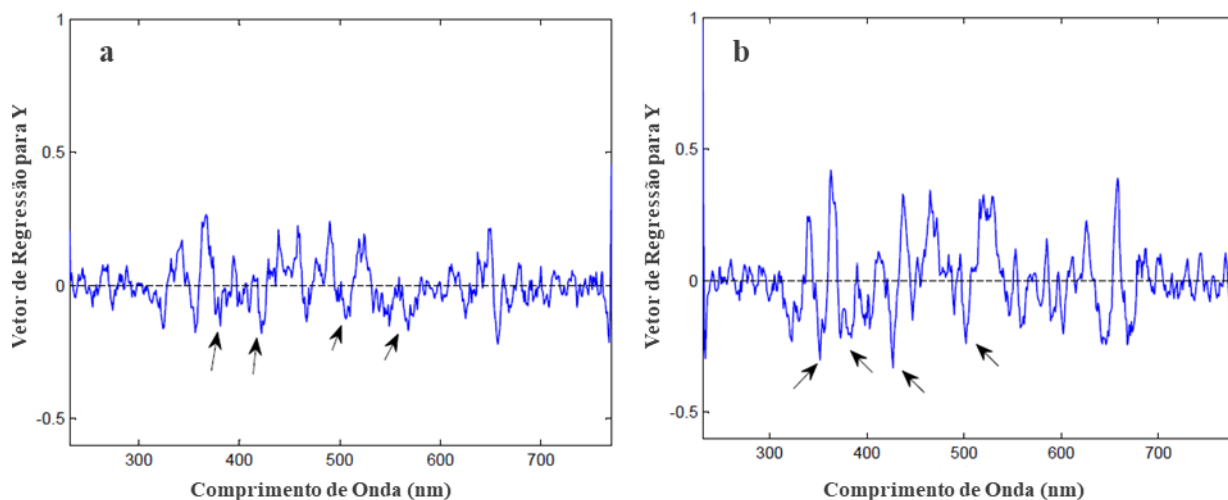


Figura 10. Gráficos dos coeficientes de regressão para os modelos com diferentes valores de $\Delta\lambda$. a) 30; b) 40 nm.

Foram então construídos novos modelos PLS a partir das matrizes normalizadas, sem prejuízo para a aquisição de dados de amostras de biodiesel. O melhor modelo normalizado para o $\Delta\lambda$ de 30 nm foi obtido utilizando o comprimento de onda de 501 nm, enquanto para o $\Delta\lambda$ de 40 nm foi o comprimento de onda de 384 nm. Isso resultou em uma melhora significativa nos valores das figuras de mérito dos modelos de PLS obtidos a partir destas normalizações para os modelos com $\Delta\lambda$ de 30 e 40 nm.

Na Tabela 4 são apresentados os valores dos erros de classificação de validação cruzada, R^2_{cal} , R^2_{val} e a razão entre os erros de calibração e validação para o planejamento realizado usando os espectros registrados para $\Delta\lambda = 30$ e 40 nm, antes (*bn*) e depois (*an*) das normalizações.

Os valores da Tabela 4 mostram que os modelos construídos a partir das matrizes normalizadas, apresentaram uma melhora significativa, tanto para os valores dos erros quanto para os R^2 de calibração e validação. A normalização proporcionou uma redução entre 14 e 31% nos valores de $RMSEC$, $RMSECV$ e $RMSEP$ para os $\Delta\lambda$ de 30 e 40 nm, mostrando a importância do pré-tratamento adequado dos dados para a obtenção de modelos com os menores erros possíveis.

Os valores de $RMSEP$ apresentados para-ambos os modelos são menores do que a máxima reprodutibilidade permitida pelo método de referência (EN 14112) para valores de estabilidade oxidativa acima de 6,4 h, indicando que está sujeito a menos erros aleatórios.

Tabela 4. Tamanho da janela (J), grau polinomial (P), coeficientes de determinação, erro quadrático médio da validação cruzada e razão entre os erros quadráticos médios da predição e classificação com e sem normalizações para $\Delta\lambda$ de 30 e 40 nm.

<i>Parâmetros</i>	$\Delta\lambda = 30 \text{ nm}$		$\Delta\lambda = 40 \text{ nm}$	
	-	$l_n = 501 \text{ nm}$	-	$l_n = 384 \text{ nm}$
<i>J</i>	19	19	11	19
<i>P</i>	4	2	2	2
<i>VL</i>	7	6	9	8
<i>RMSEC</i>	1,495	1,045	1,086	0,916
<i>RMSECV</i>	2,722	2,343	2,850	2,136
<i>RMSEP</i>	1,871	1,283	1,953	1,417
R^2_{cal}	0,808	0,912	0,894	0,917
R^2_{val}	0,749	0,822	0,699	0,843
<i>RMSEP/RMSEC</i>	1,251	1,227	1,797	1,517

Por apresentar um *RMSEP* e uma razão *RMSEP/RMSEC* ligeiramente menores, o modelo normalizado em 501 nm com o $\Delta\lambda$ de 30 nm se mostra o mais adequado para a previsão dos valores da estabilidade oxidativa das amostras de B100. A relação entre os valores do y medido pelo y previsto para este modelo (Figura 11) resultaram em uma boa correlação, com coeficientes de correlação de 0,912 e 0,822 para os modelos de calibração e previsão, respectivamente.

Quando se comparam as figuras de mérito do modelo proposto com outros descritos na literatura observa-se uma maior faixa de aplicação, 6 a 18 h, que abrange satisfatoriamente o limite estabelecido na legislação brasileira de 13 h.

Apesar de resultar em erros de calibração e de previsão ligeiramente maiores que outros modelos obtidos por diferentes técnicas, o modelo proposto foi obtido com amostras comerciais oxidadas naturalmente, conferindo uma maior representatividade, uma vez que alguns modelos foram construídos com amostras oxidadas artificialmente, às vezes a partir de uma mesma amostra, não refletindo necessariamente a realidade dos processos oxidativos que o B100 é submetido no mercado.

Desta forma, o modelo pode ser utilizado satisfatoriamente para a determinação do período de indução do B100 e na separação das amostras que tenham valores próximos ao limite estabelecido na legislação para o ensaio segundo o método previsto na legislação, agilizando o processo analítico.

Tabela 5. Tabela com trabalhos encontrados na literatura.

Autores / Técnica Espectral	Matéria Prima	Tipos de Amostras	Nº de Amostras	Estabilidade Oxidativa	RMSEP (h)
Este Trabalho – SFC	Indefinida	Comerciais	65	6,0 a 18,0 (h) Rancimat	1,4
[20] – SFC	Soja	Não comerciais	10	0 a 5,9 (h) Rancimat	0,63
[51] - FT-NIR	Óleo de cozinha usado, colza e girassol	Não comerciais	133	18 a 128 (min) PetroOxy	RMSECV 0,14
[53] - FT-NIR	Olive, girassol, amendoim, noz, milho, avelã, colza, soja, linhaça e algas e colza	Não comerciais	11	0,34 a 3,97 (h) Rancimat	1,1
[52] - FT-NIR	Óleo de cozinha usado, soja, palma, colza e girassol	Não comerciais	199 (sem antioxidantes)	0,66 a 17,75 (h) Rancimat	0,60

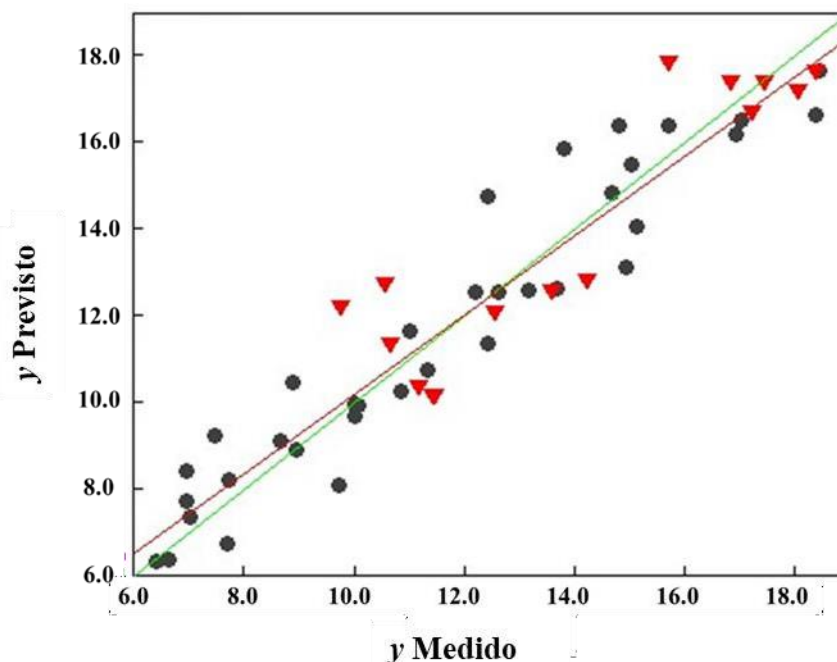


Figura 11. Gráfico do y predito pelo y medido para as amostras de B100 para diferentes modelos normalizados para $\Delta\lambda$ de 30 nm. (●) Calibração; (▼) Validação.

3.4 CONCLUSÃO

- A otimização do pré-tratamento dos espectros por meio do desenho experimental de face central e na estratégia de normalização permitiu a obtenção de modelos quimiométricos com erros de calibração e validação mais baixos.
- Os modelos obtidos com $\Delta\lambda$ de 30 e 40 nm apresentaram as melhores avaliações das figuras de mérito $RMSEC$, $RMSECV$, $RMSEP$, R^2 .
- O uso de amostras comerciais de biodiesel conferiu ao modelo uma maior robustez, ao considerar as variabilidades reais dos métodos de produção, matérias-primas e processos de oxidação natural.
- É uma técnica multivariada que utiliza pequenos volumes de amostra e respostas rápidas. O modelo quimiométrico é capaz de prever os valores de Estabilidade Oxidativa de forma mais ágil do que o método de referência e com eficiência similar às outras descritas na literatura.

4.0 CLASSIFICAÇÃO E QUANTIFICAÇÃO DE CONTAMINANTES EM QUEROSENE DE AVIAÇÃO

4.1.Introdução

A adição de solventes orgânicos, como hidrocarbonetos alifáticos leves, alifáticos pesados e aromáticos, aos combustíveis é uma prática comum, mas ilegal [65]. Essa atividade, conhecida como adulteração, tem como objetivo modificar a composição do combustível para obter lucro financeiro. Além disso, os combustíveis podem sofrer alterações em sua composição original devido a contaminações, que podem ser provocadas tanto por solventes quanto por outros tipos de combustíveis.

A modificação da matriz original de um combustível, seja ela por adulteração ou por contaminação, pode causar diversos problemas à sociedade, como risco ambiental devido à emissão de vapores e gases tóxicos, ou seja, CO e NO_x, menor durabilidade ao motor dos veículos, problemas mecânicos frequentes e até mesmo causar algum tipo de acidente [66].

Segundo os dados do Sistema de Investigação e Prevenção de Acidentes Aeronáuticos (SIPAER), o ano de 2024 foi o que mais registrou acidentes aéreos nos últimos dez anos no Brasil, foram cerca de 148 vidas perdidas, e em segundo lugar fica o ano de 2016 com 104, contabilizando um total de 773 mortes neste período [67]. Estes acidentes muitas vezes têm relação ao mal funcionamento dos motores das aeronaves, podendo estar relacionado com a qualidade dos combustíveis usados durante o voo, assim como a falta de manutenção adequada destes motores.

O querosene de aviação (QAV) é um derivado de petróleo obtido por destilação direta com faixa de temperatura de 150 a 300 °C, com predominância de hidrocarbonetos parafínicos de 9 a 15 átomos de carbono, sendo 33-61% de parafinas, 33-45% naftênicos, 12-25% aromáticos e 0,5-5% de olefinas. Em concentrações mais baixas apresentam ainda compostos de enxofre, oxigênio e nitrogênio [68].

No Brasil, a frota da aviação comercial é responsável pelo consumo anual estimado de cerca de 7 bilhões de litros de querosene. Entre 2003 e 2011, houve um aumento de 75% no uso deste combustível, fazendo com que fosse necessária sua importação, pois o país não é autossuficiente nessa produção [69]. No ano de 2023 foram produzidos 5,44 milhões de m³, e importados cerca de 992,1 mil m³ [70].

A regulamentação e fiscalização dos combustíveis de aviação são realizadas pela ANP por meio da Resolução n° 856, de 22 de outubro de 2021 [2], garantindo assim sua qualidade por meio da padronização das suas propriedades físico-químicas.

Porém, ainda assim, o combustível de aviação esta suscetível a contaminações, que podem ser provenientes de outros tipos de combustíveis ou até mesmo por solventes, e podem ocorrer durante o transporte ou até mesmo durante a estocagem. A contaminação em QAV representa um grande problema atualmente, uma vez que causa danos nos componentes do motor das aeronaves além de colocar em risco a segurança das operações. Os solventes mais comuns de serem encontrados em um combustível adulterado/contaminado incluem: etanol em quantidades excessivas, diesel, querosene, petroquímicos refinados, tolueno, xileno e hexano, entre outros [71].

Atualmente a literatura científica tem relatado metodologias alternativas na detecção de adulteração em combustíveis, em grande maioria para diesel e gasolina [65-66] e em menor proporção para os demais combustíveis como o querosene e gasolina de aviação [68]. Essas metodologias empregam ferramentas quimiométricas combinadas com diversas técnicas convencionais de análise para combustíveis.

Mendes e Barbeira [66] desenvolveram modelos de PCA (*Principal Component Analysis*) e PLS-DA (*Partial Least Squares Discriminant Analysis*) juntamente com curvas de destilação (ASTM D86) [72], para discriminar amostras de gasolina adulterada e não adulterada, bem como, a determinação do solvente usado na adulteração.

De Paulo e Barbeira [73] usaram espectroscopia de emissão de chama (FES) combinada com o método de análise discriminante por mínimos quadrados parciais (PLS e PLS-DA) para a predição e quantificação de adulterações em gasolina automotiva a partir de diversos solventes (terebintina, tiner, querosene, solvente de borracha e etanol).

Câmara et al. [74] desenvolveu um trabalho buscando avaliar o impacto na qualidade do querosene de aviação, quando adulterado com querosene iluminante, utilizando técnicas espectroscópicas de infravermelho (MIR/NIR) aliados às técnicas quimiométricas de análise exploratória (PCA), análise discriminante linear (PCA-LDA) e por resolução de curvas multivariadas com mínimos quadrados alternados (MCR-ALS) a fim de se obter um modelo estatístico que conseguisse prever a presença de adulterante e a quantidade presente deste no querosene de aviação. Em um outro trabalho, Câmara e sua equipe elaborou um trabalho usando Espectroscopia de fluorescência de excitação-emissão e métodos PARAFAC e MCR-ALS para investigação da contaminação de querosene de aviação Jet A1 por querosene comercial [68].

Para tentar detectar contaminações e inibir essas adulterações, determinadas abordagens podem ser realizadas, algumas sendo oficialmente adotadas pela ANP, como os ensaios físico-químicos normalizados, que podem evidenciar a presença de contaminantes nos produtos derivados de petróleo [75]. Porém, nem sempre apenas parâmetros físico-químicos são suficientes para caracterizar a adulteração e ou contaminação. Diante disso, novas metodologias

estão sendo aplicadas com finalidade de identificar e quantificar compostos diferentes em matrizes de combustíveis. Essas metodologias, baseiam-se na utilização de técnicas de analíticas convencionais associadas a um conjunto de ferramentas estatísticas, como: PCA, PLS e PLS-DA, integrando aspectos quimiométricos, viabilizando assim uma poderosa ferramenta de detecção e quantificação de contaminantes [67].

A avaliação da discriminação nos modelos PLS-DA se dá através da sensibilidade, especificidade e taxa de confiabilidade dos modelos. A sensibilidade se refere à capacidade do método em identificar amostras verdadeiramente positivas como positivas, sendo a taxa de sensibilidade a probabilidade numérica dessa habilidade. De maneira análoga, a especificidade se refere à capacidade do método em identificar amostras reais negativas como negativas, sendo a taxa de especificidade a probabilidade numérica dessa identificação. A taxa de falso positivo (*TFP*) refere-se à proporção de resultados de falso-positivos (*FP*) em relação à soma dos falsos positivos (*FP*) e verdadeiros-negativos (*VN*). Já a taxa de falsos negativos (*TFN*) é a proporção de resultados falsos negativos (*FN*) em relação à soma dos falsos negativos (*FN*) e verdadeiros positivos (*TP*). Por fim, a eficiência se refere a uma medida geral da capacidade de classificação e/ou previsão do modelo [31].

O objetivo desse trabalho foi desenvolver modelos de discriminação (PLS2-DA) para identificar amostras contaminadas e modelo de regressão (PLS) para quantificar os contaminantes utilizando a Espectroscopia de Fluorescência Sincronizada para amostras de querosene de aviação não contaminadas e contaminadas com: gasolina, gasolina de aviação, querosene iluminante, querosene comercial e diesel.

4.2 Materiais e Métodos

4.2.1. Amostras

Foram analisadas 67 amostras de QAV dopadas com diferentes tipos de contaminantes em proporções que variaram de 0,5%, 1%, 2%, 5%, 7% e 10% (v/v) de contaminação e 60 amostras de QAV não contaminados. As amostras de QAV foram fornecidas pelo Laboratório de Ensaios de Combustíveis do Departamento de Química da Universidade Federal de Minas Gerais (LEC-UFMG), provenientes de diferentes refinarias.

Foram usados quatro tipos de QAV diferentes para as dopagens e cinco tipos de contaminantes: Gasolina automotiva (GA), Querosene comercial (QC), Querosene Iluminante (QL), Gasolina de Aviação (GAV) e Diesel (D).

4.2.2. Instrumentação

As medidas espectrofluorimétricas foram realizadas usando um equipamento Shimadzu RF-5301PC (Duisburg, Alemanha) equipado com uma lâmpada de xenônio. As amostras foram analisadas em uma cubeta de quartzo com quatro lados polidos e dimensões de 10 mm x 10 mm x 45 mm. O equipamento foi configurado com uma largura de fenda de excitação de 5 nm e largura de fenda de emissão de 10 nm, e resolução espectral de 1 nm. Os espectros foram coletados através da varredura simultânea dos monocromadores de excitação e emissão na faixa de 230 a 500 nm, com $\Delta\lambda$ constante entre eles, de 10 a 100 nm, em intervalos de 10 nm. As medidas de fluorescência foram realizadas em triplicata, e o espectrofluorímetro foi verificado diariamente usando uma amostra de controle.

A cromatografia gasosa com detector de espectrometria de massas (GC-MS) foi realizada usando um equipamento GC-17A da Shimadzu com detector CGMS-QP5050. A coluna usada foi a db-petro 100% dimetilpolisiloxano, de dimensão 100m x 0,25mm x 0,5um e a identificação do composto foi realizada usando o Software Labsolutions.

4.2.3. Tratamento de dados

O software utilizado foi o MATLAB versão 7.10.0 - R2010A (The MathWorks, Natick, MA, EUA) e a ferramenta PLS Toolbox versão 5.2.2 (Eigenvector Technologies, Manson, WA, EUA) foram utilizados para conduzir as análises multivariadas.

Para a construção do modelo PLS e PLS-DA, os dados das amostras de QAV contaminadas foram organizados em uma matriz bidimensional \mathbf{X} (amostras $\times \lambda_{em}$) sendo que, para cada $\Delta\lambda$, os tratamentos de dados foram realizados separadamente. Os espectros de fluorescência foram pré-processados usando a primeira derivada com suavização usando o filtro *Savitzky-Golay*, e os dados foram centralizados na média. Parâmetros como grau polinomial e largura do filtro foram selecionados com base em uma análise dos vales das Taxas de confiabilidade dos modelos e do planejamento do composto central. As combinações para o tamanho da janela e o grau de polinômio foram: 13,17 e 21 para a largura da janela e 2, 4 e 6 para os graus de polinômios.

As amostras foram separadas em conjuntos de calibração e validação usando o algoritmo de *Kennard-Stone*, de forma separada para cada classe de contaminação e aquelas que apresentaram simultaneamente resíduo Q e T^2 de *Hotelling* acima dos valores críticos, com nível de confiança de 95%, foram consideradas outliers pelo modelo e, conseqüentemente, excluídas.

Além da avaliação do modelo através dos valores dos erros (*RMSE*) e do coeficiente R^2 , performance do modelo PLS2-DA também foi realizada por meio de alguns parâmetros de qualidade como sensibilidade (*SEN*), especificidade (*SPE*), taxa de falso negativo (*TFN*), taxa de falso positivo (*TFP*) e eficiência (*EEF*)

4.3. RESULTADOS E DISCUSSÃO

4.3.1. Espectrofluorimetria e Cromatografia

Para que os modelos fossem representativos, foram escolhidas quatro matrizes de QAV de refinarias diferentes, nas Figuras 12 e 13 são apresentados os perfis espectrais e cromatográficos, respectivamente, das quatro matrizes (QAV1, QAV2, QAV3 e QAV4) antes da dopagem e os cinco contaminantes (GA, QL, QC, GAV e D) usados para dopar as amostras de QAV.

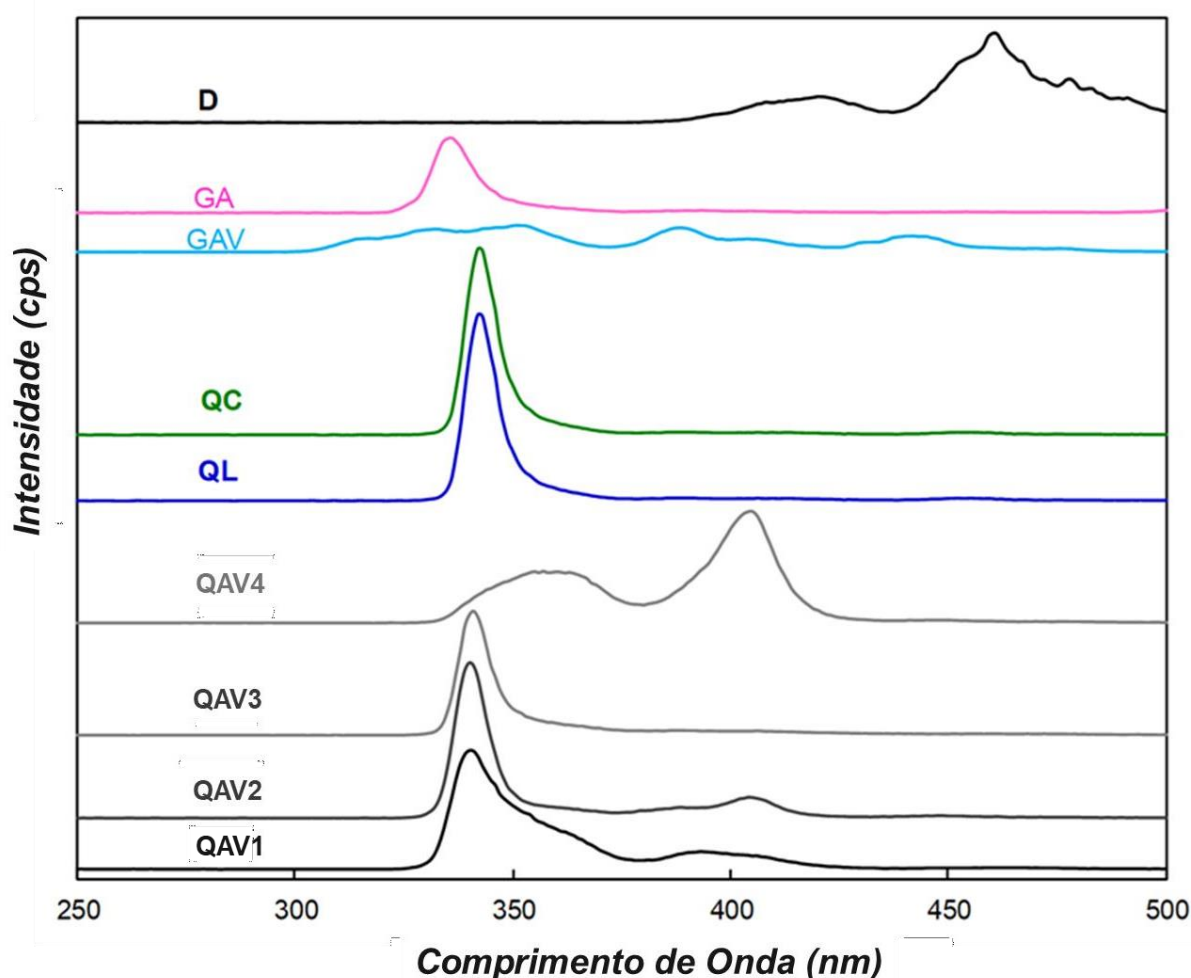


Figura 12. Espectros de fluorescência sincronizada ($\Delta\lambda$ 10 nm) das matrizes de QAV (QAV1, QAV2, QAV3 e QAV4) e os contaminantes (GA, QL, QC, GAV e D).

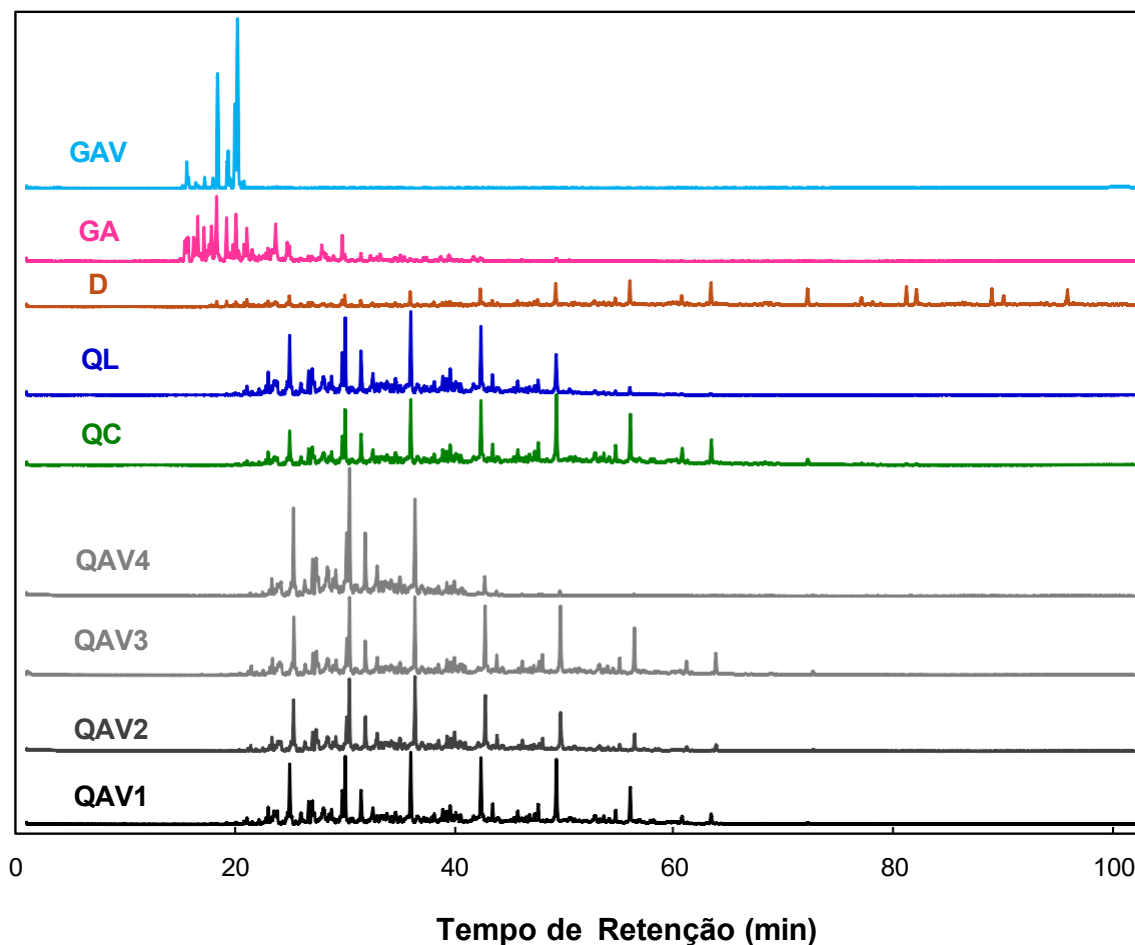


Figura 13. Cromatograma dos QAVs antes da dopagem e dos continentes: Azul-QL, Rosa-QC, Perto-QAV, Vermelho-D, Verde-GA, e Azul escuro-GAV.

Cada matriz de QAV possui perfil diferentes, sendo que as matrizes de QAV matrizes (QAV1, QAV2, QAV3 e QAV4) possuem um pico mais intenso entre 330 e 370 nm e um de menos intensidade por volta de 400 nm, já O QAV4 tem um perfil diferente das demais, pois seu pico de maior intensidade é o de 400 nm e no intervalo de 330 a 370 nm tem-se uma banda de menor intensidade, o que traz aos modelos propostos uma maior representatividade.

Os contaminantes QL e QC possuem um perfil bem parecido com as matrizes de QAV (QAV1, QAV2 e QAV3), a GA apresentou um pico um pouco antes da região observada para o QL e QC, por volta de 340 nm e com uma menor intensidade e o GAV possui em sua composição o corante azul 1,4 dialquilamino antraquinona segundo a resolução 901 da ANP de 09 de março de 2022, esse corante possui grupos fluoróforos em sua composição por conter grupos aromáticos policíclicos [76].

Para a aquisição do espectro do diesel, foi necessário modificar as condições de leitura do equipamento, pois nas condições pré-estabelecidas para as demais amostras, o espectro do

diesel estourava a escala de leitura do equipamento. O espectro obtido em outras condições de leituras, 5 mm de janela de excitação e 5 mm de janela de emissão, mostra que o seu perfil é totalmente diferente das demais amostras, apresentando picos apenas após 400 nm.

Na Figura 13 é apresentado o perfil cromatográfico dos cinco contaminantes: querosene iluminante, querosene comercial, gasolina automotiva, gasolina de aviação e do querosene de aviação puro.

Em uma análise qualitativa, pode-se dizer que o cromatograma do QAV é muito semelhante aos cromatogramas dos contaminantes QL e QC, uma vez que se trata de compostos muito semelhantes, portanto nenhum dos espectros de massa mostra uma região que poderia ser usada para identificar as diferentes composições destes três compostos. Já para os contaminantes GA e GAV percebemos uma similaridade apenas para os componentes chamados de leves, em contrapartida, o diesel já possui um perfil diferente dos demais contaminantes, mostrando a presença de uma quantidade muito maior de compostos, quando comparado aos demais tipos de contaminantes e até mesmo quando comparado ao próprio QAV.

Dessa forma não houve necessidade de se obter um modelo quimiométrico para distinguir as amostras contaminadas com o diesel pois, nas condições estabelecidas para as leituras no fluorímetro, essas amostras estouravam na escala de leitura das demais, mostrando um perfil visivelmente muito diferente das demais amostras contaminadas e não contaminadas, sendo possível então já classificá-las como contaminadas por diesel.

A Figura 14 mostra as quatro matrizes dopadas com os 4 tipos de tipos de contaminantes em diferentes $\Delta\lambda$, o perfil dos espectros varia de acordo com o delta usado. Na utilização da fluorescência sincronizada, a emissão de um fluoróforo em um comprimento de onda λ_{em} apenas será detectada se obedecer à relação $\lambda_{em} = \lambda_{ex} + \Delta\lambda$, resultando em perfis de fluorescência diferentes em cada delta, com picos de diferentes intensidades e posições, contendo ora mais, ora menos informações químicas sobre as amostras. Isso mostra a necessidade de obtenção e avaliação de modelos quimiométricos individuais para cada delta, uma vez que as informações de cada delta são independentes entre si.

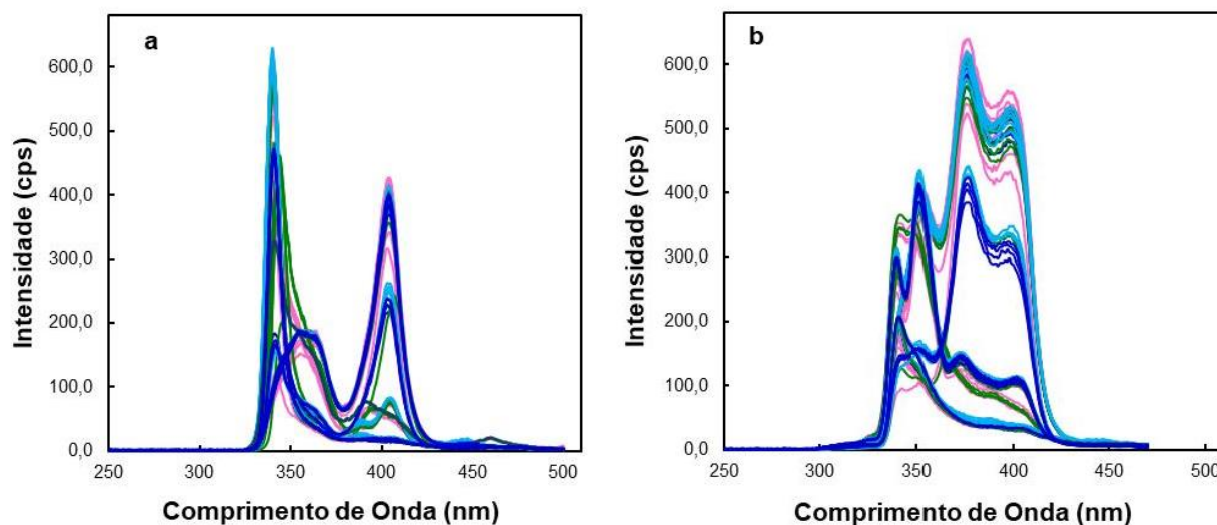


Figura 14. Espectros de fluorescência sincronizada para matrizes de QAV dopadas com GA, GAV, QC e QL obtidos em diferentes valores de $\Delta\lambda$: a) 10; b) 40 nm.

4.3.2. Otimização dos modelos PLS e PLS2-DA

Durante a construção de modelos multivariados, o pré-processamento é uma etapa importante e deve ser otimizada. No presente estudo, utilizou a primeira derivada com alisamento por usando filtro de *Savitzky-Golay* visou atenuar os desvios da linha de base, o ruído espectral e aumentar a variação das bandas espectrais, sendo empregado para pré-processar os espectros de fluorescência síncrona de maneira a não exacerbar a distorção do sinal

Neste estudo, os espectros de fluorescência síncrona das amostras de QAV contaminadas foram pré-processados por meio da primeira derivada com alisamento usando o algoritmo de *Savitzky-Golay* que, após suavização e derivatização, foram centrados na média.

Para determinar os parâmetros de suavização mais adequados de *Savitzky-Golay* (tamanho da janela e grau polinomial) para cada $\Delta\lambda$, foram realizados planejamentos de face centrada, visando uma condição ótima de pré-processamento para cada delta, levando em consideração os erros mínimos de classificação nos conjuntos de treinamento, validação cruzada e teste, além as sensibilidade e especificidade para os modelos.

Desta forma, foram construídos e avaliados nove modelos para cada $\Delta\lambda$ estudado (10 a 100 nm), totalizando 90 modelos. Para exemplificar, na Tabela 6 estão apresentados os valores para o tamanho e da janela e o grau do polinômio que apresentou melhores condições de sensibilidade, especificidade, e taxa de confiabilidade para cada delta. A Figura 15 apresenta a superfície de resposta para o planejamento de face centrada para o $\Delta\lambda$ de 20 nm, indicando que a melhor combinação para este delta foi para uma janela de 21 pontos e polinômio de ordem 4.

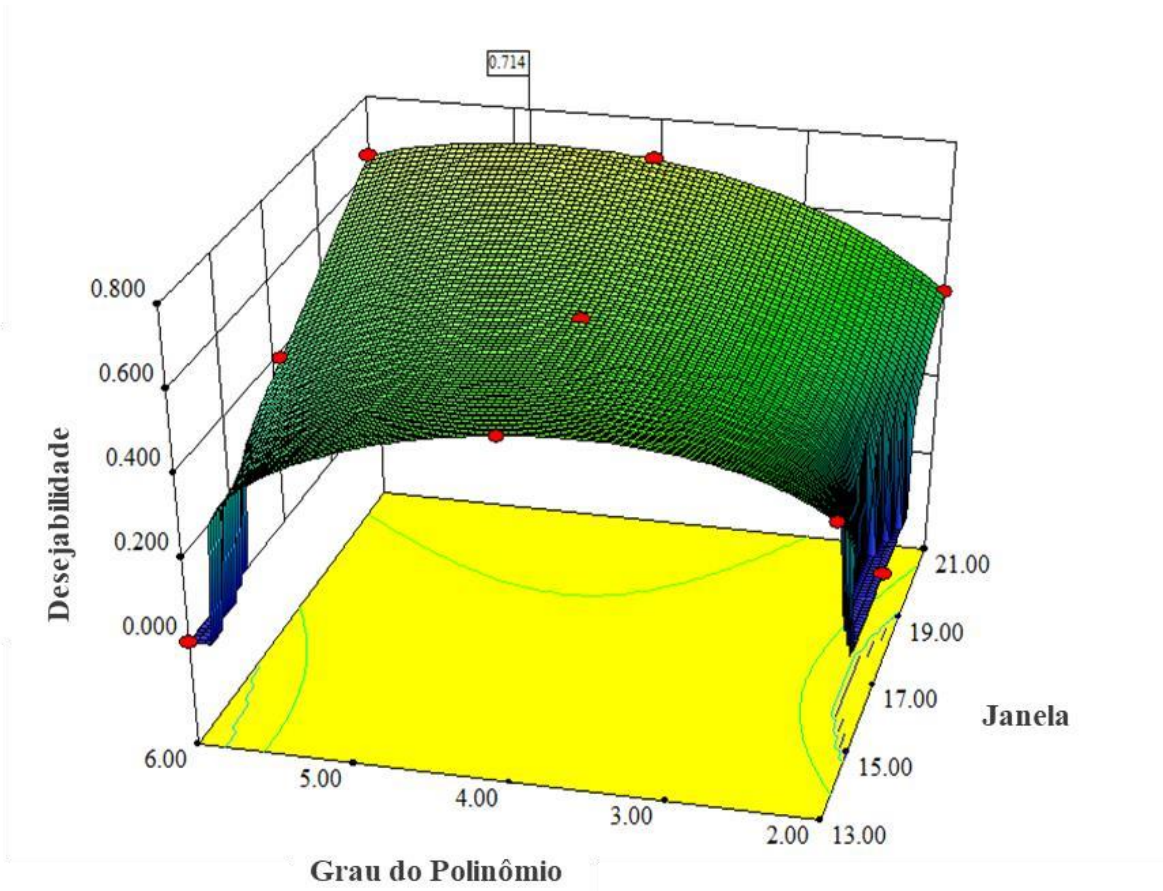


Figura 15. Gráfico de superfície de contorno do planejamento de face centrada $\Delta\lambda$ de 20 nm.

Tabela 6. Tamanho de Janela e grau de polinômio para os melhores modelos em cada $\Delta\lambda$ (nm).

$\Delta\lambda$	J	P	$\Delta\lambda$	J	P
10	17	4	60	17	2
20	21	4	70	21	2
30	13	2	80	21	2
40	13	2	90	17	2
50	21	2	100	13	2

A Tabela 6 foi obtida a partir dos resultados de *SEM*, *SPE* e Erros de classificação nos conjuntos de calibração e teste para as classes, ao total foram usados para esta parte do trabalho modelos com quatro classes de QAV contaminados com: GA, QL, QC e GAV.

Os dados obtidos na Tabela 6 foram usados para a construção de todos os modelos de classificação deste trabalho, sendo comprovada sua eficiência por meio dos parâmetros de qualidade como sensibilidade (*SEN*), especificidade (*SPE*), taxa de falso negativo (*TFN*), taxa de falso positivo (*TFP*) e eficiência (*EEF*).

4.3.4. PLS2-DA classificação entre QAV-NC e as quatro classes de contaminantes: QL, QC, GA e GAV

O modelo tem cinco classes, QAV-NC (querosene não contaminado) e QAV dopados com quatro contaminantes, cada classe com 17 amostras no conjunto de treinamento e oito no conjunto de teste. O modelo que melhor separou as quatro classes foi o $\Delta\lambda$ de 10 nm, pré-processados com janela de 17 pontos e polinômio de 4ª ordem, foi construído utilizando 10 variáveis latentes para um modelo multiclasse, que responderam por 99,9% e 48,1% da variância total nos blocos **X** e **Y** respectivamente, e apresentou cinco amostras anômalas segundo o critério descrito anteriormente.

Devido à similaridade química entre os contaminantes QL/QC e GA/GAV e a complexidade entre as classes, neste modelo só foi possível a discriminação com uma boa eficiência para a classe QAV-NC em relação as demais classes de contaminantes, a discriminação entre classes de contaminantes não ocorreu com uma taxa de eficiência satisfatória, mostrando que o modelo, ao usar o mesmo número de variáveis latentes para todas as classes pode apresentar um certo sobreajuste em relação à estas classes. Os valores de *SEM*, *SPE*, *TFN*, *TFP* e *EEF* estão apresentados na Tabela 7.

A Figura 16 apresenta as previsões de *Y* para cada contaminante em relação aos demais, assim como os valores de limite de decisão (*threshold*), esses valores são estimados pela teoria Bayesiana. Valores de **Y** acima deste limite de decisão indicam que a amostra pertence à respectiva classe, enquanto valores abaixo dele indicam que a amostra não pertence à classe. Para cada classe, é definido um vetor **y** [77]. O modelo PLS-DA apresentado apresentou eficiência máxima na separação de amostras de combustível de aviação contaminada de amostras não contaminadas, sendo então um modelo de extrema relevância para o auxílio do controle de qualidade desse tipo de combustível.

Tabela 7. Parâmetros de desempenho calculados para o modelo PLS2-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 10 nm para classificar QAV contaminado com quatro classes de contaminantes e QAV-NC.

Classes	TREINAMENTO				
	SEM	SPE	TFN	TFP	EEF
GA	0,86	0,75	13,1	24,3	62,4
QL	0,70	0,92	29,0	7,58	63,0
QC	0,82	0,93	18,0	6,00	76,1
GAV	0,82	0,92	18,6	8,00	74,9
QAV-NC	1,00	1,00	0,00	0,00	100

Classes	TESTE				
	SEM	SPE	TFN	TFP	EEF
GA	0,81	0,53	20,0	46,1	33,8
QL	0,42	0,83	57,0	17,1	26,8
QC	0,14	0,91	86,0	8,1	6,00
GAV	0,42	0,95	57,0	4,8	39,8
QAV-NC	1,00	1,00	0,00	0,00	100

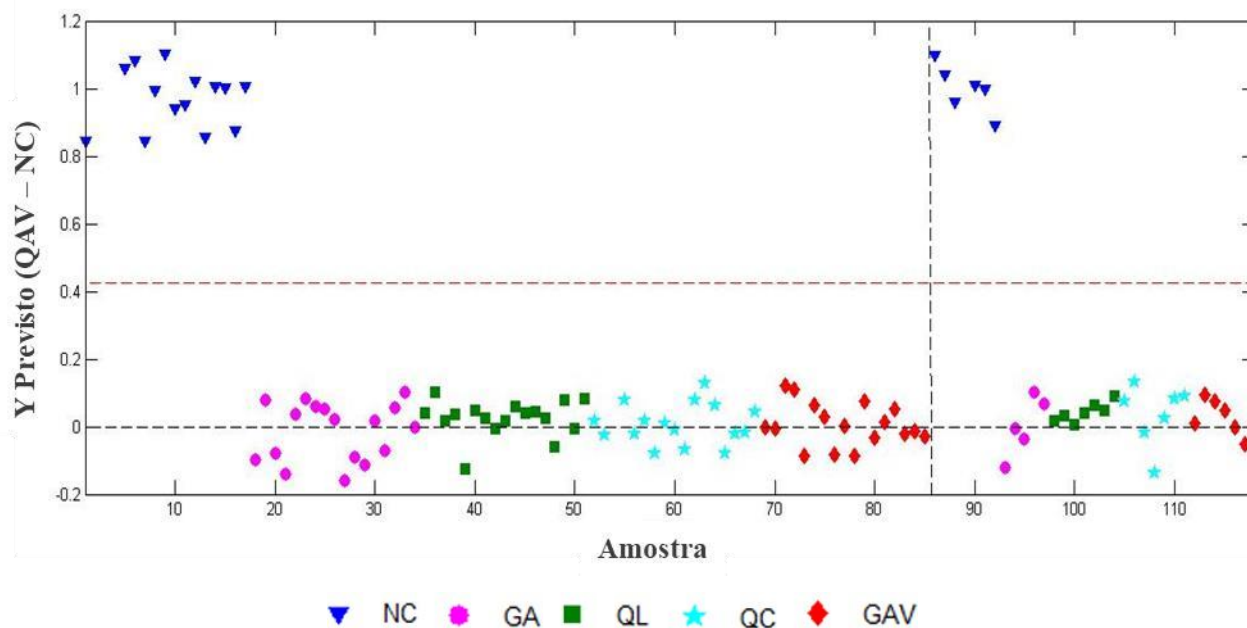


Figura 16. PLS-DA para discriminação de QAV-NC e QAV contaminados com: GA, QL, QC e GAV. Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.

Visto que para o modelo construído, foi possível diferenciar apenas as amostras de QAV não contaminado das amostras contaminadas, não sendo possível dizer a qual classe de contaminante o QAV dopado pertence. Sendo assim, foram feitos novos modelos PLA-DA com o agrupamento das classes de acordo com a similaridade de composição química entre elas, Classe 1 (GA e GAV) e Classe 2 (QL e QC).

4.3.5. PLS-DA classificação entre as classes de QAV contaminado com GA+GAV (Classe 1) e QAV contaminado com QL+QC (Classe 2)

Para discriminar a qual classe de contaminante o QAV dopado pertence, obteve-se um modelo PLS-DA agrupando os contaminados em duas grandes classes, QAV contaminado com GA e GAV (Classe 1) versus QAV contaminado com QL e QC. O modelo que apresentou os melhores valores para as figuras de mérito como *SEM*, *SPE*, *TFN*, *TFP* e *EEF* foi o $\Delta\lambda$ de 10 nm, pré-processados usado foi com janela de 17 pontos e polinômio de 4ª ordem, foi construído utilizando nove variáveis latentes, que responderam por 99,9% e 68,7% da variância total nos blocos *X* e *Y*.

Foram usadas um total de 60 amostras no conjunto de treinamento (30 amostras para cada classe e no conjunto de teste e 30 amostras (15 para cada classe). As amostras foram separadas nestes dois conjuntos através do algoritmo de Kennard-Stone.

A Tabela 8 apresenta os resultados obtidos para o presente modelo, a taxa de eficiência para essa discriminação foi de 86% para o conjunto de treinamento e 70,3% para o conjunto de teste.

Tabela 8. Parâmetros de desempenho calculados para o modelo PLS-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 10 nm para classificar QAV contaminado com duas classes de contaminantes :GA+GAV (classe 1) versus QL+QC (classe 2).

CLASSES	TREINAMENTO				
	<u>SEM</u>	SPE	TFN	TFP	EEF
Classe 1	0,92	0,93	8,12	7,23	86,0
CLASSES	TESTE				
	<u>SEM</u>	SPE	TFN	TFP	EEF
Classe 1	0,84	0,85	15,0	14,1	70,3

Na Figura 17 temos as previsões de Y para cada contaminante em relação aos demais, assim como os valores de limite de decisão (*threshold*) obtidos para o modelo.

Tão importante quanto a qual classe de contaminante pertence o QAV, é saber exatamente qual o contaminante foi usado para dopar as amostras, portanto foram feitos mais dois modelos PLS-DA, um para separar os contaminantes da GA e GAV e outro para QL e QC.

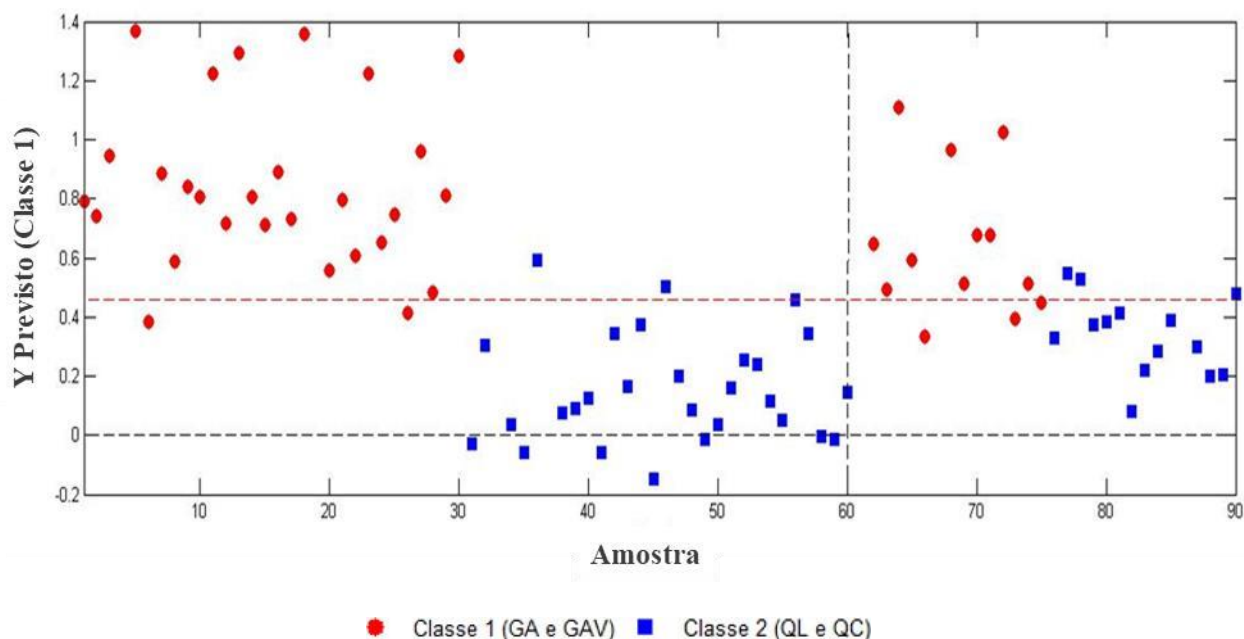


Figura 17. PLS-DA para discriminação de QAV contaminados com: GA+GAV (Classe 1) e QL+QC (Classe 2). Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.

4.3.6. PLS-DA classificação entre as classes de QAV contaminado GA e GAV

O modelo PLS-DA que apresentou os melhores valores para as figuras de mérito como *SEM*, *SPE*, *TFN*, *TFP* e *EEF* foi o $\Delta\lambda$ de 40 nm, pré-processados usando uma janela de 13 pontos e polinômio de 2ª ordem, foi construído utilizando 15 variáveis latentes, que responderam por 99,9% e 89,0% da variância total nos blocos X e Y respectivamente, e apresentou uma (M1GAV 7%) amostra anômala segundo o critério descrito anteriormente.

Foram usadas um total de 34 amostras no conjunto de treinamento (17 amostras para cada classe GA e GAV) e no conjunto de teste e 14 amostras (sete para cada classe). As amostras foram separadas nestes dois conjuntos através do algoritmo de Kennard-Stone.

A Tabela 9 mostra os resultados obtidos para o presente modelo, tendo 100% de eficiência na separação das 2 classes, como se trata de amostras de composições químicas bem distintas entre si como mostra o cromatograma (Figura 18), bons resultados já eram esperados.

Na Figura 10 temos as previsões de Y para cada contaminante em relação aos demais, assim como os valores de limite de decisão (*threshold*).

Tabela 9. Parâmetros de desempenho calculados para o modelo PLS2-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 40 nm para classificar QAV-NC e QAV contaminados com GA e GAV.

CLASSES	TREINAMENTO				
	SEM	SPE	TFN	TFP	EEF
Classe 1(GA)	1	1	0	0	100

CLASSES	TESTE				
	SEM	SPE	TFN	TFP	EEF
Classe 1(GA)	1	1	0	0	100

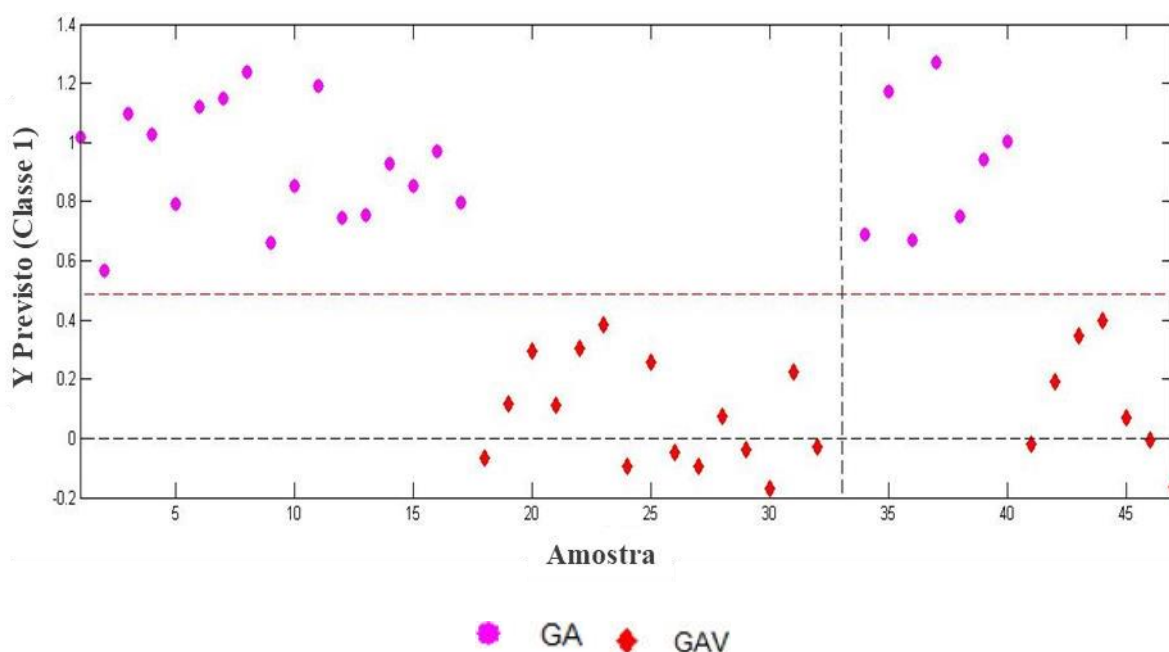


Figura 18. PLS-DA para discriminação de QAV contaminados com GA (Classe1) e GAV (Classe 2). Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.

Este modelo se mostrou eficiente para discriminar as amostras contaminadas com GA e GAV, sendo possível classificar as amostras quanto ao tipo de contaminante, ou seja, se a amostra pertence a classe dos contaminados com GA ou se ela pertence a classe dos contaminados com GAV com 100% de eficiência.

O mesmo tipo de modelo foi feito para discrimina se amostras foram dopadas com QL ou com QC.

4.3.7. PLS-DA classificação entre as classes de QAV contaminado QL (Classe 1) e QC (Classe 2)

O modelo PLS2-DA com $\Delta\lambda$ de 20 nm, pré-processados com uma janela de 21 pontos e polinômio de 4ª ordem, foi construído utilizando seis variáveis latentes, que responderam por 99,9% e 70,9% da variância total nos blocos X e Y respectivamente, duas amostras foram consideradas como anômala neste modelo. O número de amostras e o algoritmo usado para separar os conjuntos em treinamento e teste foram os mesmos do item 3.5.

A Tabela 10 mostra os resultados obtidos para o presente modelo, tendo 93% de eficiência para conjunto de treinamento e 86% para o conjunto teste na separação das 2 classes, neste caso as amostras possuem um perfil muito parecido, como pode ser observado através dos cromatogramas (Figura 13), o que dificulta obter maior eficiência para o modelo, mais ainda sim trata-se de um modelo com valores satisfatórios.

Tabela 10. Parâmetros de desempenho calculados para o modelo PLS-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 20 nm para classificar QAV contaminado com: QL (Classe 1) e QC (Classe 2).

CLASSES	TREINAMENTO				
	SEM	SPE	TFN	TFP	EEF
Classe 1	1	0,93	0	7	93
CLASSES	TESTE				
	SEM	SPE	TFN	TFP	EEF
Classe 1	1	0,85	0	14	86

Na Figura 19 temos as previsões de Y para cada contaminante em relação aos demais, assim como os valores de limite de decisão (*threshold*).

Os modelos para discriminação de contaminantes em amostras de QAV apresentou alta sensibilidade, sendo capaz de discriminar contaminações por volta de 0,5%, trata-se também de modelos representativos, uma vez que as amostras usadas para dopagem foram amostras comerciais de QAV e que possuíam diferentes perfis tanto para os seus espectros de fluorescência quanto para seus cromatogramas.

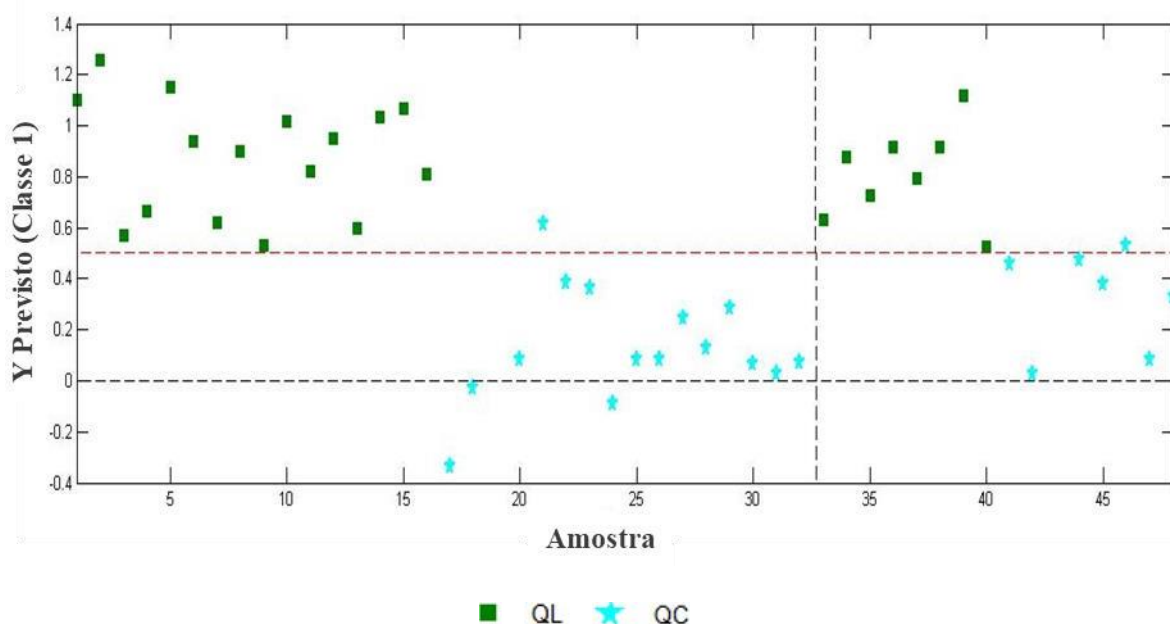


Figura 19. PLS-DA para discriminação de QAV contaminados com QL e QC. Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.

4.3.8. Gráficos de importância variável na projeção VIP scores para os modelos PLS-DA e PLS2-DA

Os comprimentos de onda mais importantes para a classificação dos modelos PLS-DA e PLS2-DA estão dispostos na Tabela 11 assim como os gráficos de importância variável na projeção VIP scores (Figura 20). A atribuição desses comprimentos de onda específicos é difícil, pois o SFS não é uma técnica comumente usada para impressão digital, mas todos os modelos apresentaram picos em comum nas regiões de 335 a 440 nm e na região de 400 a 411 nm, sendo então as duas regiões de maior importância em comum para todos os modelos de classificação.

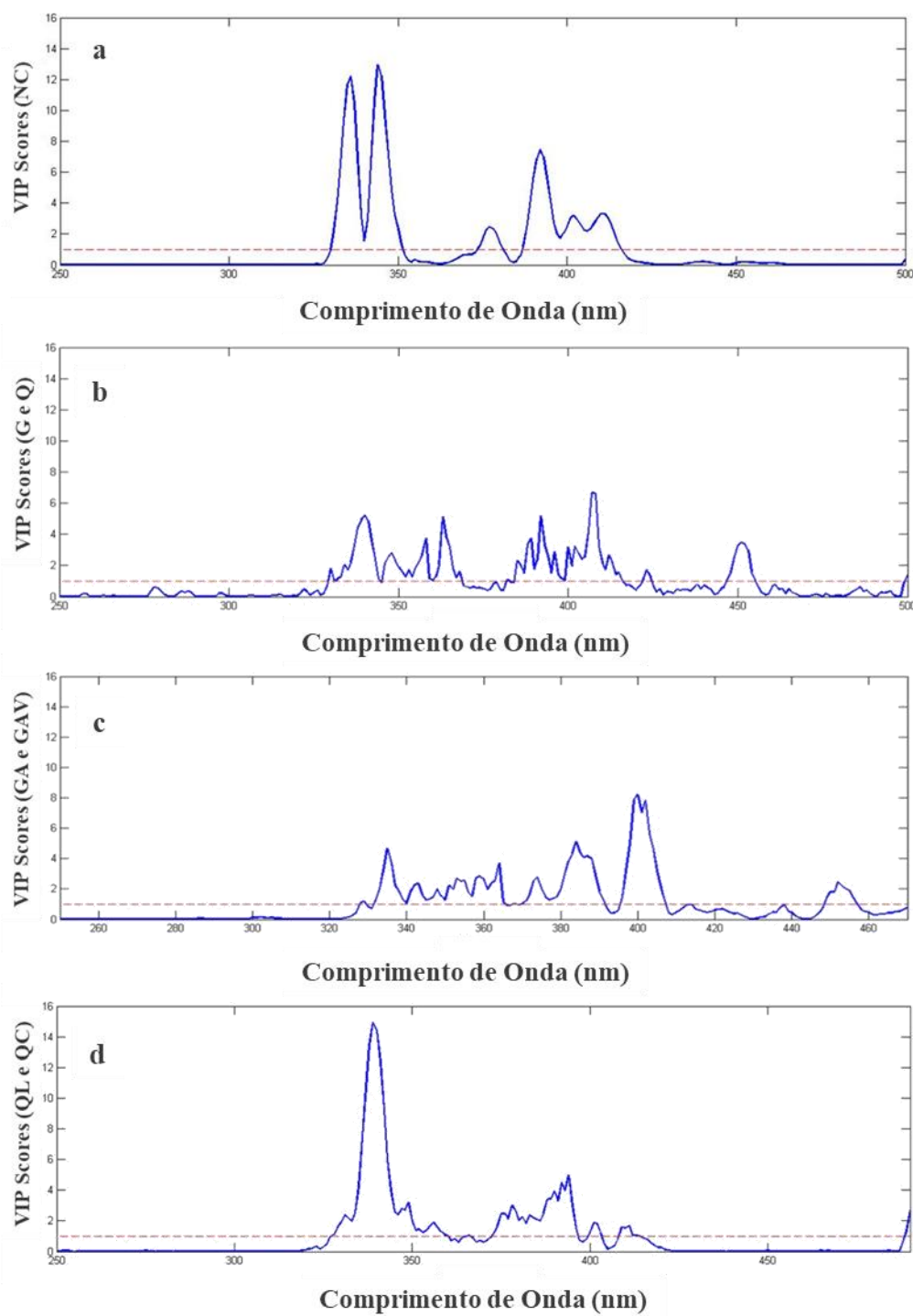


Figura 20. Gráficos VIP scores dos modelos PLS-DA e PLS2-DA de fluorescência síncrona para classificação das amostras contaminadas com: a) QAV(NC) x QAV (GA, GAV, QL e QC), b) QAV(NC) x QAV(C), c) QAV(GA) x QAV(GAV) e d) QL x QC.

Tabela 11. Valores dos comprimentos de onda (nm) presentes nos gráficos de importância variável na projeção *VIP* escores de todos os modelos PLS-DA e PLS2-DA para classificação de amostras de QAV contaminados.

QAV(NC)	QAV(GA e GAV)	QAV(GA)	QAV(QL)
x	x	x	x
QAV (GA, GAV, QL e QC)	QAV(QL e QC)	QAV(GAV)	QAC(QC)
336	340	335	339
344	348	343	349
337	358	353	378
392	363	358	390
402	389	364	392
411	392	374	394
	408	384	401
	452	400	411
		402	
		452	

4.4. PLS - Quantificação dos Contaminantes

Modelos PLS para quantificação dos teores dos contaminantes nas amostras de QAV foram feitos, os resultados obtidos para estes modelos estão apresentados na Tabela 6, e sua avaliação foi realizada a partir dos valores do coeficiente de determinação (R^2) da curva obtida entre os valores previstos versus medidos e os erros de validação cruzada e predição. Assim, quanto maior seu valor, maior a variabilidade explicada pelo modelo de regressão linear (Figura 21).

Os modelos PLS para quantificar os contaminantes deram bons resultados (Tabela 12), exceto para as amostras contaminadas com QL e QC, que apresentou baixos valores para os coeficientes de determinação. Cada contaminante apresentou uma condição de tamanho e janela e grau de polinômios ótimos, os deltas também variaram, sendo que os melhores modelos para os contaminantes GA, GAV e QL foi o $\Delta\lambda$ de 10 nm e para o QC foi o de 40 nm, tendo como pré-processamento apenas os dados centrados na média, quanto que para os demais foi necessário o uso do alisamento de Savitzky-Golay.

Tabela 12. Parâmetros de desempenho dos modelos PLS calculados para os modelos de quantificação de contaminantes em amostras de QAV.

Parâmetro	<i>GA</i>	<i>GAV</i>	<i>QL</i>	<i>QC</i>
<i>$\Delta\lambda$</i>	10	10	10	40
<i>J</i>	17	15	15	CM
<i>P</i>	4	2	2	CM
<i>RMSEC</i>	0,814	0,514	0,955	1,777
<i>RMSECV</i>	1,459	2,531	2,512	3,439
<i>RMSEP</i>	1,067	1,142	1,831	1,787
<i>R²_{CAL}</i>	0,938	0,979	0,937	0,741
<i>R²_{VAL}</i>	0,911	0,814	0,874	0,727
<i>VL</i>	4	6	5	5

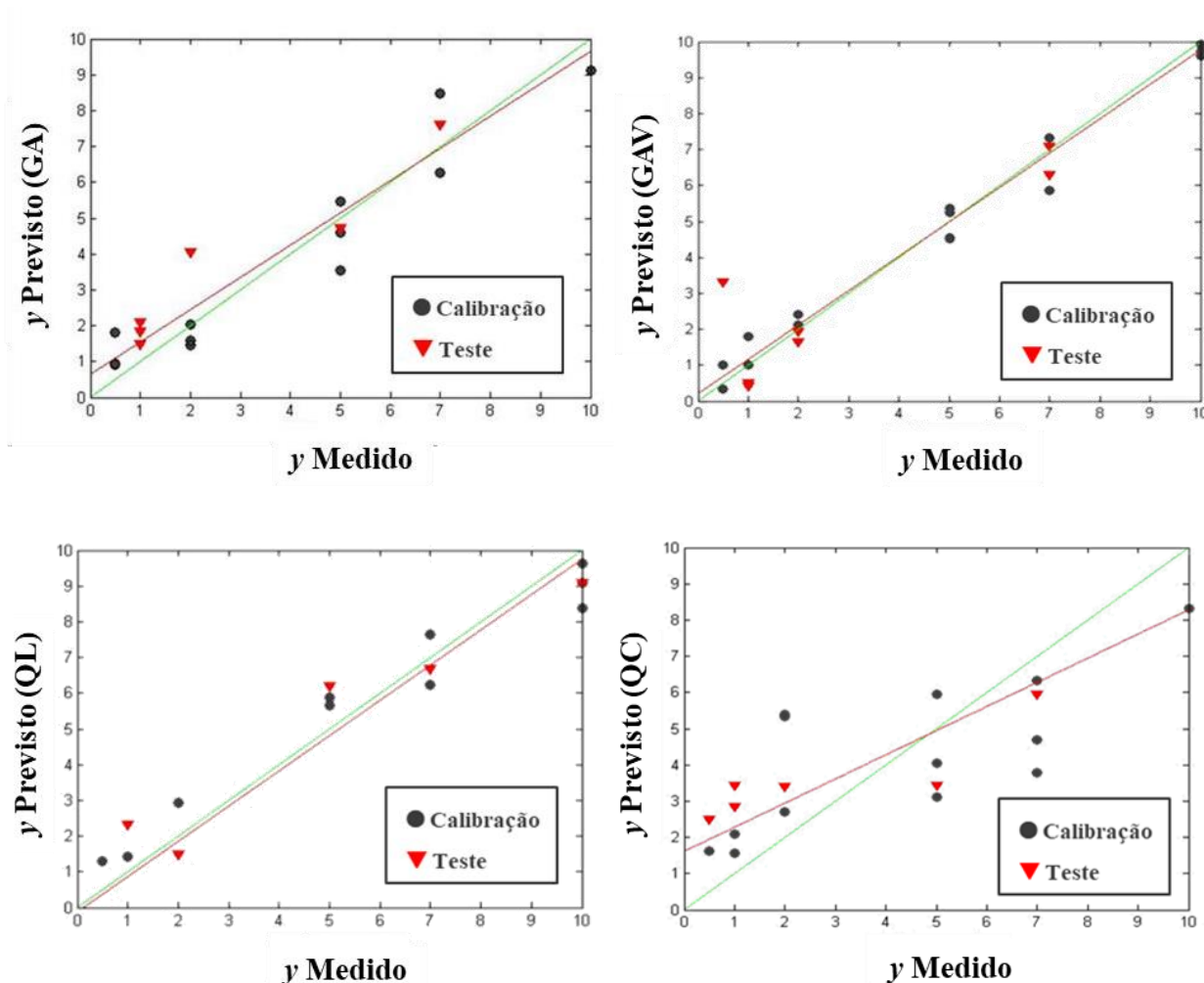


Figura 21. Gráficos dos modelos PLS para os y previstos pelo y medidos para as amostras de QAV.

Os comprimentos de onda mais importantes para quantificação dos contaminantes nas amostras de QAV dos modelos PLS estão dispostos na Tabela 13 assim como os gráficos de importância variável na projeção *VIP* scores (Figura 23). A atribuição desses comprimentos de onda específicos é difícil, pois o SFS não é uma técnica comumente usada para impressão digital, mas todos os modelos apresentaram picos em comum nas regiões de 330 a 340 nm e na região de 400 a 410 nm, sendo então as duas regiões de maior importância em comum para todos os modelos de classificação. As amostras contaminadas com GAV apresentou maior quantidade de picos com importância para a construção do modelo, pois o corante usado nesse tipo de contaminante que apresenta cor azul conhecido como 1,4 dialquilamino antraquinona [76] (Figura 22) um composto policromático que são grupos fluoróforos.

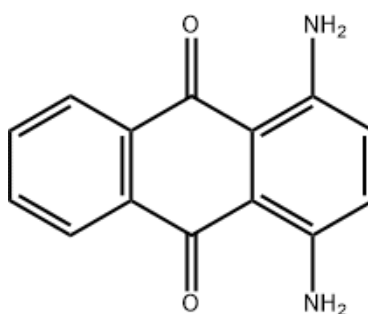


Figura 22. Estrutura química da 1,4 dialquilamino antraquinona.

Tabela 13. Valores dos comprimentos de onda (nm) presentes nos gráficos de importância variável na projeção *VIP* escores de todos os modelos PLS para quantificação de amostras de QAV contaminados.

QAV (GA)	QAV (GAV)	QAV (QL)	QAV (QC)
339	329	331	335
358	339	338	337
370	349	347	344
388	359	357	350
399	381	353	363
409	92	395	374
	400	410	377
	407	414	401
	417		
	434		
	452		

Os resultados apresentados na Tabela 13 mostram que a quantificação para os contaminantes GA e GAV, além de apresentar bons valores de R^2 tanto para o conjunto de calibração quanto para o conjunto de validação (0,938 e 0,911 para GA e 0,979 e 0,814 para GAV), apresentou erros com vales baixos (*RMSEC* e *RMSEP*) sendo respectivamente 0,814 e 1,067 para GA e 0,514 e 1,142 para GAV. Estes valores estão de acordo com a outros modelos obtidos por meio de técnicas diferentes (Tabela 14).

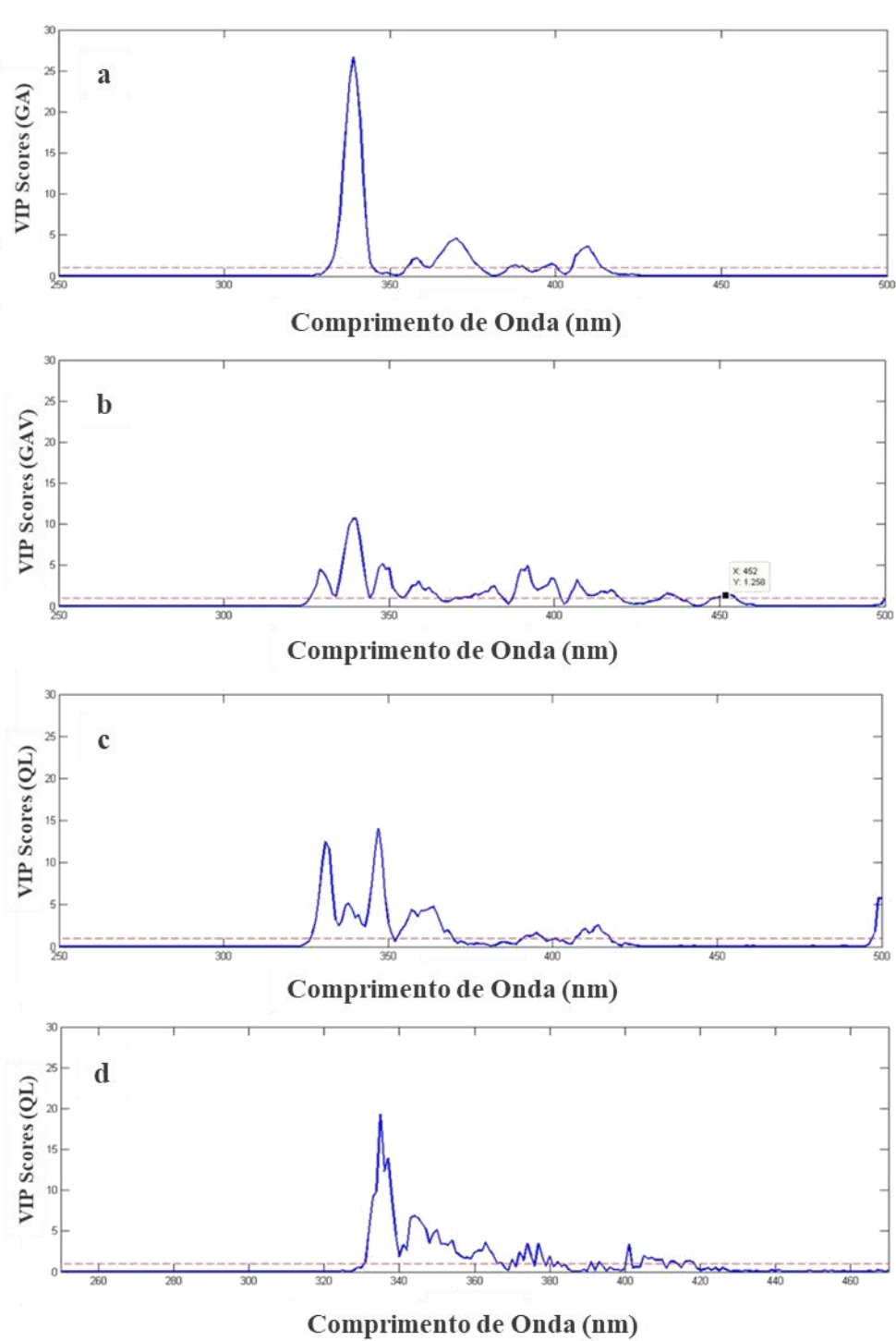


Figura 23. Gráfico das variáveis significativas da regressão (VIP scores) para os modelos de quantificação de contaminantes em QAV.

Tabela 14. Tabela com trabalhos encontrados na literatura.

<i>Autores</i>	<i>Técnica Espectroscópica</i>	<i>Técnica Quimiométrica</i>	<i>Tipo de combustível</i>	<i>Contaminantes</i>	<i>% contaminação</i>	<i>RMSEP</i>	<i>R²_{CAL}</i>	<i>R²_{VAL}</i>
Este trabalho	SFS	PLS	QAV	GA	0,5 a 10 %	1,067	0,938	0,911
				GAV		1,142	0,979	0,814
				QL		1,831	0,937	0,874
				QC		1,787	0,741	0,727
[68]	EEM	MCR-ALS	JET-A1	QC	2,0 a 100%	6,19	--	0,969
		PARAFAC				5,36	--	0,918
		PLS				12,13	--	0,978
[74]	MIR/NIR	PLS	JET-A1	QC	20 a 60%	1,64	--	0,995
[78]	EEM	PARAFAC PCA /PCR	JET-A1	Lubrificante de turbina	10 a 100 ppm	--	-- 0,978	--

4.5 CONCLUSÃO

- Os modelos para discriminação de contaminantes em amostras de QAV são eficientes para serem usados como “*screening*” auxiliando a rotina de fiscalização e de controle de qualidade do combustível de aviação.
- A técnica analítica é rápida e não destrutiva, fazendo uso de pouco volume de amostra (de 3 a 4 mL).
- O modelo de quantificação PLS apresentou bons valores de R^2 , sendo um modelo promissor na triagem do percentual de contaminantes em amostras de QAV.
- A abrangência de tipos de contaminantes e o uso de amostras reais de QAV de diferentes perfis, torna o trabalho robusto e representativo.

5.0 MODELO QUIMIOMÉTRICO PARA A PREVISÃO DO PONTO DE CONGELAMENTO DO QAV

5.1 Introdução

Devido à importância do querosene em aplicações de aviação comercial e militar, a atenção aqui está voltada para a avaliação de sua qualidade. Atualmente 26 testes físico-químicos diferentes são usados para caracterizar os combustíveis de aviação, a fim de garantir sua qualidade, evitando assim, problemas que vão desde a segurança no seu transporte, abastecimento de aeronaves e desempenho nos voos [79].

Todos os ensaios são realizados seguindo protocolos analíticos internacionais cobertos por especificações, alguns testes envolvem operações manuais que são bastante subjetivas, tediosas e propensas a erros operacionais. Dentre os vários ensaios previstos para o controle de qualidade do QAV, o Ponto de Congelamento é sem dúvida um dos mais importantes.

O Ponto de Congelamento de um combustível é definido como sendo a temperatura na qual cristais de hidrocarbonetos formados pelo resfriamento da amostra desaparecem quando está sujeita a aquecimento, sob agitação constante. Em termos práticos, representa a menor temperatura em que o QAV permanece livre de cristais que podem restringir o seu fluxo através de filtros em turbinas aeronáuticas [80-81].

Essa propriedade é influenciada pelo tipo de hidrocarboneto e pelo ponto final de ebulição da fração. Quanto maior o teor de compostos parafínicos e o ponto final de ebulição, maior será o Ponto de Congelamento. De acordo com a ANP o ponto de congelamento mínimo para o QAV é de -47 °C [2].

Alguns trabalhos usam análises quimiométricas juntamente com técnicas analíticas como cromatografia gasosa acoplada a espectrometria de massas (GC-MS), espectrofotometria no infravermelho próximo (NIR) [80], infravermelho médio com transformada de Fourier (FT-MIR) [79], dentre outras técnicas para prever as principais propriedades físico-químicas do QAV.

O objetivo deste estudo foi avaliar o uso a espectroscopia de fluorescência sincronizada usada em conjunto com análises quimiométricas (PLS-DA) para classificar amostras comerciais de QAV em relação aos seus valores de Ponto de congelamento.

5.2 Materiais e métodos

5.2.1. Amostras

O conjunto de amostras para o modelo PLS-DA contém 30 amostras no total, com os valores de Ponto de Congelamento variando na faixa de -42,0 a -56°C, cuja distribuição das amostras é mostrada na Tabela 15.

Tabela 15. Distribuição das 30 amostras de QAV, em diferentes faixas de valores do Ponto de Congelamento para os modelos PLS2-DA

Faixa (°C)	Número de Amostras	Classe
- 42 a - 44	10	R (Reprovada)
- 45 a - 49	10	C (Crítica)
- 50 a - 56	10	A (Aprovada)

5.2.2. Instrumentação

A instrumentação e metodologia utilizada foram as mesmas descritas no item 4.2.2 deste documento.

5.2.3. Ponto de Congelamento

O ensaio para determinar o Ponto de Congelamento foi realizado através da norma ASTM D2386, utilizando um analisador manual do tipo Stanhope-seta, mantido em temperatura ambiente de 15 a 35 °C e um termômetro de imersão total contendo as seguintes especificações da ASTM N° 114C (-80 a +20 °C) com resolução de 0,5 °C, calibrado por laboratório pertencente à RBC nos pontos 0, -40, -60, -75 °C e com registro de correção da calibração P-EQ-R6.

Para montagem do sistema, um vaso de vidro externo é colocado dentro do case de madeira. A essa case é fixado uma braçadeira que contém um grampo para prender o tubo de ensaio. Dentro desse tubo de ensaio é colocado um agitador fixado a uma rolha que também contém um colar anti-umidade, por onde será introduzido a corrente de gás nitrogênio ou ar comprimido.

Uma quantidade de aproximadamente 25 mL de amostra é colocada dentro desse tubo de ensaio, o tubo é então fechado com a rolha que contém a haste de agitação e o termômetro. É necessário se atentar para a posição do termômetro que deve ser ajustada para que seu bulbo fique a uma distância de 10 a 15 mm do fundo do tubo e não toque as paredes e ao anel da haste agitadora que deve passar dentro do termômetro, tomando cuidado para que no movimento de descida ele aproxime do fundo e na subida esteja abaixo da superfície da amostra. Ligar o agitador automático e observar a amostra continuamente para o aparecimento dos cristais/névoa, neste momento é feito a aferição e a temperatura é anotada.

Caso o ponto de cristalização não puder ser alcançado pelo resfriamento contínuo do combustível até a temperatura mais baixa atingida pelo equipamento (entre -65 e -70 °C), quando usado gelo seco e álcool como refrigerante, faz-se o registro da temperatura mensurável mais baixa que o combustível atingir.

Vale ressaltar que devido às baixas temperaturas, pode ocorrer a formação de um condensado no vaso externo dificultando a visualização e, para removê-lo, deve-se utilizar uma pisseta com acetona.

Para a rotina de realização deste ensaio, as amostras a serem analisadas devem estar em temperatura ambiente, e é necessário realizar uma verificação intermediária do ensaio de ponto de congelamento com um material de referência certificado. A periodicidade mínima para a verificação intermediária é semestral.

5.3. RESULTADOS E DISCUSSÃO

5.3.1. PLS2-DA classificação entre as classes de QAV de acordo com as faixas de Ponto de congelamento (PC)

De acordo com a ANP, o valor mínimo do ponto de congelamento para uma amostra ser dita conforme é de no mínimo -47 °C de acordo com a norma ASTM D 2386, a reprodutibilidade do teste pode ser de até 2,5 °C. Sendo assim, foram estabelecidas as faixas de separação de classes para a construção do modelo PLS2-DA visando discriminar as amostras de QAV de acordo com seus valores de ponto de congelamento. A primeira classe (R) contém as amostras em uma faixa de PC não conforme, a segunda classe (C) contém as amostras em uma faixa crítica, ou seja, bem próxima do valor de referência que é de -47 °C e a terceira classe (A) compõe as amostras em uma faixa bem abaixo do valor de referência, logo seriam as amostras conformes.

O modelo que apresentou os melhores valores para as figuras de mérito como *SEM*, *SPE*, *TFN*, *TFP* e *EEF* foi o $\Delta\lambda$ de 10 nm, pré-processados usado foi com janela de 17 pontos e polinômio de 4ª ordem, foi construído utilizando nove variáveis latentes, que responderam por 99,9% e 87,67% da variância total nos blocos *X* e *Y*.

Foram usadas um total de 21 amostras no conjunto de treinamento (sete amostras para cada classe) e no conjunto de teste e nove amostras (três para cada classe). As amostras foram separadas nestes dois conjuntos através do algoritmo de Kennard-Stone.

O modelo apresentou uma taxa de eficiência de 100% para as três classes (Tabela 16), mesmo sendo construído com uma pequena quantidade de amostras, o modelo possui uma ampla faixa de valores para os PC, podendo ser usando como um “*screening*” na hora de selecionar amostras que serão analisadas pelo ensaio padronizado da ANP.

Na Figura 26 temos as previsões de *Y* para cada classe em relação as demais, assim como os valores de limite de decisão (*threshold*) obtidos para o modelo.

Tabela 16. Parâmetros de desempenho calculados para o modelo PLS2-DA construído com espectros de fluorescência síncrona registrados em $\Delta\lambda$ de 10 nm para classificar QAV em diferentes faixas de PC.

CLASSES	TREINAMENTO				
	<u>SEM</u>	SPE	TFN	TFP	EEF
R, C e A	1	1	0	0	100
CLASSES	TESTE				
	<u>SEM</u>	SPE	TFN	TFP	EEF
R, C e A	1	1	0	0	100

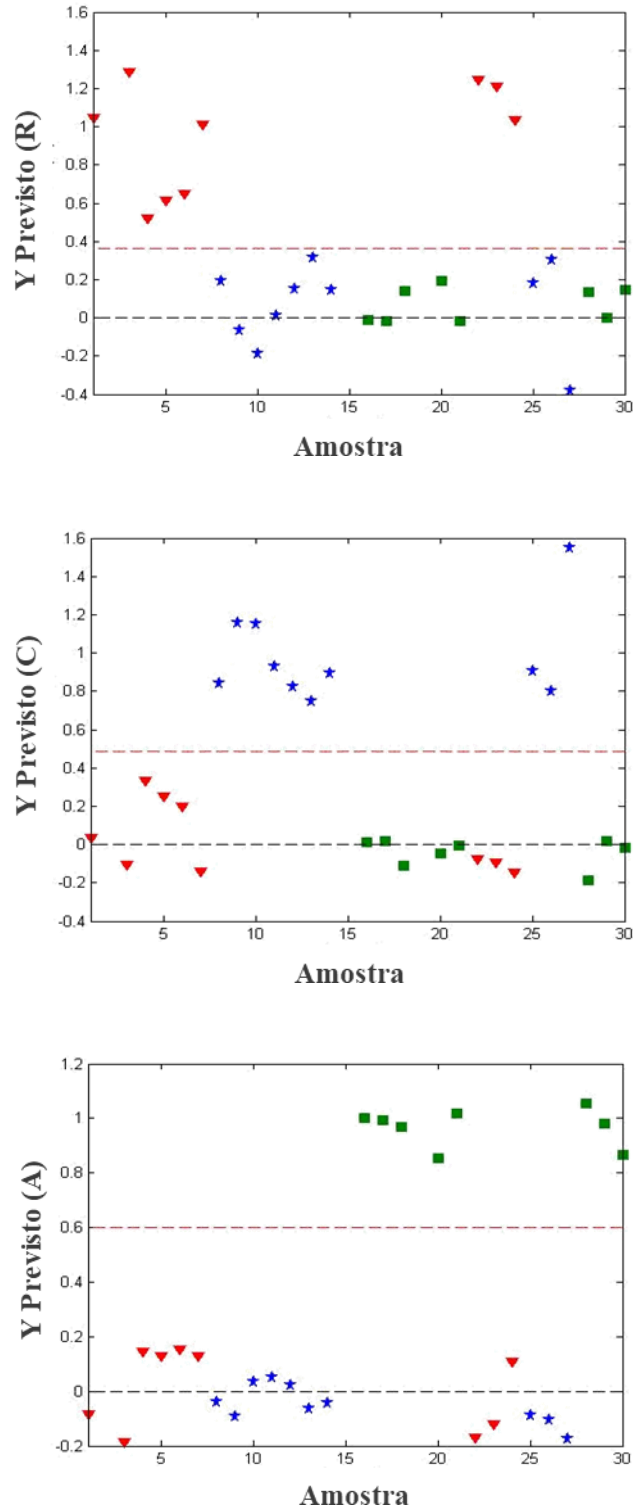


Figura 24. PLS2-DA para discriminação de QAV em diferentes faixas de Ponto de Congelamento. Linhas tracejadas verticais separam amostras de treinamento e teste, e linhas tracejadas horizontais indicam a classe limite.

Os comprimentos de onda mais importantes para a classificação dos modelos PLS2-DA estão dispostos na Tabela 17 assim como os gráficos de importância variável na projeção *VIP* escores (Figura 27). A atribuição desses comprimentos de onda específicos é difícil, pois o SFS não é uma técnica comumente usada para impressão digital, mas todos os modelos apresentaram picos em comum nas regiões de 325 a 355 e de 400 a 445 nm, sendo então as duas regiões de maior importância em comum para todos os modelos de classificação, já o para a classificação das amostras que pertencem a faixa de pontos de congelamentos abaixo de $-52\text{ }^{\circ}\text{C}$, ou seja as da classe das Aprovadas, apresentou um pico diferente das demais o 464 nm, podendo ser um diferencial dessa classe.

De acordo com os resultados apresentados para o modelo PLS2-DA para classificação quanto as faixas do Ponto de congelamento, e comparado com os poucos trabalhos que usam o PLS-DA para discriminar amostras quanto as suas propriedades físico-químicas (Tabela 18), o modelo mostrou-se eficiente como uma ferramenta auxiliar para otimização de uma rotina laboratorial, fazendo uma pré-seleção das amostras que realmente necessitam do ensaio normalizado pela ANP.

Tabela 17. Valores dos comprimentos de onda (nm) presentes nos gráficos de importância variável na projeção *VIP* escores para os modelos PLS2-DA para os faixas de PC.

R	C	A
328	328	328
333	333	333
341	343	342
351	351	351
383	380	391
408	394	445
445	408	464
	445	

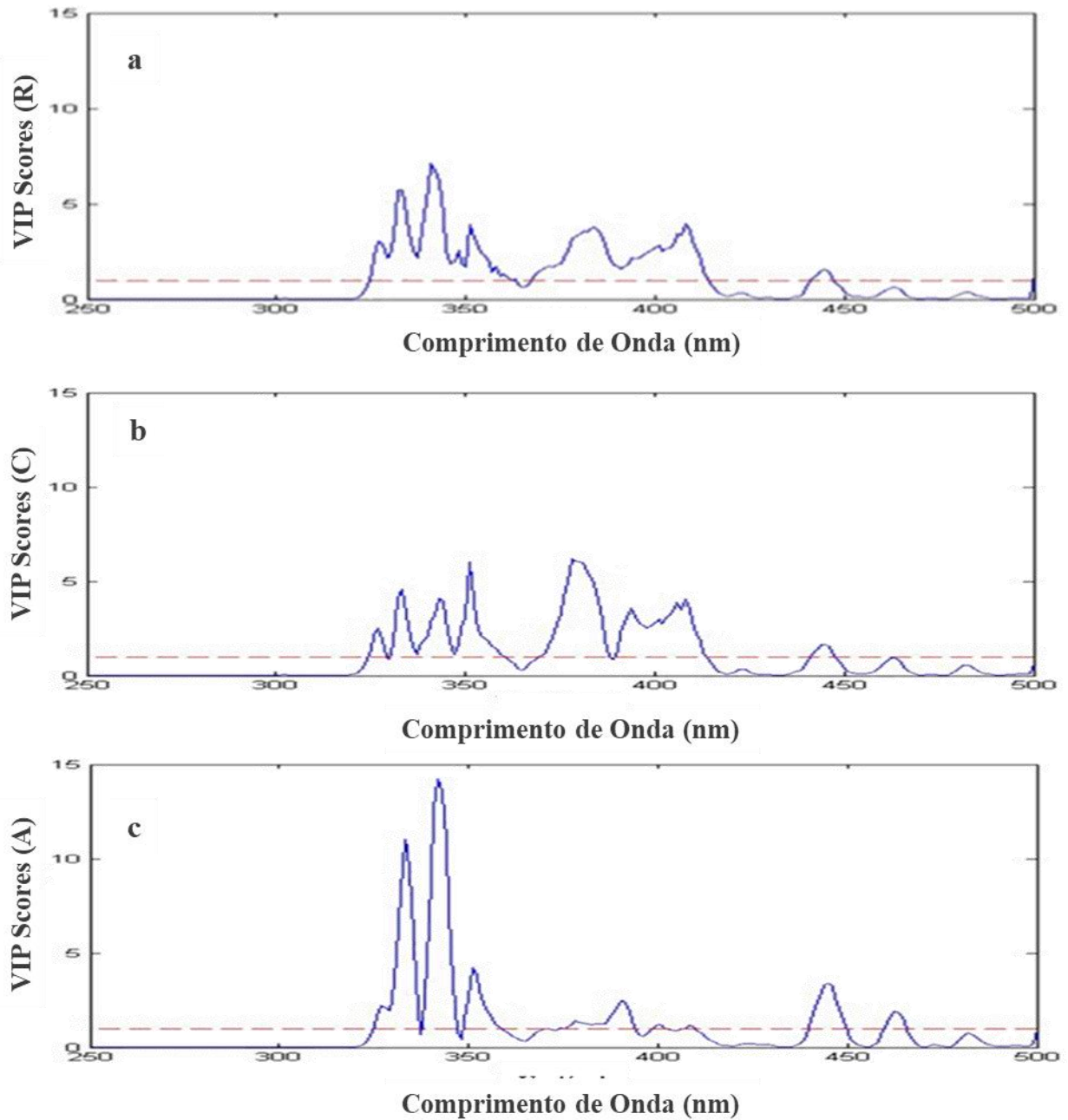


Figura 25. Gráficos VIP escores dos modelos PLS2-DA de fluorescência síncrona para classificação das amostras em diferentes faixas de Ponto de Congelamento. (a) classe das amostras Reprovadas, (b) classe das amostras com valores Críticos e (c) classe das amostras Aprovadas.

Tabela 18. Tabela com trabalhos encontrados na literatura que avaliaram o ponto de congelamento de combustível de aviação, utilizando diferentes técnicas.

<i>Autores</i>	<i>Técnica Espectroscópica</i>	<i>Técnica Quimiométrica</i>	<i>Tipo de Combustível</i>	<i>Faixa (°C)</i>	<i>% EEF</i>	<i>R²_{CAL}</i> <i>R²_{VAL}</i>
Este trabalho	EFS	PLS2-DA	QAV Comercial	-42 a -53	100%	--
[83]	NIR	PLS-DA SVM	JET-A JET-A1	-42 a -65	91% 92%	--
[84]	NIR	PLS	JET- A1	ND	--	0,968 0,952
[85]	NIR	PLS	JET- A1	-42 a -60	--	0,850
[86]	NIR	PLS	JET -A1	Média de -49,5	--	0,890

5.4 CONCLUSÃO

- Os modelos para classificar as amostras de QAV em relação a faixa de Ponto de Congelamento são eficientes para serem usados como “*screening*” auxiliando a rotina de fiscalização e de controle de qualidade do combustível de aviação.
- A técnica analítica é rápida e não destrutiva, fazendo uso de pouco volume de amostra.
- A baixa quantidade de amostra torna o modelo menos representativo, mesmo se tratando de amostras reais.

6. CONCLUSÃO GERAL

Este trabalho demonstrou a aplicabilidade e eficiência do uso da espectrofluorimetria sincronizada associada a técnicas quimiométricas (PLS e PLS-DA) na previsão e classificação de parâmetros físico-químicos relevantes para o controle de qualidade de combustíveis, com destaque para o biodiesel (B100) e o querosene de aviação (QAV). A partir da construção de modelos quimiométricos robustos, foi possível prever a estabilidade oxidativa do B100 com bons índices de desempenho ($R^2_{val} = 0,822$) e identificar diferentes tipos de contaminação em QAV, com modelos de regressão que alcançaram R^2_{val} superiores a 0,9 para contaminantes como gasolina e querosene iluminante. Além disso, os modelos de classificação obtiveram taxas de eficiência acima de 80% para a identificação do ponto de congelamento e dos tipos de contaminantes presentes nas amostras. Os resultados obtidos evidenciam que a metodologia proposta, além de ser rápida e de baixo custo, não exige preparo de amostra e apresenta alto potencial para ser utilizada como ferramenta de triagem (*screening*) nos processos de fiscalização e monitoramento da qualidade dos combustíveis, contribuindo significativamente para a identificação precoce de não conformidades. Dessa forma, o trabalho contribui para a otimização das rotinas laboratoriais, a redução de custos operacionais e o aumento da confiabilidade dos resultados analíticos, alinhando-se às exigências regulatórias estabelecidas pela ANP e às demandas por inovação na área de controle de qualidade de combustíveis.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Brasil. Lei nº 9.478, de 6 de agosto de 1997. Dispõe sobre a política energética nacional, as atividades relativas ao monopólio do petróleo, institui o Conselho Nacional de Política Energética e a Agência Nacional do Petróleo e dá outras providências. Disponível em: <http://www.planalto.gov.br>. Acesso em: 28 março 2025.
- [2] ANP. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Resolução ANP nº856, de 22 de outubro de 2021. Disponível em: < <https://www.in.gov.br/en/web/dou/-/resolucao-anp-n-856-de-22-de-outubro-de-2021-354349404>>. Acesso em: 21 jun. 2024.
- [3] ANP. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Resolução ANP nº920, de 04 de abril de 2023. Disponível em: < <https://www.gov.br/anp/pt-br/assuntos/producao-e-fornecimento-de-biocombustiveis/biodiesel/especificacao-do-biodiesel>>. Acesso em: 21 jun. 2024.
- [4] Barra, I.; Kharbach, M.; Qannari, E. M.; Hanafi, M.; Cherrah, Y.; Bouklouze, A., **Predicting cetane number in diesel fuels using FTIR spectroscopy and PLS regression**, *Vibrational Spectroscopy*, (2020). <https://doi.org/10.1016/j.vibspec.2020.103157>.
- [5] Chakravarthy, R.; Acharya, C.; Savalia, A.; Naik, GN.; Das, AK.; Saravanan, C.; Gudasi, KB., **Property Prediction of Diesel Fuel Based on the Composition Analysis Data by two-Dimensional Gas Chromatography**. *Energy & Fuels*, 32, (2018), 3760-3774.
- [6] Máquina, A. D. V.; Siteo, B. V.; Santana, F. B.; Santos, D. Q.; Borges Neto, W., **Determination of adulteration of the B10 blend of diesel and crambe biodiesel using proton nuclear magnetic resonance (¹H NMR) spectroscopy with a data driven soft independent modeling of class analogy (DD-SIMCA) model**. *Analytical Letters*, 54, (2021), 790 – 801.
- [7] Wang, S.; Cheng, Q.; Yuan, Y.; Wang, C.; Ma, S., **Determination of adulterated gasoline using fluorescence emission-excitation matrices and multivariate calibration**. *Spectrochim. Acta Mol. Biomol. Spectrosc*, 210, (2019), 260 – 265.
- [8] Sikorska, E. **Fluorescence spectroscopy and chemometrics in analysis of beverages**. *In Quality Control in the Beverage Industry: The Science of Beverages*, edition no. 1; A.M. Grumezescu, A.M. Holban, Eds.; Elsevier: Amsterdam, Netherlands, 161, (2019), 161-203. <https://doi.org/10.1016/B978-0-12-816681-9.00005-9>.
- [9] Sádecká, J.; Tóthová, J. **Fluorescence spectroscopy and chemometrics in the food classification a review**. *Czech J Food Sci*, 25, (2007), 159-173. <https://doi.org/10.17221/687>.
- [10] Flumignan, D.L. et al. **Multivariate calibrations in gas chromatographic profiles for prediction of several physicochemical parameters of Brazilian commercial gasoline**. *Chemometrics and Intelligent Laboratory Systems*, 92, (2008), 53-60.
- [11] Cooper, J.B. **Chemometric analysis of Raman spectroscopic data for process control applications**. *Chemometrics and Intelligent Laboratory Systems*, 46, (1999), 231-247.
- [12] Patra, D., Mishra, A. K. **Study of diesel fuel contamination by excitation emission matrix spectral subtraction fluorescence**. *Analytica Chimica Acta*, 454, (2002), 209-215.

- [13] John, P.; Soutar, I., **Identification of crude oils by synchronous excitation spectrofluorimetry.** *Analytical Chemistry*, vol. 48, no 3, (1976), 520-524.
- [14] Vo-Dinh, T. **Multicomponent analysis by synchronous luminescence spectrometry.** *Analytical Chemistry*, vol. 50, no 3, (1976), 396-401.
- [15] Patra, D., Mishra, A. K. **Study of diesel fuel contamination by excitation emission matrix spectral subtraction fluorescence.** *Analytica Chimica Acta*, 454, (2002), 209-215.
- [16] Patra, D.; Mishra, A. K. **Effect of sample geometry on synchronous fluorimetric analysis of petrol, diesel, kerosene and their mixtures at higher concentration.** *Analyst*, 125, (2000), 1383-1386.
- [17] Taksande, A.; Hariharan, C. **Synchronous fluorescence method to check adulteration of petrol and diesel by kerosene.** *Spectroscopy Letters*, 39, (2006), 345-356.
- [18] Patra, D. e Mishra, A. K., **Investigation on Simultaneous Analysis of Multicomponent Polycyclic Aromatic Hydrocarbon Mixtures in Water Samples: a Simple Synchronous Fluorimetric Method.** *Talanta*, 558, (2001), 143-153.
- [19] Wang, Z. et al. **Application of excitation-emission matrix fluorescence spectroscopy and chemometrics for quantitative analysis of emulsified oil concentration.** *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 328, (2025). <https://doi.org/10.1016/j.saa.2024.125423>.
- [20] Meira, M. et al. **Determination of the oxidation stability of biodiesel and oils by spectrofluorimetry and multivariate calibration.** *Talanta*, 85, (2011), 430-434.
- [21] Divya, O., Mishra, A.K. **Combining synchronous fluorescence spectroscopy with multivariate methods for the analysis of petrol–kerosene mixtures.** *Talanta*, 72, (2007), 43-48.
- [22] Patra, D.; Mishra, A. K., **Concentration Dependent Red Shift: Qualitative and Quantitative Investigation of Motor Oils by Synchronous Fluorescence Scan.** *Talanta*, 53, (2001), 783-790.
- [23] Abbas, O. et al. **Application of chemometric methods to synchronous UV fluorescence spectra of petroleum oils.** *Fuel*, 85, (2006), 2653-2661.
- [24] Li, J. et al. **Matching fluorescence spectra of oil spills with spectra from suspect sources.** *Analytica Chimica Acta*, 514, (2004), 51-56.
- [25] Corgozinho, C.N.C., Barbeira, P.J.S. **Quantification of Biodiesel in Biodiesel–Diesel Blends Using Spectrofluorimetry and Multivariate Calibration.** *Journal of Analytical Chemistry*, 70, (2015), 936-942.
- [26] Barros Neto, B. et al. **25 Anos de Quimiometria no Brasil.** *Quimica nova*, 29, (2006), 1401-1406.
- [27] Almeida, Mariana R. **Espectroscopia Raman e quimiometria como ferramentas analíticas para química forense e paleontologia.** *Tese (Doutorado em Química) – UNICAMP*, Campinas,

2015.

[28] Dunn, Kevin G. **Process Improvement using Data**. 2023, 381. Disponível em: <<https://learnche.org/pid/PID.pdf>>. Acesso em: 18 de dezembro 2024.

[29] Wise, B.M.; Gallagher, N.B.; Bro, R.; Shaver, J.M.; Windig, W.; Koch, R.S. **Chemometrics Tutorial for PLS_Toolbox and Solo**, *Eingenvector Reserch, Inc., Wenatchee*, 2006.

[30] Gondim, C. S.; Junqueira, R. G.; Souza, S. V. C. **Trends in implementing the validation of qualitative methods of analysis**. *Revista Instituto Adolfo Lutz*. 70, (2011), 433-447.

[31] Botelho, B. G.; Reis, N.; Oliveira, L. S.; Sena, M. M. **Development and analytical validation of a screening method for simultaneous detection of five adulterants in raw milk using mid-infrared spectroscopy and PLS-DA**. *Food Chemistry*. 181,(2015), 31-37.

[32] Johnson, R. A., Wichern, D. W., *Applied Multivariate Statistical Analysis* – second edition, Prentice Hall, Englewood Cliffs, New Jersey, 1988.

[33] Breitreitz, M.C.; Souza, A.M.; Poppi, R.J. **Experimento didático de quimiometria para planejamento de experimentos: avaliação das condições experimentais na determinação espectrofotométrica de ferro II com o-fenantrolina. Um tutorial, parte III**. *Quim Nova*, 37(3) (2014) 564-573. <http://dx.doi.org/10.5935/0100-4042.20140092>.

[34] Barros Neto, B.; Scarminio, I.S.; Bruns, R.E. **Como fazer experimentos. Pesquisa e desenvolvimento na ciência e na indústria**. *Editora Unicamp*: Campinas, 2001.

[35] Biodiesel Market Size, Share & Trends Analysis Report By Feedstock (Vegetable Oils, Animal Fats), By Application (Fuel, Power Generation), By Region (Europe, APAC), And Segment Forecasts, 2022 – 2030, 2022.

[36] Ramos, L. P.; da Silva, F. R.; Mangrich, A. S.; e Cordeiro C. S. **Biodiesel Production Technologies**, *Revista Virtual de Química*, vol. 3, no. 5, (2011). doi: 10.5935/1984-6835.20110043.

[37] Nair, J. N.; Deepthi, J.; Kalyani, S. K., **Study of biodiesel blends and emission characteristics of biodiesel**, *Int J Innov Res Sci Eng Technol*, vol. 3297, (2007), 2319–8753.

[38] BRASIL. Conselho Nacional de Política Energética. Resolução CNPE nº 3, de 20 de março de 2023. Disponível em: <<https://www.gov.br/mme/pt-br/assuntos/conselhoscomites/cnpe/resolucoes-do-cnpe/2023>>. Acesso em: 15 jun. 2025.

[39] Brasil, “Lei nº 11.097, de 13 de janeiro de 2005”. Brasília, 13 jan. 2005.

[40] Suarez, P. A. Z; and Abreu, F. R, **O biodiesel no Brasil**, *Senatus*, vol. 4, no. 1, (2005), 49–52.

[41] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, “Anuário estatístico brasileiro do petróleo, gás natural e biocombustíveis 2022,” Rio de Janeiro, 2022. Disponível em: <https://www.gov.br/anp/pt-br/centrais-de-conteudo/publicacoes/anuario-estatistico/arquivos/anuario-estatistico-2022/anuario-2022.pdf>. Acesso em: 28 janeiro 2025.

[42] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, “Painel Dinâmico Produtores de Biodiesel.

[43] Lin, C.Y; Chiu, C, **Effects of Oxidation during Long-term Storage on the Fuel Properties of Palm Oil-based Biodiesel**, *Energy & Fuels*, vol. 23, (2009), 3285–3289. doi: 10.1021/ef900105t.

[44] Meher,L;Vidyasagar, D.I; Naik,S, **Technical aspects of biodiesel production by transesterification—a review**, *Renewable and Sustainable Energy Reviews*, vol. 10, no. 3, (2006), 248–268. doi: 10.1016/j.rser.2004.09.002.

[45] Lobo, I. P; Ferreira, S. L. C; Cruz, R. S, **Biodiesel: Parâmetros de qualidade e métodos analíticos**, *Quim Nova*, vol. 32, no. 6, (2009). Disponível em: <https://doi.org/10.1590/S0100-40422009000600044>. Acesso em: 22 outubro 2023.

[46] Dunn, R. O, **Effect of oxidation under accelerated conditions on fuel properties of methyl soyate (biodiesel)**, *J Am Oil Chem Soc*, vol. 79, no. 9, (2002), 915–920. doi: 10.1007/s11746-002-0579-2.

[47] Neto, R. C. R. *et al.*, **Thermo-oxidative stability of mineral naphthenic insulating oils: Combined effect of antioxidants and metal passivator**, *Ind Eng Chem Res*, vol. 43, no. 23, (2004), 7428–7434. doi: 10.1021/ie049645o.

[48] EN 14112 Fat and oil derivatives – fatty acid methyl esters (FAME) Determination of oxidation stability (accelerated oxidation test).

[49] EN 15751 Automotive Fuels, Fatty Acid Methyl Ester (FAME) Fuel and blends with Diesel Fuel – Determination of Oxidation Stability by Accelerated Oxidation Method.

[50] Tsuchiya, T.; Shiotani, H.; Goto, S.; Sugiyama, G. and Maeda, A., **Diesel Fuel Containing 5% FAME: Investigation of Acid Generation in FAME Blended Diesel Fuels and Its Impact on Corrosion**, *Japanese Standards*, (2006). doi: 10.4271/2006-01-3303.

[48] Comité Européen de Normalisation, **Fat and oil derivatives - Fatty Acid Methyl Esters (FAME) - Determination of oxidation stability (accelerated oxidation test)**. Bruxelas, 2016.

[51] Velvarská, R.; Vráblík, A.; Fiedlerová, M. and Černý, R., **Near-infrared spectroscopy for determining the oxidation stability of diesel, biodiesel and their mixtures**, *Chemical Papers*, vol. 73, (2019), 2987–2993. doi: 10.1007/s11696-019-00852-4.

[52] Canha, N.; Felizardo, P.; Menezes, J.C. and Correia, M., **Multivariate near infrared spectroscopy models for predicting the oxidative stability of biodiesel: Effect of antioxidants addition**, *Fuel*, vol. 97, (2012), 352–357. doi: 10.1016/j.fuel.2012.02.017.

- [53] Varghese, G.; Saeed, K. and Rutt, K.J., **Determination of the oxidative stability of biodiesel fuels by near-infrared spectroscopy**, *Fuel*, vol. 290, (2021). doi: 10.1016/j.fuel.2020.120015.
- [54] De Lira, L. F. B. *et al.*, **Infrared spectroscopy and multivariate calibration to monitor stability quality parameters of biodiesel**, *Microchemical Journal*, vol. 96, no. 1, (2010), 126–131. doi: 10.1016/j.microc.2010.02.014.
- [55] Pavia, D. L.; Lampman, G.M.; Kriz, G.S. and Vyvyan, J.R., **Introduction to spectroscopy**, 5th ed. *Stamford: Cengage Learning*, 2015.
- [56] John, P. and Soutar, I., **Identification of crude oils by synchronous excitation spectrofluorimetry**, *Anal Chem*, vol. 48, no. 3, (1976), 520–524. doi: 10.1021/ac60367a013.
- [57] Insausti, M. *et al.*, **Screening analysis of biodiesel feedstock using UV–vis, NIR and synchronous fluorescence spectrometries and the successive projections algorithm**, *Talanta*, vol. 97, (2012), 579–583. doi: 10.1016/j.talanta.2012.04.056.
- [58] Insausti, M.; Romano, C.; Pistonesi, M. F. and Band, B. S. F., **Simultaneous determination of quality parameters in biodiesel/diesel blends using synchronous fluorescence and multivariate analysis**, *Microchemical Journal*, vol. 108, (2013), 32–37. doi: 10.1016/j.microc.2012.12.007.
- [59] Kumar, K. and Mishra, A. K., **Quantification of Ethanol in Ethanol-Petrol and Biodiesel in Biodiesel-Diesel Blends Using Fluorescence Spectroscopy and Multivariate Methods**, *J Fluoresc*, vol. 22, no. 1, (2012), 339–347. doi: 10.1007/s10895-011-0964-6.
- [60] Magalhães, K. F.; Caires, A. R. L.; Silva, M. S.; Alcantara, G. B. and Oliveira, S. L., **Endogenous fluorescence of biodiesel and products thereof: Investigation of the molecules responsible for this effect**, *Fuel*, vol. 119, (2014), 120–128. doi: 10.1016/j.fuel.2013.11.024.
- [61] Sikorska, E.; Górecki, T.; Khmelinskii, I. V.; Sikorski, M. and Koziół, J., **Classification of edible oils using synchronous scanning fluorescence spectroscopy**, *Food Chem*, vol. 89, no. 2, (2005), 217–225. doi: 10.1016/j.foodchem.2004.02.028.
- [62] De Santana *et al.*, **Experimento didático de quimiometria para classificação de óleos vegetais comestíveis por espectroscopia no infravermelho médio combinado com análise discriminante por mínimos quadrados parciais: Um tutorial, parte V**, *Quim Nova*, 43, no. 3, (2020), 371–381. doi: 10.21577/0100-4042.20170480.
- [63] Silveira, A. L.; Barbeira, P. J. S., **Synchronous fluorescence spectroscopy and multivariate classification for the discrimination of cachaças and rums**, *Spectrochim Acta Mol Biomol Spectrosc*, 270, (2022). doi: 10.1016/j.saa.2021.120821.
- [64] Vivó-Truyols, G. and Schoenmakers, P.J., **Automatic Selection of Optimal Savitzky–Golay Smoothing**, *Anal Chem*, 78, no. 13, (2006), 4598–4608. doi: 10.1021/ac0600196.
- [65] Moreira, L.; d’Avila, L. & Azevedo, D., **Automotive Gasoline Quality Analysis by Gas Chromatography: Study of Adulteration**. *Chromatographia*, 58, (2003), 501–505, <https://doi.org/10.1365/s10337-003-0065-z>.
- [66] Mendes, G.; Barbeira, P. J. S., **Detection and quantification of adulterants in gasoline**

using distillation curves and multivariate methods. *Fuel*, 112, (2013), 163-171.
<https://doi.org/10.1016/j.fuel.2013.04.077>.

[67] Ocorrências Aeronáuticas Na Aviação Cível Brasileira, PAINEL SIPAER , Disponível em: <https://painelsipaer.cenipa.fab.mil.br/extensions/Sipaer/acidentes.html>. Acesso em: 20 de janeiro de 2025.

[68] Câmara, A. B. F.; Da Silva, W. J. O.; Neves, A. C. O.; Moura, H.O. M. A.; De Lima, K. M. G.; De Carvalho, L.S., **Excitation-emission fluorescence spectroscopy coupled with PARAFAC and MCR-ALS with area correlation for investigation of jet fuel contamination.** *Talanta*, 266, (2024). <https://doi.org/10.1016/j.talanta.2023.125126>.

[69] Carvalho, I., Consumo gerenciado. 2013. Disponível em: https://aeromagazine.uol.com.br/artigo/consumo-gerenciado_809.html. Acesso em: 13 jul. 2024.

[70] ANP. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Evolução da produção e importação querosene de aviação no Brasil. 2024. Disponível em: <http://www.anp.gov.br/?id=561>. Acesso em 14 de março. de 2025.

[71] Pereira, R.C.C.; Skrobot, V. L.; Castro, E. V. R.; Fortes, I. C. P.; Pasa, V. M. D., **Determination of Gasoline Adulteration by Principal Components Analysis-Linear Discriminant Analysis Applied to FTIR Spectra.** *Energy & Fuels*, 20, (2006), 1097-1102.

[72] ASTM – American Society for Testing Materials. ASTM D 86, Distillation of Petroleum Products at Atmospheric Pressure – Significance, Instrumentation and Standard Test. Disponível em: <https://analytical.com/methods/d86/>. Acesso em: 20 de jul. de 2024.

[73] De Paulo, J. M.; Barros, J. E. M. and Barbeira, P. J. S., **Differentiation of Gasoline Samples Using Flame Emission Spectroscopy and Partial Least Squares Discriminate Analysis.** *Energy & Fuels*, 28, (2014), 4355-4361.

[74] Câmara, A. B. F. et al. **Multivariate strategy for identifying and quantifying jet fuel contaminants by MCR-ALS/PLS models coupled to combined MIR/NIR spectra.** *Analytical And Bioanalytical Chemistry*, 414 (2022), 7897-7909.

[75] 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*, 83, (2004), 917-923.

[76] ANP. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Resolução ANP n°901, de 18 de novembro de 2022. Disponível em: <https://www.legisweb.com.br/legislacao/?id=438718>. Acesso em: 23 abr. de 2024.

[77] Barker, M.; Rayens; W. **Partial least squares for discrimination.** *Journal of Chemometrics*. 17, (2003) 166-173.

[78] Andrews, N.L.P.; Fan, J.Z.; Omrani, H.; Dudelzak, A.; Loock, H. P., **Comparison of lubricant oil antioxidant analysis by fluorescence spectroscopy and linear sweep voltammetry.** *Tribology International*, 94, (2016), 279-287.

[79] Gómez-Carracedo, M. P.; Andrade, J. M.; Calviño, M.; Fernández, E.; Prada, D.; Muniategui, S., **Multivariate prediction of eight kerosene properties employing vapour-phase mid-**

infrared spectrometry. *Fuel*,82, (2003), 1211-1218.

[80] ASTM – American Society for Testing Materials. ASTM D 2386 – 19 , Standard Test Method for Freezing Point of Aviation Fuels. Disponível em: <https://compass.astm.org/document/?contentCode=ASTM%7CD2386-19%7Cen-S&proxycl=https%3A%2F%2Fsecure.astm.org&fromLogin=true> . Acesso em: 20 de jul. de 2023.

[81] ASTM – American Society for Testing Materials. ASTM D 5972 – 22, Standard Test Method for Freezing Point of Aviation Fuels (Automatic Phase Transition Method. Disponível em: <https://compass.astm.org/document/?contentCode=ASTM%7CD5972-23%7Cen-US&proxycl=https%3A%2F%2Fsecure.astm.org&fromLogin=true> . Acesso em: 20 de jul. de 2023.

[82] Johnson, K. J.; Morris, R. E.; Rose-Pehrsson, S. L., **Evaluating the Predictive Powers of Spectroscopy and Chromatography for Fuel Quality Assessment.** *Energy & Fuels*, 20, (2006), 727-733.

[83] Xu, Z.; Bunker, C. E.; Harrington, P. DE B., **Classification of Jet Fuel Properties by Near-Infrared Spectroscopy Using Fuzzy Rule-Building Expert Systems and Support Vector Machines.** *Applied Spectroscopy*. 64, (2010), 1251-1258. doi:[10.1366/000370210793335115](https://doi.org/10.1366/000370210793335115).

[84] Li, K.; Zhang, X.; Zhang, j.; Du, B.; Song, X.; Wang, G.; Li, Q.; Zhang, Y.; Liu, F.; Zhang, Z., **Simultaneous Rapid Detection of Multiple Physicochemical Properties of Jet Fuel Using Near-Infrared Spectroscopy.** *ACS Omega*, 9, (2024), 16138–16146.

[85] Brouillette, C.; Smith, W.; Shende, C.; Gladding, Z.; Farquharson, S.; Morris, R. E.; Cramer, J. A.; Schmitgal, J., **Analysis of Twenty-Two Performance Properties of Diesel, Gasoline, and Jet Fuels Using a Field-Portable Near-Infrared (NIR) Analyzer.** *Applied Spectroscopy*. 70, (2016), 746-755.

[86] 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.** *Chemometrics and Intelligent Laboratory Systems*. 110, (2012), 64-73.