

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

**ESTUDO DA VIABILIDADE DO USO DE ESPECTROSCOPIA POR
FLUORESCÊNCIA 2D PARA QUANTIFICAR TEOR DE ENXOFRE
EM ÓLEO DIESEL**

DISSERTAÇÃO DE MESTRADO

Lucas Ranzan

Porto Alegre

2014

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

**ESTUDO DA VIABILIDADE DO USO DE ESPECTROSCOPIA POR
FLUORESCÊNCIA 2D PARA QUANTIFICAR TEOR DE ENXOFRE
EM ÓLEO DIESEL**

LUCAS RANZAN

Dissertação de Mestrado apresentada como requisito parcial para obtenção do título de Mestre em Engenharia

Área de concentração: Pesquisa e Desenvolvimento de Processos

Linha de Pesquisa: Engenharia de Sistemas – Projeto, Modelagem, Controle e Otimização de Processos

Orientador:

Prof. Dr. Jorge Otávio Trierweiler

Co-orientadora:

Profa. Dra. Luciane Ferreira Trierweiler

Porto Alegre

2014

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

A Comissão Examinadora, abaixo assinada, aprova a Dissertação *Estudo da Viabilidade do Uso de Espectroscopia por Fluorescência 2D para Quantificar Teor de Enxofre em Óleo Diesel*, elaborada por Lucas Ranzan, como requisito parcial para obtenção do Grau de Mestre em Engenharia.

Comissão Examinadora:

Dr^a. Daniele Misturini Rossi

Prof. Dr. Marco Flôres Ferrão

Prof. Dr. Pedro Rafael Bolognese Fernandes

Resumo

A dispersão de óxidos sulfurados no meio ambiente, proveniente da queima de combustíveis fósseis, contribui significativamente para a poluição atmosférica. A presença de compostos sulfurados em combustíveis também apresenta desvantagens práticas na operação das unidades, corroendo tubulações e unidades de armazenamento, além de causar danos aos motores. De acordo com a Resolução nº 50, de 23 de dezembro de 2013 da ANP, todo o diesel metropolitano vendido no Brasil deve conter no máximo 10 ppm de enxofre, forçando uma adaptação dos processos produtivos, que, por sua vez, forçam as indústrias a investirem em sistemas de controle e consequentemente metodologias de análise on-line de correntes de processo. As análises atuais certificadoras de enxofre em diesel requerem preparação amostral, equipamentos de alto custo e possuem alto tempo morto associado. Surge a necessidade de estudo de uma metodologia capaz de ser aplicada em analisadores de linha para predição de enxofre em tempo real. Com este intuito, este trabalho visa estudar a viabilidade do uso de espectroscopia por fluorescência 2D para predição de enxofre em óleo diesel. A espectroscopia por fluorescência é uma técnica rápida, que dispensa preparação de amostra e possui alta sensibilidade para compostos naturalmente fluorescentes. Um estudo envolvendo quatro componentes sulfurados presentes em diesel foi realizado por meio de metodologias não supervisionadas - PCA, e supervisionadas - PSCM. Não foi possível segmentar as quatro soluções-padrão por meio de gráfico de escores da PCA, mas com PSCM foi factível o ajuste de modelos multilíneares para predição de enxofre nas soluções-padrão, com coeficientes de determinação superiores a 0,97. Ainda, é possível definir a região de fluorescência mais significativa para cada um dos padrões, constatando que todos possuem regiões de emissão de fluorescência similares. Com relação às amostras de diesel, foram analisados dois grupos distintos de óleo diesel, (i) diesel HDT com média de 100 ppm de enxofre, e (ii) diesel S10 com média de 6,5 ppm de enxofre. Avaliando os resultados da PCA, foi possível segmentar os dois conjuntos de diesel utilizando os dados de escores. Por PSCM, foi possível ajustar modelos baseados em pares de fluorescência capazes de prever satisfatoriamente concentrações de enxofre em amostras de diesel S10. Os modelos ajustados para diesel HDT apresentaram resultados menos significativos. Assim, a viabilidade do uso de espectroscopia por fluorescência 2D para a caracterização de enxofre em correntes de diesel foi confirmada, viabilizando a construção de sensores de processos baseados nesta técnica analítica.

Palavras-chave: Diesel, espectroscopia por fluorescência 2D, enxofre, PCA, PSCM.

Abstract

The dispersion of sulfur oxides in the environment from the burning of fossil fuels contributes significantly to air pollution. The presence of sulfur compounds in fuel also presents disadvantages in operation units, corroding pipes and storage units and may even damage the engine. According to ANP Resolution No. 50, dated December 23, 2013 the entire metropolitan diesel sold in Brazil must contain no more than 10 ppm sulfur, forcing an adaptation of production processes. The current analysis to certify sulfur content in diesel requires sample preparation, expensive equipment and have high dead time associated. Therefore, it is important the study of a method capable of being applied to online analyzers for predicting sulfur in real time. This work aims to study the feasibility of using 2D fluorescence spectroscopy for prediction of sulfur in diesel fuel. Fluorescence spectroscopy is a rapid technique that, usually, does not require sample preparation and has high sensitivity to naturally fluorescent compounds. Initially, a study involving four sulfur containing compounds present in diesel was carried out by means of unsupervised methods - PCA and supervised - PSCM. It was not possible to segregate the four standard solutions by score plot from the PCA, but PSCM was able to fit general multilinear models for the prediction of sulfur in standard solutions with coefficients of determination greater than 0.97. It was possible to define the region of most significant fluorescence for each standard, noting that all standards present similar fluorescence emission regions. After, two groups of different diesel, diesel HDT averaging 100 ppm sulfur and diesel S10 averaged 6.5 ppm sulfur were analyzed. Evaluating the results of PCA was possible to segregate both sets of diesel data using score plot. With PSCM was possible to find models based on fluorescence able to satisfactorily predict concentrations of sulfur in diesel samples S10. The models adjusted for diesel HDT showed less significant results. Thus, the feasibility of prediction of sulfur in diesel using 2D fluorescence spectroscopy was confirmed.

Key-words: sulphur, diesel, 2D fluorescence spectroscopy, PCA, PSCM.

*“Home is behind, the world ahead, And there are many paths to tread.
Through shadows to the edge of night, Until the stars are all alight.
Then world behind and home ahead, We’ll wander back to home and bed.
Mist and twilight, cloud and shade, Away shall fade! Away shall fade!”*

J. R. R. Tolkien

Agradecimentos

Aos meus pais pelo apoio incondicional durante toda a minha vida. Sempre pude contar com vocês como exemplos de boa conduta, dedicação e amor. Um agradecimento especial ao meu pai por sempre me incentivar a estudar, e me lembrar constantemente que o estudo e o esforço podem conquistar qualquer coisa. Um agradecimento especial para minha mãe, sempre vigilante, sempre presente, que se preocupou talvez mais com este trabalho do que eu mesmo durante estes anos.

Um agradecimento especial ao meu irmão Cassiano, sem o qual este trabalho não existiria. Não poderia escolher uma pessoa melhor para ser meu modelo. Espero algum dia poder retribuir todo o carinho e ajuda.

Aos meus demais familiares, que sempre acreditaram no meu potencial.

Ao meu orientador Prof. Dr. Jorge Otávio Trierweiler e a minha co-orientadora Prof. Dra. Luciane Ferreira Trierweiler, pela sabedoria e suporte, fundamentais para o desenvolvimento desta dissertação.

Aos meus queridos amigos da Família GM. Quase duas décadas de amizade não podem ser expressas em poucas palavras. Não sei o que teria feito nas noites de estresse se não fossem as risadas intermináveis, as discussões sobre a vida e os planos mirabolantes para todas as oportunidades que tínhamos de passar juntos. Sou grato a todos por não ter deixado a distância e os caminhos da vida enfraquecerem nossa amizade.

Aos meus amigos da UFRGS, os que ficaram durante os anos do mestrado e aos que se foram por outros caminhos. Um muito obrigado especial ao Felipe, ao Fernando, ao João Paulo, a Carol e os eternos amigos da 08/1 que estiveram diretamente envolvidos com este trabalho. As contribuições de vocês serão para sempre lembradas.

A REFAP e seus colaboradores, por usarem seu valioso tempo para me fornecer amostras e responder às minhas perguntas.

O meu agradecimento a todas as pessoas que durante os últimos anos foram de importância vital, direta ou indiretamente, para a conclusão deste trabalho.

Finalmente, à CNPq, pelo fomento durante toda a pesquisa.

SUMÁRIO

Capítulo 1 – Introdução.....	17
1.1 Motivação.....	17
1.2 Objetivo	19
1.3 Estrutura da Dissertação	19
1.4 Contribuições	20
Capítulo 2 – Revisão bibliográfica.....	21
2.1 Compostos Sulfurados em Diesel.....	21
2.2 Metodologias Padrão para Medição de Enxofre em Diesel.....	25
2.2.1 ASTM D2622 (2010) – Teste padrão de enxofre em produtos de petróleo através de Espectroscopia fluorescente dispersiva com comprimentos de onda de raios X	26
2.2.2 ASTM D4294 (2008) – Teste padrão para enxofre em petróleo e produtos de petróleo por espectroscopia fluorescente dispersiva de Energia de Raios X.....	26
2.2.3 ASTM D5453 (2012) – Método de Teste padrão para determinação de enxofre total em hidrocarbonetos leves, combustíveis para motores com ignição a centelha, combustíveis para motores ciclo Diesel e Óleo para motores utilizando Fluorescência Ultravioleta.....	27
2.2.4 ASTM D7039 (2013) – Metodologia de teste padrão de enxofre em gasolina, diesel combustível, querosene, biodiesel, misturas de biodiesel e misturas gasolina-etanol utilizando comprimentos de onda dispersivos monocromáticos de Raio X.....	27
2.2.5 ASTM D7212 (2013) - Metodologia padrão para teste de amostras com baixo conteúdo enxofre de combustíveis automotivos através de espectroscopia fluorescente de energia dispersiva de raio X usando contador proporcional de baixo ruído de fundo	28
2.2.6 ASTM D7220 (2012) - Metodologia de teste padrão de enxofre em combustíveis automotivos, de aquecimento e de jatos utilizando espectroscopia fluorescente de energia monocromática dispersiva de Raio X.....	28
2.3 Métodos Óticos para Caracterização de Compostos	29
2.3.1 Fundamentos da Espectrometria no Infravermelho Próximo - NIR	30
2.3.2 Espectroscopia por Fluorescência	32
2.4 Quimiometria	36
2.4.1 Uso de dados espectrais não segmentados	37
2.4.2 Seleção de Variáveis - Métodos de Otimização Global e Segmentação de Dados Espectrais.....	38
Capítulo 3 – Metodologia.....	41
3.1 Tratamento dos Dados	41
3.2 Métricas de Qualidade dos Modelos	42
3.3 Desenvolvimento de Modelos e Seleção de Variáveis.....	43
3.3.1 Principal Component Analysis – Análise dos Componentes Principais	43
3.3.2 Pure Spectra Chemometric Modeling - Modelagem Quimiométrica com Componentes Espectrais Puros.....	43
Capítulo 4 – Estudo de Soluções Padrão	49
4.1 Preparo das Soluções-Padrão	50
4.2 Coleta dos Espectros de Fluorescência em 2D.....	52
4.3 Resultados e Discussões.....	53
4.3.1 Metodologia Não assistida - PCA.....	53
4.3.2 Metodologia Assistida - PSCM.....	57
Capítulo 5 – Estudo de amostras de Diesel	63
5.1 Materiais	63
5.1.1 Diesel HDT.....	63

5.1.2 Diesel S10.....	66
5.2 Resultados e Discussões	67
5.2.1 Metodologia Não Assistida - PCA.....	67
5.2.2 Metodologia Assistida - PSCM	73
Capítulo 6 – Conclusões e Sugestões para Trabalhos Futuros.....	79
Referências	83

LISTA DE FIGURAS

Figura 2-1: Comparação entre cromatogramas SCD e FID obtidos de uma injeção simples de petróleo leve. (Chawla e Di Sanzo, 1992)	22
Figura 2-2: Cromatografia tradicional GC-SCD de uma amostra comum de diesel com 7000 ppm de enxofre total. (Wang <i>et al.</i> , 2003)	23
Figura 2-3: Cromatografia GCxGC-SCD da mesma amostra apresentada na Figura 2-2. (Wang <i>et al.</i> , 2003)	24
Figura 2-4: Cromatograma GCxGC-SCD de uma amostra de diesel após hidrotratamento com composição total de enxofre 120 ppm. (Wang <i>et al.</i> , 2003) 25	
Figura 2-5: Estados eletrônicos do spin das moléculas. Em (a) é mostrado o estado eletrônico fundamental Singleto. Em (b) e (c) são mostrados os estados eletrônicos excitados. Quando os spins permanecem emparelhados, a molécula encontra-se em um estado excitado singleto (b). Se os spins tornam-se desemparelhados (c), a molécula está em um estado excitado tripleto. Fonte: (Skoog, Holler e Crouch, 2007).....	33
Figura 2-6: Diagrama parcial de níveis de energia para um sistema fotoluminescente. Fonte: (Sotomayor <i>et al.</i> , 2008).....	34
Figura 2-7: Espectros de Luminescência Total para uma mistura de antraceno e ovaleno (a), e para 8-hidroxibenzopireno (b). Fonte: (Skoog, Holler e Crouch, 2007).....	35
Figura 3-1: Representação esquemática do algoritmo ACO para seleção de grupos de elementos espectrais. Fonte: (Ranzan <i>et al.</i> , 2014).....	46
Figura 3-2: Exemplo de seleção de elementos espectrais utilizando a estratégia de Dorigo e Gambardella (1997). F_i corresponde à trilha de feromônios simulada e C_{Fi} à sua respectiva curva de densidade de feromônios acumulada. No exemplo, o gatilho randômico é acionado em 0,43, levando a seleção do respectivo elemento espectral 670. Fonte: (Ranzan <i>et al.</i> , 2013)	47
Figura 4-1: Estrutura molecular dos padrões avaliados. Fonte: Sigma-Aldrich	50
Figura 4-2: Espectro por fluorescência do Hexano.	51
Figura 4-3: (a) Espectrofluorômetro Fluoromax-4 e (b) cubeta de quartzo com caminho ótico de 1cm.	52
Figura 4-4: Espectros de fluorescência de duas amostras de dibenzotiofeno em hexano. (a) B1 (153 ppm de S) (b) B9 (5 ppm de S).	53
Figura 4-5: Variância Acumulada do grupo de padrões obtida com o aumento do número de componentes principais utilizados para: BDT: dibenzotiofeno; DEDBT: 4,6-dietildibenzotiofeno; DMDBT: 4,6-dimetildibenzotiofeno; TNTH: tionafteno..	54
Figura 4-6: Gráfico de Escores dos padrões avaliados individualmente.	55
Figura 4-7: Gráfico de Escores da PCA com uso de todos os espectros de soluções-padrão.	56
Figura 4-8: Mapeamento de regiões espectrais significativas por análise PSCM baseado nos valores da trilha de feromônios adimensional. (a) Dibenzotiofeno, (b) 4,6 –Dietildibenzotiofeno, (c) 4,6 – Dimetildibenzotiofeno e (d) Tionafteno.	58
Figura 4-9: Pares de fluorescência selecionados pelos modelos localizados nas regiões do espectro de fluorescência 2D.....	60

Figura 4-10: Comparação entre pares espectrais selecionados pelos modelos lineares e a trilha de feromônios para cada solução-padrão.	61
Figura 5-1: Equipamentos utilizados para coleta de espectros por fluorescência do diesel HDT: (a) o equipamento HORIBA Fluoromax®-4 com o módulo para fibra ótica ; (b) a câmara escura e (c) o tipo de frasco utilizado.....	65
Figura 5-2: Espectro de fluorescência típico de uma amostra de Diesel HDT.....	66
Figura 5-3: Espectro de fluorescência da amostra DS10-10.....	67
Figura 5-4: Gráfico de Variância Acumulada para análises individuais de diesel S10 (D_S10) e diesel HDT (D_HDT).	69
Figura 5-5: Gráfico de Escores do conjunto amostras Diesel HDT e Diesel S10, avaliados de forma individual.	69
Figura 5-6: Comparativo entre exemplo de espectro por fluorescência de diesel S10 (a) e Amostra DS10-11 (b).	70
Figura 5-7: Gráfico de Escores do conjunto de amostras Diesel HDT e Diesel S10, avaliados de forma conjunta.....	71
Figura 5-8: Gráfico de Escores do conjunto de amostras de Diesel HDT, Diesel S10 e novas medições de Diesel HDT com cubeta.....	72
Figura 5-9: Mapeamento de regiões espectrais significativas por análise PSCM baseado nos valores da trilha de feromônios adimensional. (a) Diesel S10 e (b) Diesel HDT.....	73
Figura 5-10: Pares de fluorescência selecionados pelos modelos nas respectivas posições no espectro por fluorescência para diesel HDT e diesel S10.	75
Figura 5-11: Comparação entre os pares selecionados pelos modelos lineares para diesel HDT e diesel S10 e suas respectivas trilhas de feromônios.	76
Figura 5-12: Pares de fluorescência selecionados pelos modelos lineares de um a quatro pares para busca de enxofre nos padrões da família DBT e diesel S10, plotados sobre as regiões mais significativas para diesel S10.....	76

LISTA DE TABELAS

Tabela 2-1: Distribuição de componentes sulfurados em diferentes óleos diesel (concentração em ppm) . Fonte: (Hua <i>et al.</i> , 2003)	25
Tabela 2-2: Resumo dos testes ASTM apresentados e suas aplicações.....	29
Tabela 4-1: Concentração de enxofre em cada uma das amostras produzidas e a segmentação do conjunto de amostras grupo de teste e calibração (amostras em negrito e sublinhado).....	51
Tabela 4-2: Número mínimo de PCs necessários para expressar mais de 95% da variância contida no conjunto original de dados.....	54
Tabela 4-3: Seleção de pares por PSCM e métricas dos modelos para DBT.	59
Tabela 4-4: Seleção de pares por PSCM e métricas dos modelos para DEDBT.....	59
Tabela 4-5: Seleção de pares por PSCM e métricas dos modelos para DMDBT.	59
Tabela 4-6: Seleção de pares por PSCM e métricas dos modelos para TNTH.....	60
Tabela 5-1: Concentração de enxofre nas amostras de Diesel provenientes do processo de hidrotreamento (HDT). Amostras em fundo azul fazem parte do grupo de testes.....	64
Tabela 5-2: Amostras de diesel S10. Amostras com fundo azul foram selecionadas para grupo de teste.	66
Tabela 5-3: Número de PCs necessários para descrever 95% da variância acumulada dos dados originais.....	68
Tabela 5-4: Pares de fluorescência selecionados por PSCM e valores das métricas de avaliação dos modelos para Diesel HDT.	74
Tabela 5-5: Pares de fluorescência selecionados por PSCM e valores das métricas de avaliação dos modelos para Diesel S10.	74

ABREVIações

ACO	Otimização Colônia de Formigas (<i>Ant Colony Optimization</i>)
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ASTM	Associação Americana para Testes e Materiais (<i>American Society for Testing and Materials</i>)
CG	Cromatografia Gasosa
CNT	Confederação Nacional do Transporte
DBT	Dibenzotiofeno
DEDBT	4,6 - DiEtilDibenzotiofeno
DMDBT	4,6 - DiMetilDibenzotiofeno
Ex/Em	Excitação/Emissão
FID	Detecção por Ionização de Chama (<i>Flame Ionization Detection</i>)
GA	Algoritmo Genético (<i>Genetic Algorithm</i>)
HDT	Hidrotratamento
IR	Infra vermelho (<i>infrared</i>)
NIR	Infra vermelho próximo (<i>Near Infrared</i>)
PC	Componente Principal (<i>Principal Component</i>)
PCA	Análise dos Componentes Principais (<i>Principal Component Analysis</i>)
PLS	Mínimos Quadrados Parciais (<i>Partial Least Squares</i>)
ppm	Parte por milhão
Proconve	Programa de Controle da Poluição do Ar por Veículos Automotores
PSCM	Modelagem Quimiométrica com os Componentes Espectrais Puros (<i>Pure Spectra Chemometric Modeling</i>)

R^2	Coeficiente de Determinação
REFAP	Refinaria Alberto Pasqualine
RMSEC	Raiz quadrada do erro médio quadrático de calibração (<i>Root-Mean Square Error of Calibration</i>)
RMSEP	Raiz quadrada do erro médio quadrático de predição (<i>Root-Mean Square Error of Prediction</i>)
SCD	Detecção de Enxofre por Quimiluminescência (<i>Sulfur Chemiluminescence Detector</i>)
SNV	Padrão Normal de Variação (<i>Standard Normal Variate</i>)
SSR	Soma dos Resíduos Quadráticos (<i>Sum of Square Residuals</i>)
TNTH	Tionafteno
ULSD	Ultra Low Sulfur Diesel
λ	Comprimento de onda

Capítulo 1 – Introdução

1.1 Motivação

A dispersão de óxidos sulfurados no meio ambiente, proveniente da queima de combustíveis fósseis, contribui significativamente para poluição do ar, sendo há muito tempo um tópico importante nas discussões sobre defesa ambiental (Chandra Srivastava, 2012). Além do fator ambiental, componentes sulfurados ainda são considerados venenos para os catalisadores utilizados durante o processo de refino. A presença de compostos sulfurados também apresentam desvantagens práticas na operação das unidades, pois corroem tubulações e unidades de armazenamento e ainda podem causar danos aos motores e rotores (Murakami, 1995).

Durante o processo de combustão, o enxofre presente no combustível se oxida a SO_2 e SO_3 , este último, em contato com a água, forma ácido sulfúrico que corrói as partes metálicas dos motores, diminuindo sua vida útil. Quanto maior a concentração deste elemento, maior a emissão de particulados e óxidos (poluentes primários) que trazem prejuízos à saúde humana. O dióxido de enxofre, em contato com a umidade atmosférica, gera ácido sulfúrico e contribui consideravelmente para a chuva ácida, acidificando o solo e provocando o arraste de metais pesados do solo para rios e lagos.

Fora o dano ambiental, o aumento na dispersão de poluentes na atmosfera leva a um grande gasto do governo com o sistema de saúde pública, sobretudo em grandes centros urbanos. Estima-se que um aumento de $10 \mu\text{g}/\text{m}^3$ na concentração de material particulado no ar leva a um aumento em 1,5% nas internações por doença isquêmica do coração em idosos e mais de 4% por doenças pulmonares em crianças (Cnt, 2012).

Em países desenvolvidos, o teor de enxofre máximo encontrado em combustíveis já é controlado severamente há mais de uma década. Em países como Japão, EUA e na União Europeia (em 2005) todo o diesel comercializado já possuía uma concentração máxima de 10, 15 e 50 ppm respectivamente.

No Brasil, até o ano de 1980 (quando não havia regulamentação), os veículos a diesel utilizavam óleo com até 13.000 ppm de enxofre. Atualmente, o país se encontra na chamada fase P7 do Programa de Controle da Poluição do Ar por Veículos Automotores (Proconve). De acordo com a legislação atual, regulamentada pela Resolução nº 50, de 23 de dezembro de 2013 da Agência Nacional do Petróleo (ANP), todo o diesel para uso de frotas cativas de ônibus urbanos nos municípios e regiões metropolitanas deve conter até no máximo 10 mg/Kg de enxofre (denominado Diesel S10). Para uso rodoviário, agrícola, marítimo e industrial, fica permitido a venda de diesel com até 500 ppm de enxofre (Diesel S500).

A evolução nas leis que regulamentam a comercialização do óleo diesel no território nacional afeta diretamente as refinarias e o sistema produtivo do mesmo. Para atender aos novos limites máximos de enxofre no combustível, o processo de hidrotreamento para remoção deste contaminante precisou ser intensificado e ganhou um maior destaque na busca por uma maior eficiência econômica.

Atualmente, todos os testes certificadores para qualificar o óleo diesel são definidos pela ANP (2013). Dentre os indicados para certificar enxofre em diesel S10, pode-se citar ASTM D2622 (baseado em espectroscopia fluorescente dispersiva de raio X) e ASTM D5453 (procedimento baseado em quimiluminescência). Estas normas requerem preparação de amostra por pessoal qualificado, equipamentos especializados de alto custo, alto custo de operação e manutenção, são invasivos e destrutivos e possuem um tempo morto significativo para obtenção de resultados. Usualmente, apenas uma amostra de cada ponto de interesse (seja a saída final do processo de hidrotreamento ou o tanque final de armazenagem de diesel) é retirada por dia e analisada, uma vez que o mesmo equipamento utilizado para quantificar enxofre em diesel é utilizado para certificar também outras correntes de combustíveis dentro da refinaria.

O tempo morto associado a este processo de análise inviabiliza uma ação de controle efetiva sobre o processo produtivo, tanto da unidade de hidrotreamento, quando da mistura de diferentes correntes que formam o óleo diesel final (denominado *blending*), uma vez que o tempo decorrido entre a amostragem e a obtenção de resultado da análise é alto.

O desenvolvimento de um analisador de processo, instalado diretamente nas linhas produtivas e capaz de quantificar em tempo real a concentração de enxofre da corrente em questão, poderia trazer benefícios expressivos tanto financeiros quanto ambientais, subsidiando sistemas de controle avançado para a produção de combustível com maior lucratividade e menores índices de reprocessamento. Esta seria uma ferramenta poderosa para reduzir significativamente a dependência de ensaios laboratoriais e fornecer informações valiosas, em tempo real, que não estão disponíveis atualmente por meio de outras técnicas.

Dentre as técnicas analíticas com maior possibilidade para uso em analisadores de processo, se destacam os métodos óticos como espectroscopia no infravermelho próximo (NIR) e o uso de espectroscopia de fluorescência. Tais técnicas apresentam características

interessantes, tais como: rapidez, baixa ou nenhuma necessidade de preparação de amostras, não geração de resíduos, instrumentação robusta e aplicável a sistemas explosivos, não serem invasivas ou destrutivas, além do fato de que em uma mesma análise, com a manipulação correta dos espectros, mais de uma propriedade de interesse pode ser quantificada.

São diversas as pesquisas realizadas no ramo petroquímico que utilizam espectroscopia NIR para avaliar parâmetros de interesse em combustíveis. Como apresentado em *Breitkreitz* (2003), a técnica se mostrou capaz de quantificar satisfatoriamente concentrações de enxofre total em diesel em uma faixa de 70 - 3300 ppm, com resultados comparáveis aos testes certificadores tradicionais. Porém, em sistemas menos concentrados (< 15 ppm de enxofre), a técnica se mostra pouco efetiva e não apresenta bons resultados para predição e quantificação do mesmo, tornando a técnica pouco atrativa para avaliação de correntes de baixo enxofre (como diesel S10) (Bueno, 2011).

As técnicas analíticas baseadas em espectroscopia por fluorescência, por sua vez, possuem grande sensibilidade e limites de detecção de até três ordens melhores que os encontrados em espectroscopia de absorção (Skoog, Holler e Crouch, 2007). Esta metodologia já é empregada em uma grande gama de processos industriais e é dominante nas áreas da biotecnologia. Na indústria dos combustíveis, muitos são os trabalhos que avaliam o uso desta técnica, uma vez que grande parte dos compostos de interesse emitem naturalmente fluorescência, dispensando a preparação das amostras. Dentre os compostos naturalmente fluorescentes, encontram-se grande parte dos compostos sulfurados presentes em diesel, tornando esta metodologia altamente atrativa para o desenvolvimento de analisadores de processos para quantificação de enxofre (Aburto *et al.*, 2014).

1.2 Objetivo

O objetivo deste trabalho é estudar a viabilidade do uso de espectroscopia fluorescente 2D como uma alternativa para quantificar teor de enxofre em óleo diesel.

Como objetivos secundários, estão incluídos o estudo de compostos sulfurados presentes em óleo diesel e a busca por enxofre nos mesmos com uso de espectroscopia por fluorescência e a capacidade em usar esta metodologia para segmentar óleos diesel que passaram por processos de hidrotreatamento de diferentes intensidades.

O estudo dos compostos sulfurados presentes em óleo diesel também visa avaliar se as regiões do espectro por fluorescência mais significativas para predição de enxofre em cada um desses compostos tem correlação com as regiões mais significativas para quantificação de enxofre em diesel.

1.3 Estrutura da Dissertação

No Capítulo 2 é apresentada a revisão bibliográfica pertinente aos assuntos abordados de forma geral no decorrer do trabalho. Nesse capítulo é feita uma revisão a respeito de Óleo Diesel e compostos sulfurados presentes no mesmo, além dos métodos de caracterização usuais utilizados na indústria do petróleo para certificação do óleo, definidos pelas normas Resolução ANP 50 e as referidas normas ASTM. Além disso, este capítulo apresenta um apanhado geral a respeito dos métodos ópticos com aplicabilidade

na caracterização de diesel, além das principais técnicas quimiométricas utilizadas neste trabalho.

No Capítulo 3 é apresentada a metodologia empregada neste trabalho, bem como as métricas utilizadas para preparação de amostras, ajuste e teste de modelos matemáticos, avaliação de modelos e procedimento para realização das análises quimiométricas em dados de espectroscopia por fluorescência 2D.

O Capítulo 4 esta focado no estudo de soluções padrão contendo moléculas que possuem enxofre, consideradas responsáveis pela presença de enxofre em correntes de diesel oriundas do processo de hidrotreatamento.

Capítulo 5 é composto por análises quimiométricas de dados de espectroscopia por fluorescência de amostras de Diesel S10 (ULSD - *Ultra Low Sulphur Diesel*) e Diesel HDT. Nesse capítulo são apresentadas métricas qualitativas para classificação de amostras, além de métricas quantitativas para determinação do teor de enxofre presente em amostras de diesel. Ambas as métricas apresentadas são baseadas em dados de espectroscopia por fluorescência.

O Capítulo 6 apresenta as principais conclusões obtidas no desenvolver deste trabalho, além de apresentar sugestões para trabalhos futuros. Por fim são apresentadas as referencias bibliográficas.

1.4 Contribuições

- Desenvolvimento de Inferências utilizando o Método de Otimização Colônia de Formigas publicado nos anais do Congresso Brasileiro de Eng. Química 2014 - COBEQ 2014.
- Ultra-low Sulfur Diesel Classification with 2D Fluorescence Spectroscopy and PCA submetido para - Energy & Fuels.
- Wheat flour characterization using NIR and spectral filter based on Ant Colony Optimization. Chemometrics and Intelligent Laboratory Systems, v. 132, n. 0, p. 133-140, 3/15/ 2014. ISSN 0169-7439

Capítulo 2 – Revisão bibliográfica

2.1 Compostos Sulfurados em Diesel

Com base em todos os fatores industriais e ambientais referentes a este tema, passa a ser de extrema importância a quantificação da concentração de enxofre presente nos combustíveis comercializados, no caso deste trabalho, em especial à adequação das correntes de diesel à legislação ambiental vigente e consequente viabilidade de comércio de óleo diesel.

Existem diversas maneiras de abordar a caracterização de enxofre em diesel. Muitas delas objetivam exclusivamente medir a quantidade total de enxofre presente, como por fluorescência de raio-x (classe de testes padrões realizados para certificar diesel combustível), uma vez que o uso de espectroscopia fotoeletrônica de raio-x é capaz de distinguir diferentes estados de oxidação do enxofre (West *et al.*, 2013).

Durante as últimas décadas, diferentes técnicas de cromatografia gasosa (CG) e líquida (CL), combinadas com técnicas de espectroscopia de massa ou sistemas especializados para detecção de enxofre, em especial a técnica de Detecção de Enxofre por Quimiluminescência (*Sulfur Chemiluminescence Detector - SCD*), serviram como base para pesquisa de métricas de quantificação e separação de compostos sulfurados. No entanto, cada uma destas técnicas possui limitações que variam desde requerimentos de amostragem específicos até limites de detecção insatisfatórios (Chawla e Di Sanzo, 1992; Bernard *et al.*, 1994).

A combinação entre CG com SCD possibilitou um grande avanço na detecção de compostos sulfurados. O SCD possui grande seletividade e sensibilidade para enxofre, tornando possível a quantificação de compostos sulfurados que coeluem com hidrocarbonetos durante a separação cromatográfica. Assim, a seletividade do SCD foi capaz de superar problemas de coeluição e passou a ser aplicado em situações onde

existe grande interferência de hidrocarbonetos que limitavam a detecção com CG-MS (Chawla e Di Sanzo, 1992).

O avanço em especificação alcançado com o uso de CG-SCD foi notável, a Figura 2-1 apresenta a comparação entre técnicas convencionais de Detecção por Ionização de Chama (Flame Ionization Detection - FID) e SCD.

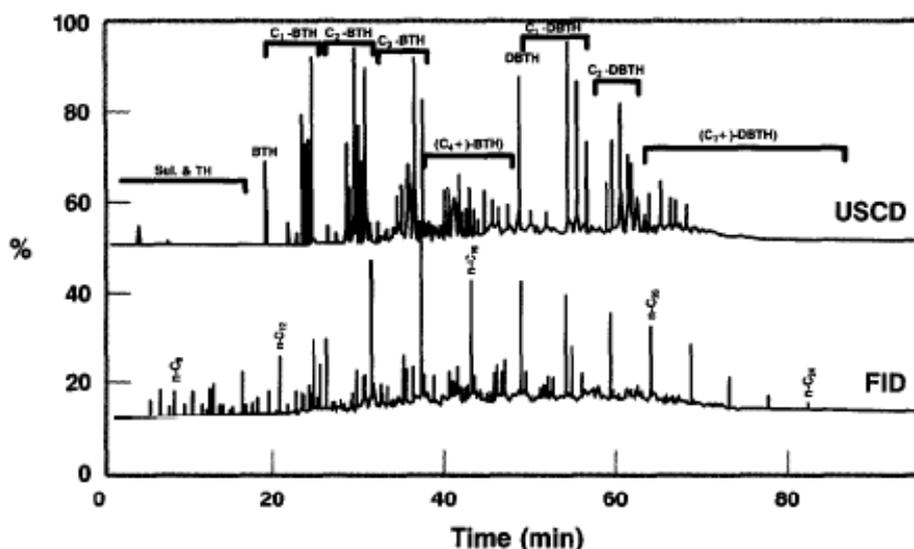


Figura 2-1: Comparação entre cromatogramas SCD e FID obtidos de uma injeção simples de petróleo leve. (Chawla e Di Sanzo, 1992)

O uso das técnicas de CG-SCD otimizou a busca por componentes sulfurados minimizando problemas de coeluição com hidrocarbonetos, porém, a coeluição de diferentes componentes sulfurados torna inviável a separação e quantificação dos mesmos, uma vez que a resolução dos picos e a sensibilidade da técnica não apresentam resultados suficientes.

Com a evolução tecnológica das últimas duas décadas, técnicas de CG bidimensional (CGxCG) se tornaram possíveis. A literatura apresenta resultados promissores do uso desta tecnologia acoplada a detectores FID (Beens *et al.* 1998), deste modo, a extensão da técnica para uso de detectores específicos para enxofre (CGxCG-SCD) foi capaz de elucidar boa parte das perguntas relativas ao complexo problema de compostos sulfurados encontrados em diesel.

Componentes sulfurados podem ser segmentados em classes, de acordo com sua estrutura, como benzotiofenos ou dibenzotiofenos. Os análogos alquil em cada classe podem ser agrupados de acordo com seu número de carbonos. A relativa concentração de cada classe, assim como seu número de carbonos, pode ser quantificada por integração dos picos correspondentes aos mesmos nos cromatogramas. Em aplicações práticas, a técnica pode ser usada para monitorar desde composição de uma

determinada amostra, até medir a eficiência de um processo específico de dessulfurização (Wang *et al.*, 2003).

As Figura 2-2 e Figura 2-3 ilustram resultados das técnicas CG-SCD e CGxCG-SCD aplicadas a uma amostra comum de diesel com aproximadamente 7000 ppm de enxofre, realizadas nas mesmas condições experimentais. Os picos foram identificados comparando tempos de retenção com padrões ou com o isolamento cuidadoso dos compostos, juntamente com outras técnicas como CL e identificados com MS.

Na análise da Figura 2-2, entre os compostos encontrados, duas grandes classes se fazem presentes: benzotiofenos (BT) e dibenzotiofenos (DBT). Porém, a grande coeluição de compostos sulfurados forma um grande "envelope" entre o tempo de retenção de 30-80 min. Este envelope impossibilita a análise quantitativa dos picos e inviabiliza a identificação dos componentes.

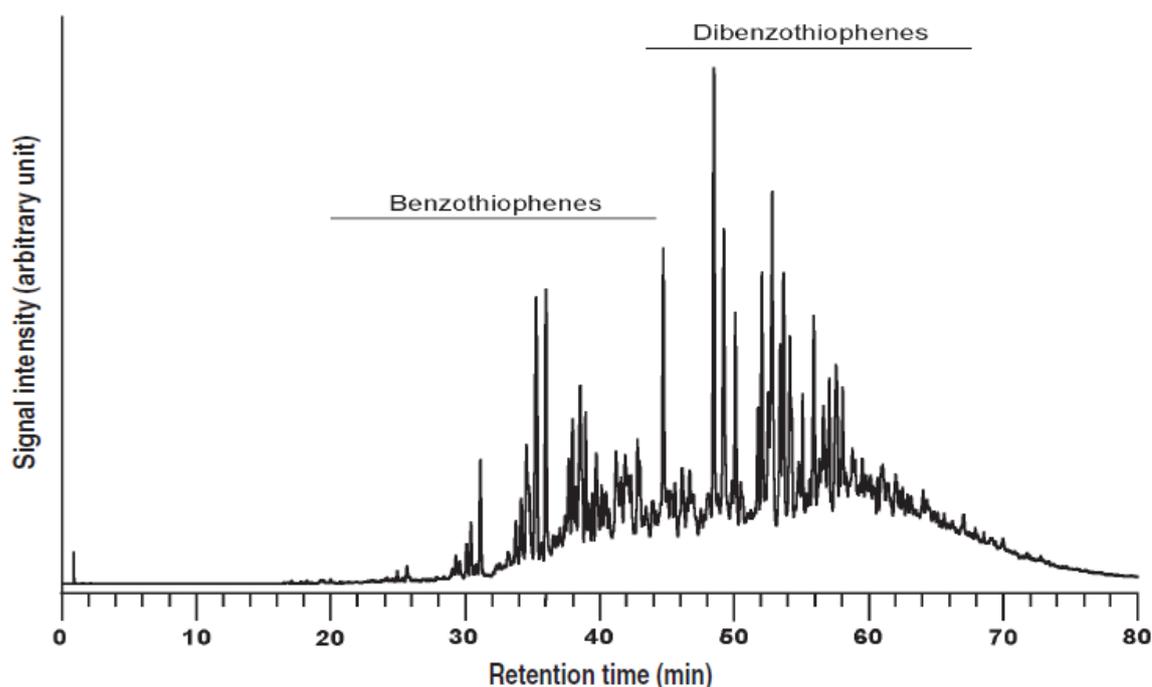


Figura 2-2: Cromatografia tradicional GC-SCD de uma amostra comum de diesel com 7000 ppm de enxofre total. (Wang *et al.*, 2003)

A Figura 2-3 por sua vez apresenta duas grandes bandas, a superior (associada aos dibenzotiofenos) e a inferior (associada aos benzotiofenos). Dentro de cada banda, ainda é possível a subdivisão em função do número de carbonos presente nas cadeias. Por causa da grande resolução, é possível realizar a integração dos picos de cada grupo e sua quantidade pode ser avaliada.

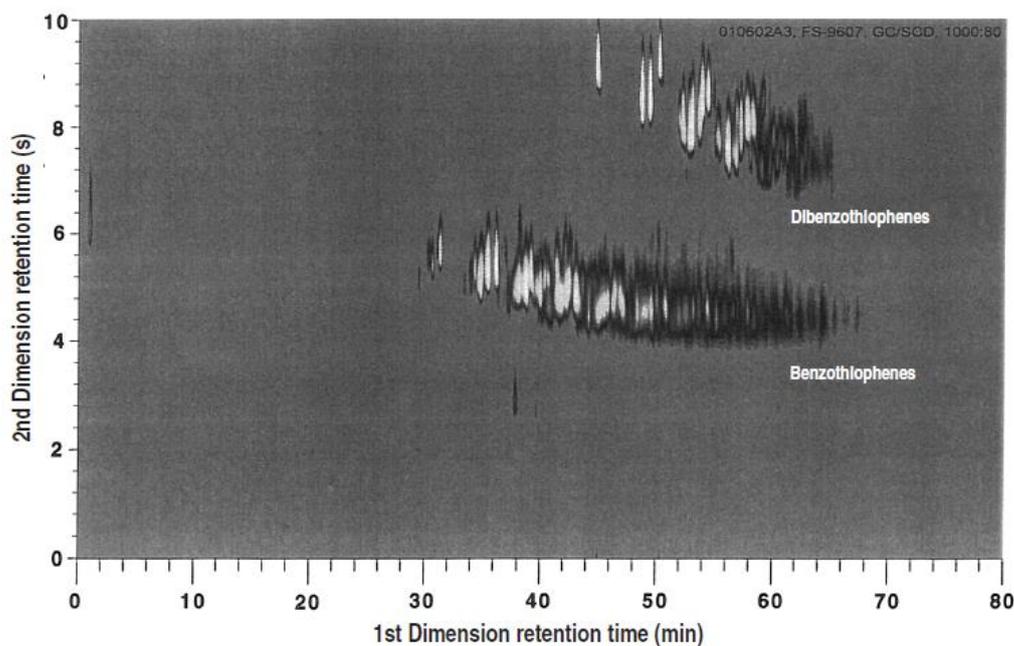


Figura 2-3: Cromatografia GCxGC-SCD da mesma amostra apresentada na Figura 2-2. (Wang *et al.*, 2003)

A Figura 2-4 traz o cromatograma de uma amostra de diesel após processos de hidrotratamento para retirada de enxofre. Fica claro pela avaliação dos picos que os compostos presentes são em sua maioria dibenzotiofenos. O processo de desulfurização remove quase que completamente todos os benzotiofenos e a maioria dos dibenzotiofenos (DBT), e os compostos remanescentes podem então ser identificados por comparação entre tempo de retenção com padrões, e são em sua maioria DBT com grupos alquila entre 2 e 4 carbonos (Wang *et al.*, 2003).

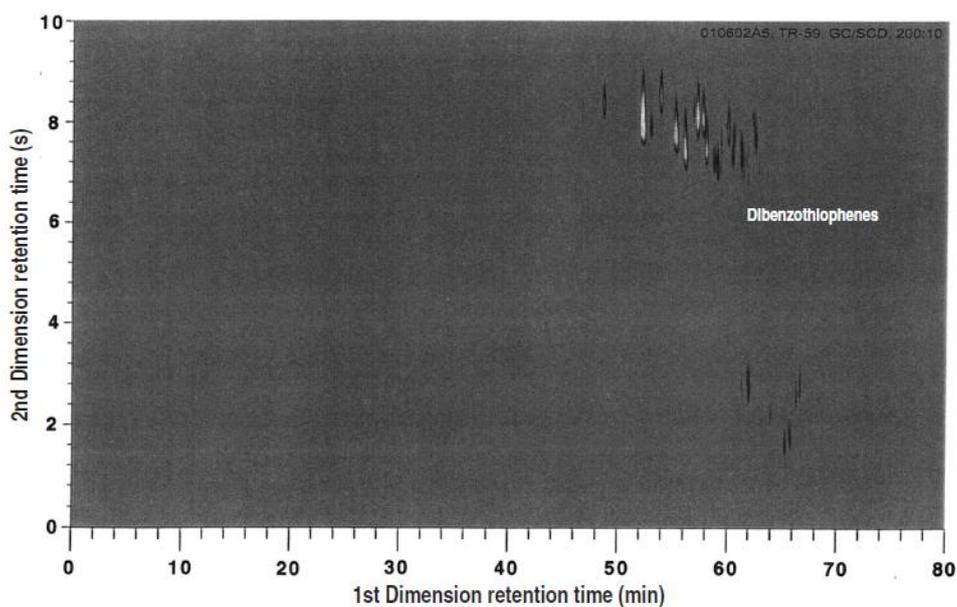


Figura 2-4: Cromatograma GCxGC-SCD de uma amostra de diesel após hidrotratamento com composição total de enxofre 120 ppm. (Wang *et al.*, 2003)

A distribuição de componentes sulfurados em diesel está atrelada diretamente à origem do petróleo a ser tratado. Trabalhos como o de Hua *et al.*, (2003) trazem um comparativo dos componentes sulfurados presentes em diferentes óleos diesel provenientes de diversas fontes e processos de refino, avaliados com o uso de CGxCG-SCD. A Tabela 2-1 apresenta um apanhado geral dos componentes sulfurados provenientes de diferentes amostras de óleo e que sofreram diferentes processos de refino. Na Tabela 2-1, a nomenclatura utilizada faz referência a: FCC (*Fluid Catalytic Cracking*), RFCC (*Residuum Fluid Catalytic Cracking*), VB (*Visbreaking*) e DC (*Delayed Coking*). HFU: Óleo Diesel Hidrorefinado. Alkyl S: tiols, sulfidos e disulfidos. TPs: tiofenos. BT: benzotiofenos. DBTs: dibenzotiofenos. BNTs: Benzonaftotiofenos.

Tabela 2-1: Distribuição de componentes sulfurados em diferentes óleos diesel (concentração em ppm) . Fonte: (Hua *et al.*, 2003)

Sulfur compound group	Stright run 3A1	Stright run 3B1	Stright run 3A2	FCC	RFCC	VB	DC	HFU
Crude oil origin	Cabinda	Masila	Banama	Masila Zafiro	Oman Kole	Duri	Duri	Diesel feed
DBTs	18.8	37.3	25.4	214.9	261.8	8.8	16.8	4.7
4-MeDBTs	60.6	71.9	78.3	199.6	291.7	19.0	24.9	43.3
2-MeDBTs	48.1	50.8	58.8	354.3	362.9	14.7	23.4	6.8
3-MeDBTs	1.5	1.2	0.9	56.3	97.9	9.3	21.2	0.2
1-MeDBTs	38.1	38.9	53.9	97.1	114.8	9.3	13.7	9.4
4-EtDBTs	21.1	15.4	28.6	49.7	51.2	10.7	9.7	8.9
4,6-DMeDBTs	52.5	38.0	83.0	116.5	133.2	23.8	17.4	35.1
Alkyl S + TPs	420	550	1000	220	210	2900	1000	28
Group BTs	1500	1900	2100	3500	6700	1900	1600	71
Group DBTs	2000	2900	1300	4400	5000	330	700	540
Group BNTs	44	20	130	140	7	–	61	4
Total S value	3900	5400	4500	8300	12000	5100	3400	650
Sulfur (XRF)	4200	5400	4600	8400	11300	5400	3700	600
R.S.D. (%)	4.4	0.3	1.8	0.8	3.8	3.8	6.0	5.3
DBTs/total S (%)	51.4	54.0	28.5	52.8	41.8	6.5	20.5	84.2

Fica claro pela avaliação da Tabela 2-1 que o óleo diesel hidrotratado possui em sua maioria compostos do grupo dos DBT, praticamente 85% do enxofre total. Estes compostos são os de maior dificuldade de remoção, e em alguns casos catalisadores e parâmetros operacionais devem ser ajustados para levarem em consideração a concentração destes compostos, e não somente do enxofre total (Hua *et al.*, 2003).

2.2 Metodologias Padrão para Medição de Enxofre em Diesel

A aferição dos limites de enxofre presente no diesel comercial é uma certificação exigida em diversos países. Assim, foram desenvolvidos testes considerados padrões, que são aceitos mundialmente como confiáveis e precisos, para esta quantificação. Uma das instituições mais respeitadas na certificação destes testes é a Associação Americana para Testes e Materiais (*American Society for Testing and Materials - ASTM*), a qual rege grande parte dos testes exigidos pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) para produção e distribuição de óleo diesel no Brasil. Dentre as

diversas normas técnicas publicadas pela ASTM acerca deste assunto, seis delas são classificadas como sendo mais importantes para a aferição da concentração de enxofre em amostras de diesel, sendo elas: ASTM D2622, ASTM D4294, ASTM D5453, ASTM D7039, ASTM D7212 e ASTM D7220.

As normas citadas anteriormente apresentam os testes certificadores de quantidade de enxofre em óleo diesel exigidos pela regulamentação da ANP estabelecida na portaria ANP 50/2013 para comercialização do mesmo no território brasileiro. A seguir as principais normas são apresentadas resumidamente.

2.2.1 ASTM D2622 (2010) – Teste padrão de enxofre em produtos de petróleo através de Espectroscopia fluorescente dispersiva com comprimentos de onda de raios X

Esta norma aborda a determinação de enxofre total em petróleo ou derivados de petróleo que são líquidos em condições ambiente, líquidos em temperaturas moderadas, ou solúveis em hidrocarbonetos. A técnica pode ser aplicada para amostras com enxofre total em faixa desde 3 mg/Kg até 5.3% da massa total. Amostras com concentração mássica maior que 5% em enxofre podem ser diluídas para se enquadrar ao método.

A amostra é posicionada sobre a incidência de feixe de raios X, e a intensidade do pico da linha $K\alpha$ do enxofre em 0,5373 nm é coletada. A intensidade de fundo, analisada no comprimento de onda recomendado de 0,519 nm é subtraída da intensidade do pico. O resultado é então comparado com uma curva de calibração previamente preparada ou uma equação adequada para obter a concentração de enxofre em porcentagem de massa.

Esta técnica, por basear sua análise em medidas óticas, propicia uma rápida e confiável medida de enxofre total com um mínimo de preparo de amostra. Por outro lado, o equipamento necessário para caracterização das amostras tem alto custo e não possui fatores de correção para interferências, como no caso de amostras com alto grau de oxigênio (que absorve a radiação $K\alpha$ na mesma faixa que o enxofre), ou outros elementos dopantes, tais como fósforo, zinco, bário, dentre outros.

2.2.2 ASTM D4294 (2008) – Teste padrão para enxofre em petróleo e produtos de petróleo por espectroscopia fluorescente dispersiva de Energia de Raios X

Essa metodologia baseia-se em princípios de fluorescência de raios-X, capazes de determinar enxofre total em produtos de petróleo líquidos a temperatura ambiente, líquidos a temperaturas moderadas ou solúveis em hidrocarbonetos. A metodologia descrita nesta norma pode ser aplicada a amostras com enxofre total variando desde 150 mg/Kg até 4.6% em massa total.

A amostra é irradiada por um feixe de energia emitido de um tubo de raios-X. A excitação X característica resultante é medida e comparada com padrões previamente

calculados. Um mínimo de três grupos de calibração com distintas concentrações de enxofre é requerido para a criação da curva de concentração: 0.0 - 0.1 % mássicos, 0.1 - 1.0 % mássicos e 1.0 - 5.0 % mássicos de enxofre.

Interferências espectrais podem ser causadas por elementos com picos característicos de raio-X próximos ao do enxofre, e os limites destes interferentes nas amostras devem ser respeitados para resultados confiáveis.

As condições do filme utilizado como janela devem ser rigorosamente avaliadas, o mesmo deve possuir baixa rugosidade e ser mantido perfeitamente limpo.

2.2.3 ASTM D5453 (2012) – Método de Teste padrão para determinação de enxofre total em hidrocarbonetos leves, combustíveis para motores com ignição a centelha, combustíveis para motores ciclo Diesel e Óleo para motores utilizando Fluorescência Ultravioleta

Essa norma descreve o procedimento para determinação da quantidade de enxofre presente em amostras de hidrocarbonetos leves com base em medidas de fluorescência ultravioleta. Este teste pode ser aplicado em amostras que tenham ponto de ebulição entre 25 - 400 graus Celsius, com viscosidade entre 0.2 e 20 cSt em temperatura ambiente. Amostras com concentração de enxofre total entre 1.0 e 8000 mg/Kg podem ser analisadas com este procedimento.

A amostra de hidrocarboneto é inserida em uma câmara de combustão de alta temperatura, onde todo o enxofre é oxidado a dióxido de enxofre (SO₂), em uma atmosfera rica em oxigênio. Água produzida durante a combustão é removida e o restante dos gases provenientes da combustão são expostos à luz ultravioleta. O SO₂ absorve energia da luz UV e passa ao seu estado excitado (SO₂*). A fluorescência emitida quando SO₂* retorna ao seu estado estável é detectada por um fotomultiplicador e o sinal resultante é comparado com padrões para determinação da quantidade de enxofre total da amostra.

2.2.4 ASTM D7039 (2013) – Metodologia de teste padrão de enxofre em gasolina, diesel combustível, querosene, biodiesel, misturas de biodiesel e misturas gasolina-etanol utilizando comprimentos de onda dispersivos monocromáticos de Raio X

Metodologia baseada em medidas de fluorescência de raios-X com o uso de monocromador. A técnica descrita pode ser usada em uma faixa desde 3.2 até 2822 mg/Kg de concentração de enxofre total nas amostras. A amostra é irradiada por um feixe de raios-X com comprimento de onda apropriado para excitar a camada K do enxofre. A radiação K de comprimento de onda de 0,5373 nm emitida pelo enxofre é captada por um monocromador, e sua intensidade pode ser convertida em concentração total de enxofre através de uma equação de calibração. O uso do monocromador reduz o ruído de fundo, simplificando a matriz de correlação e aumentando a razão sinal/ruído de fundo, quando comparada com a excitação policromática de técnicas de comprimento de onda dispersiva de raios-X convencionais.

2.2.5 ASTM D7212 (2013) - Metodologia padrão para teste de amostras com baixo conteúdo enxofre de combustíveis automotivos através de espectroscopia fluorescente de energia dispersiva de raio X usando contador proporcional de baixo ruído de fundo

Metodologia baseada em medidas de energia dispersiva de fluorescência de raios-X, em amostras com concentração total de enxofre variando entre 7 - 50 mg/Kg.

A amostra é posicionada sob um feixe oriundo de uma fonte de raios-X com alvo de titânio e filtro simples, de forma que a radiação emitida seja essencialmente monocromática de 4.51 keV e virtualmente sem ruído de fundo em 2.3 keV. Um contador proporcional de baixo ruído de fundo mede então a intensidade da fluorescência da camada K do enxofre e a intensidade da camada K do argônio (ar residual) e a intensidade acumulada é comparada com a intensidade previamente medida de padrões de calibração, permitindo obter a concentração do enxofre em mg/Kg. Caso traços de cloro sejam esperados na amostra, outras regiões do espectro devem ser avaliadas para compensar a banda de sobreposição espectral deste elemento. O método requer padrões de calibração em uma faixa de concentração entre 0 e 150 mg/Kg de enxofre.

2.2.6 ASTM D7220 (2012) - Metodologia de teste padrão de enxofre em combustíveis automotivos, de aquecimento e de jatos utilizando espectroscopia fluorescente de energia monocromática dispersiva de Raio X

Metodologia de medição de enxofre baseada em medidas de energia dispersiva de fluorescência de raios-X. Proposto para amostras com concentração total de enxofre entre 3 e 942 mg/Kg. Por definição, uma radiação monocromática de raios-X é aquela que possui energia pré-determinada de fótons, com largura de banda entre $\pm 5\%$ da energia selecionada.

A amostra é posicionada em um feixe monocromático de raios-X, e a área do pico da camada K do enxofre em 2.307 keV é medida. O espectro de fundo, medido com um padrão de óleo sem enxofre (ou outra matriz semelhante para o branco de amostra) é adaptado ao espectro medido usando regiões de ajuste seguindo as orientações do fabricante do equipamento, sendo subtraído do espectro coletado da amostra. A intensidade resultante é então comparada com curvas de calibração obtidas a partir de padrões e a quantidade total de enxofre em mg/Kg é calculada.

A Tabela 2-2 apresenta um resumo do tipo de teste realizado, faixa de concentrações e em para que categoria de diesel é utilizado cada teste ASTM.

Tabela 2-2: Resumo dos testes ASTM apresentados e suas aplicações.

Metodo	Tipo de Teste	Faixa	Aplicado
D 2622	Espectroscopia fluorescente de comprimento de onda dispersivo de raios X	3 – 53 000 ppm	D S10 D S500
D 4294	Espectroscopia fluorescente de energia dispersiva de raios X	150 – 50 000 ppm	D S500
D 5453	Combustão da amostra; Fluorescência ultravioleta para quantificar dióxido de Enxofre	1.0 – 8000 ppm	D S10 D S500
D 7039	Espectroscopia fluorescente de comprimentos de onda monocromáticos dispersivos de raios X	3 – 2800 ppm	D S10 D S500
D 7212	Espectroscopia fluorescente de energia dispersiva de raios X usando contador proporcional de baixo ruído de fundo	7 – 50 ppm	D S10
D 7220	Espectroscopia fluorescente de energia monocromática dispersiva de raios X	3 – 940 ppm	D S500 D S10

2.3 Métodos Óticos para Caracterização de Compostos

Métodos óticos envolvem medidas baseadas em luz e outras formas de radiação eletromagnética. As interações da radiação com a matéria são o objetivo principal de estudo da espectroscopia. Métodos espectroscópicos são aqueles que analisam a medida da quantidade de radiação produzida ou absorvida por moléculas ou espécies atômicas de interesse. Estes métodos são classificados de acordo com a região do espectro eletromagnético envolvido na medida (Settle, 1997).

Dentre os tipos de interação estudados pela espectroscopia, os de maior interesse envolvem as transições entre diferentes níveis energéticos das espécies químicas. Outros tipos de interações como reflexão, refração, espalhamento elástico, interferência e difração são frequentemente mais relacionados às alterações das propriedades globais dos materiais do que com os níveis energéticos de moléculas ou átomos. A energia da radiação empregada e o modo de detecção interferem fortemente no tipo de interação observada.

O maior apelo para o uso de métodos óticos como técnicas analíticas se deve a sua: (i) rapidez, (ii) baixa necessidade de preparação de amostras, uma vez que em geral são técnicas não invasivas e não destrutivas, e (iii) a possibilidade do desenvolvimento de sensores em linha.

Entre os métodos óticos de maior destaque no ramo industrial, é possível citar espectroscopia de luminescência molar (incluindo medidas por fluorescência,

fosforescência e quimiluminescência), a espectroscopia vibracional (espectrometria no infravermelho próximo, médio e distante), e as técnicas baseadas em espectroscopia *Raman* (Skoog, Holler e Crouch, 2007).

2.3.1 Fundamentos da Espectrometria no Infravermelho Próximo - NIR

É inquestionável o fato de que pesquisas envolvendo técnicas baseadas em espectroscopia no infravermelho têm crescido consideravelmente no decorrer das últimas décadas. Atualmente, pode-se dizer que se trata de um dos ramos da química analítica de desenvolvimento mais intenso.

Diversas aplicações podem ser relatadas, tanto para análises qualitativas, como quantitativas. Este aprimoramento pode ser diretamente correlacionado com o desenvolvimento de novos equipamentos robustos e de alto desempenho (tanto espectrofotômetros como computadores), bem como o desenvolvimento de metodologias matemáticas que permitem a calibração de sistemas multivariados de elevada complexidade.

A região espectral do infravermelho (do inglês *Infrared - IR*) compreende a radiação com comprimento de onda de 0,78 a 1000 μm . O espectro IR é normalmente subdividido em três regiões, cada qual com técnicas e aplicações consideravelmente diferentes, são elas IR - próximo (NIR), classificado como a região das harmônicas e sobretons entre 0,78 a 2,5 μm , IR - médio (MIR), classificado como a região de vibração - rotação e compreendido entre 2.5 a 50 μm e IR - distante, classificado como região de rotação e compreendido entre 50 a 1000 μm .

Os espectros de absorção, emissão e reflexão no infravermelho podem ser entendidos assumindo que todos se originam de numerosas variações de energia produzidas por transições de moléculas de um estado vibracional ou rotacional para outro (Skoog, Holler e Crouch, 2007).

Para compreender a interação que existe entre uma molécula e a radiação IR, é necessário o estudo do comportamento vibracional das ligações químicas. Em geral, o modelo de oscilador anarmônico é utilizado para ilustrar a vibração da ligação de uma molécula diatômica, onde existem diferenças entre os espaçamentos dos níveis energéticos e transições de energia não adjacentes são permitidas (Skoog, Holler e Crouch, 2007).

De acordo com a distribuição de Boltzmann, a maioria das moléculas na temperatura ambiente se encontra no chamado estado fundamental, com número quântico vibracional igual a zero ($\nu = 0$). Uma transição para um nível energético do tipo $\nu = 0$ para $\nu = 1$ é denominada fundamental. Transições do tipo $\nu = 0$ para $\nu = 2$ ou 3 são denominadas *sobretons* ou bandas de combinações entre vibrações.

Para que uma molécula absorva radiação IR é necessário que ela receba radiação em uma frequência capaz de fornecer exatamente a energia requerida para a transição entre dois níveis energéticos distintos ou para a ocorrência de sobretons. Qualquer

radiação com frequência diferente desta não será absorvida. Além disso, para que uma molécula possa absorver radiação IR, o deslocamento dos átomos em vibração precisa produzir uma alteração no momento dipolar da molécula ou grupo de átomos, sendo esta alteração de momento a responsável por permitir a interação do componente elétrico da radiação eletromagnética incidente com a molécula (Bueno, 2011).

A região espectral do IR - próximo (*Near Infrared* - NIR) estende-se desde a região final do visível de maior comprimento de onda de 780 nm até 2500 nm. As bandas de absorção nesta região são quase em sua totalidade sobretons ou combinações de vibrações fundamentais de estiramento que ocorrem na região de 3000 a 1700 cm^{-1} . Considerando que as bandas de sobretons e combinações possuem baixas absorvidades molares (de 10 a 1000 vezes mais fracas que bandas fundamentais), os limites de detecção são da ordem de 0,1%. As ligações envolvidas são geralmente C-H, N-H, O-H e S-H (Skoog, Holler e Crouch, 2007).

A grande versatilidade da técnica de espectroscopia NIR pode ser avaliada pela grande gama de diferentes áreas do conhecimento que a utilizam. Existem trabalhos publicados em ciências agrícolas (Wilson, Zhang e Kovacs, 2014), animais (Valenti *et al.*, 2013), biotecnologia (Yu *et al.*, 2014), classificação de frações de petróleo (Cramer *et al.*, 2009), química fina (Shinzawa *et al.*, 2012), ciências médicas (Vergouw *et al.*, 2014), e são usualmente utilizadas em determinações quantitativas de espécies como água, proteínas, gorduras, hidrocarbonetos de baixa massa molecular e outros produtos industriais.

Embora a grande complexidade dos espectros NIR dificultem a elucidação estrutural, algumas vantagens da espectroscopia NIR podem ser destacadas, tais como:

- (i) A não necessidade de preparo das amostras;
- (ii) rapidez na obtenção dos espectros e resultados;
- (iii) ser um método não invasivo e não destrutivo;
- (iv) grande gama de aplicações, considerando que quase a universalidade de moléculas que possuem ligações C - H , N - H, S - H e O - H podem ser analisadas; e
- (v) a possibilidade em se desenvolver analisadores de processo, diretamente em linhas de produção (Ozaki, 2012).

A possibilidade do uso de espectroscopia NIR para o desenvolvimento de analisadores em linha se deve principalmente a possibilidade de análise remota, com o uso de fibra ótica, tornando possível seu uso em sistemas tóxicos, perigosos ou explosivos (grande parte da indústria petroquímica) e o fato de um mesmo espectro NIR, com a manipulação matemática correta, conter mais de uma informação relevante, como por exemplo: concentração de enxofre em diesel (Breitkreitz *et al.*, 2003); caracterização de correntes de petróleo (Pasquini e Bueno, 2007); octanagem, concentração de aromáticos e olefinas em gasolina (Bueno, 2011).

Na questão pela busca quantitativa de enxofre em combustíveis, a espectroscopia NIR se mostra muito efetiva para amostras com alta concentração de enxofre, com concentração de no mínimo 3500 ppm, obtendo resultados comparáveis aos métodos padrões de aferição regulamentados pela ASTM (Breitkreitz *et al.*, 2003). Entretanto, mesmo sendo possível utilizar espectroscopia NIR para classificar diesel que sofreu ou não forte hidrotreatamento (Cramer *et al.*, 2009), as bandas de absorção emitidas por compostos organo-sulfurados são pouco significativas, e em baixas concentrações (<15 ppm) ainda não existem modelos ou métodos capazes de prever e quantificar satisfatoriamente enxofre total em combustíveis.

Um maior aprofundamento dos mecanismos envolvidos nas técnicas de espectrometria no infravermelho pode ser encontrado em Bueno (2011), Skoog *et al.*, (2007) e Siesler *et al.*, (2002).

2.3.2 Espectroscopia por Fluorescência

Ao longo dos últimos 50 anos, houve um aumento significativo no uso de técnicas baseadas em fluorescência em ciências biológicas, levando a espectroscopia fluorescente a ser considerada uma das ferramentas de significativa importância no campo de pesquisas bioquímicas e biofísicas. É metodologia dominante aplicada a pesquisas biotecnológicas, sequenciamento de DNA, análises genéticas e forenses, dentre outras (Lindemann *et al.*, 1998).

Dentre as vantagens apresentadas pelas técnicas, pode-se citar: alta sensibilidade para detecção de metabólitos e componentes em solução, praticidade de operação, uma vez que não exigem o emprego de metodologias caras ou mão de obra especializada, e a não necessidade do uso de traçadores radioativos para a maioria dos compostos medidos. Os métodos óticos são em sua maioria não invasivos, não destrutivos e possuem grande potencial para o desenvolvimento de sensores online (Lakowicz, 1985).

Luminescência é caracterizada como a emissão de luz de uma substância qualquer que ocorre em decorrência de estados eletrônicos excitados. Este fenômeno é geralmente subdividido em três categorias (fluorescência, fosforescência e quimiluminescência), dependendo da natureza dos estados excitados que dão origem a este fenômeno (Williams, 1981).

A fluorescência e a fosforescência são fenômenos baseados em mecanismos semelhantes, assim, frequentemente são denominados pelo termo mais geral "fotoluminescência". Os dois fenômenos diferem uma vez que a fluorescência não envolve uma mudança no spin eletrônico, ao contrário do que acontece na fosforescência. Por este motivo, os tempos de vida dos estados excitados da fluorescência são significativamente menores que os envolvidos no fenômeno da fosforescência. Usualmente, a fotoluminescência é emitida em comprimentos de onda maiores do que os da radiação de excitação (Skoog, Holler e Crouch, 2007).

A Figura 2-5 é uma representação gráfica da diferença entre os estados eletrônicos fundamental, excitado singleto e excitado tripleto (Skoog, Holler e Crouch, 2007).

Um estado molecular onde todos os spins eletrônicos estão emparelhados é denominado estado singleto, e neste caso, não existe nenhum desdobramento dos níveis eletrônicos de energia se o mesmo for exposto a um campo eletromagnético. Quando um dos pares de elétrons de uma molécula é excitado a um nível mais alto de energia, forma-se um estado excitado singleto ou tripleto. No estado excitado singleto, o spin do elétron promovido ainda está emparelhado com o elétron do estado fundamental. No caso do tripleto, os spins dos dois elétrons tornam-se desemparelhados, permanecendo paralelos.

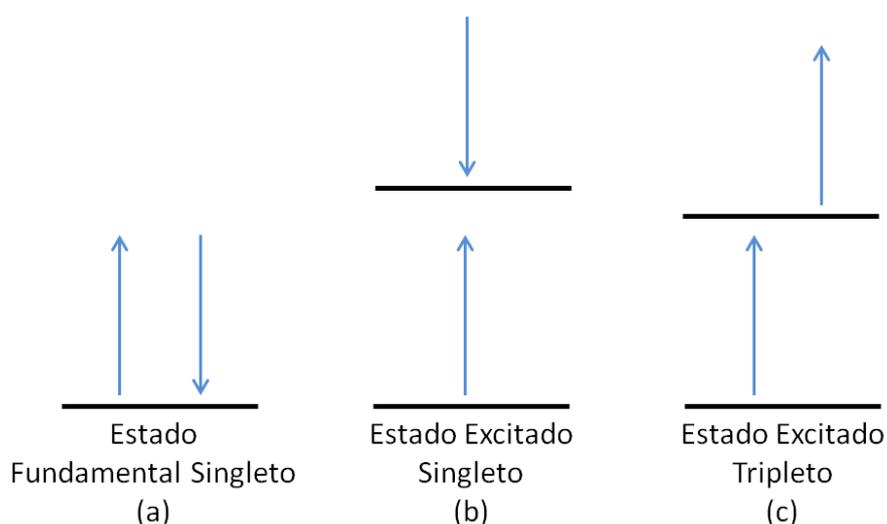


Figura 2-5: Estados eletrônicos do spin das moléculas. Em (a) é mostrado o estado eletrônico fundamental Singlete. Em (b) e (c) são mostrados os estados eletrônicos excitados. Quando os spins permanecem emparelhados, a molécula encontra-se em um estado excitado singleto (b). Se os spins tornam-se desemparelhados (c), a molécula está em um estado excitado tripleto. Fonte: (Skoog, Holler e Crouch, 2007).

Pode-se observar que a energia do estado excitado tripleto é menor que a do estado excitado singleto. As propriedades de uma molécula no estado excitado singleto diferem significativamente daquelas em estado excitado tripleto. Por exemplo, uma molécula é paramagnética em estado tripleto e diamagnética em estado singleto. Vale ressaltar que uma transição singleto-tripletto (que envolve uma mudança no estado eletrônico) é um evento significativamente menos provável que aquele correspondente à transição singleto-singleto (Lakowicz, 1985).

O Diagrama de Jablonski, apresentado na Figura 2-6, ilustra o processo entre a absorção e emissão de luz, que serve como base para o entendimento dos fenômenos de luminescência. A linha horizontal mais espessa na parte inferior do diagrama representa o estado fundamental S_0 da molécula, que a temperatura ambiente pode ser entendida como a energia média das moléculas da solução. As linhas espessas superiores representam os níveis de energia de três estados eletrônicos fundamentais excitados, são eles: S_1 e S_2 estados eletrônicos singletos e T_1 o primeiro estado eletrônico tripleto (Lakowicz, 1985).

Diversos níveis energéticos vibracionais estão associados a cada um dos estados eletrônicos, conforme ilustrado pelas linhas mais finas. Como a excitação do estado fundamental para o estado tripleto envolve uma transição de pequena probabilidade de ocorrência, a mesma não é mostrada no Diagrama (Lakowicz, 2006).

Um das características altamente atraentes para o uso de métodos luminescentes é a sua inerente sensibilidade, com limites de detecção até três ordens de magnitude melhores que aqueles encontrados na espectroscopia de absorção. Em casos específicos e sob condições controladas, é possível detectar uma única molécula em solução por espectroscopia de fluorescência. Outra característica de interesse nos métodos fotoluminescentes é a alta faixa linear de concentrações (Whitford e Julien, 2007).

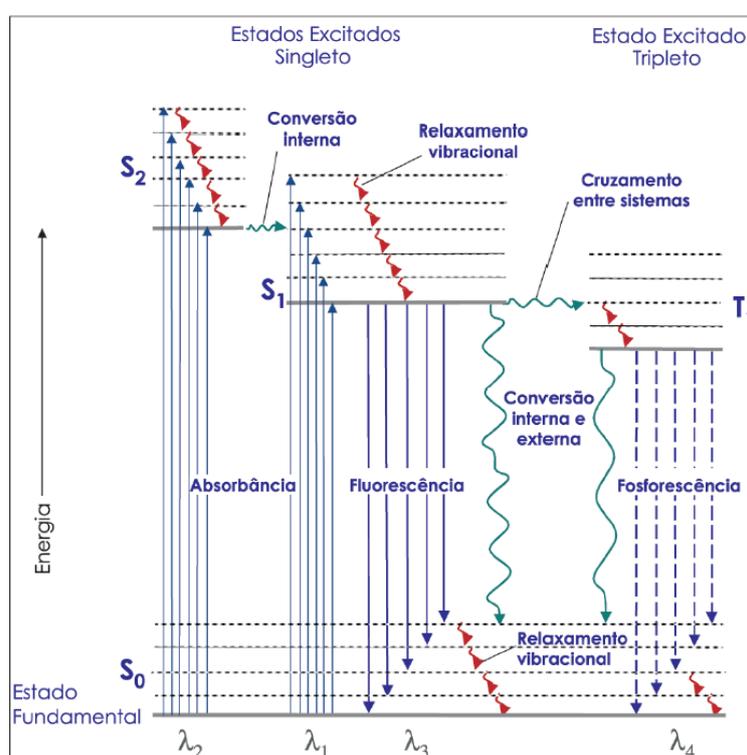


Figura 2-6: Diagrama parcial de níveis de energia para um sistema fotoluminescente. Fonte: (Sotomayor *et al.*, 2008).

O fenômeno da fluorescência é aquele observado em estados excitados do tipo singlete (aquele onde o elétron presente no orbital excitado está pareado com o elétron no estado fundamental). Ao retornar ao estado fundamental, o elétron excitado emite rapidamente energia, na forma de fóton. A taxa de emissão de fluorescência é tipicamente de 10^8 Hz, resultando em um tempo de meia vida médio de aproximadamente 10^{-8} segundos. Devido ao pequeno tempo de emissão de fluorescência, tais medidas requerem equipamentos ópticos e eletrônicos sofisticados (Wang *et al.*, 2011).

O fenômeno da fluorescência ocorre tipicamente em moléculas contendo grupos aromáticos com transições de baixa energia. Compostos alifáticos ou estruturas com ligações duplas altamente conjugadas também podem apresentar fluorescência (Croney, Jameson e Learmonth, 2001).

Os dados de espectroscopia fluorescente são no geral apresentados no chamado Espectro de Fluorescência, na forma tridimensional ou em gráficos de contorno. Ambos gráficos apresentam o sinal de luminescência em função dos comprimentos de onda de excitação e dos comprimentos de onda de emissão, de forma simultânea. O conjunto de dados que da origem a este gráfico é referido usualmente como Matriz de Excitação/Emissão. A Figura 2-7 exemplifica um de cada um destes Espectros Fluorescentes 2D, onde (a) se refere ao espectro fluorescente total de uma mistura de antraceno e de ovaleno, na forma tridimensional e (b) um gráfico de fluorescência de 8-hidroxibenzopireno na forma de linhas de contorno (Skoog, Holler e Crouch, 2007).

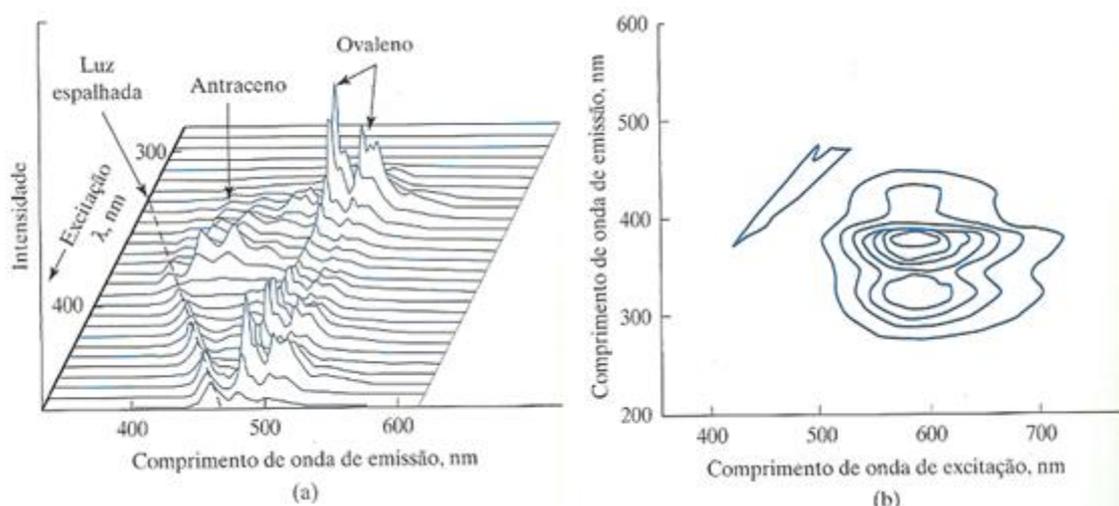


Figura 2-7: Espectros de Luminescência Total para uma mistura de antraceno e ovaleno (a), e para 8-hidroxibenzopireno (b). Fonte: (Skoog, Holler e Crouch, 2007).

O uso de espectroscopia de fluorescência na área de análise de combustíveis tem apresentado bons resultados, uma vez que diversos dos componentes de interesse emitem fluorescência, descartando a necessidade de preparação de amostra e possibilitando o desenvolvimento de pesquisas para construção de analisadores em linha (Riveros *et al.*, 2006; Pantoja *et al.*, 2011).

Em trabalhos como o apresentado por Aburto *et al.*, (2014), resultados promissores são discutidos para busca específica de enxofre em óleo diesel com o uso de espectros de fluorescência e métodos quimiométricos. No referido trabalho, porém, o uso de enzimas para prévia oxidação dos compostos sulfurados torna o método impróprio para a aplicação na caracterização on-line de processos industriais.

2.4 Quimiometria

A quimiometria diz respeito à aplicação de ferramentas matemáticas e estatísticas para o planejamento, otimização e extração de informação útil partindo de um conjunto de dados multivariáveis. Foi introduzida no final dos anos 60, por grupos de pesquisadores das áreas de química analítica e físico-química orgânica, devido à disponibilidade de uma gama de equipamentos de análise que geravam respostas multivariadas (Geladi, 2003).

O surgimento da quimiometria se deve ao fato das análises estatísticas monovariáveis não serem capazes de descrever satisfatoriamente processos químicos multivariados, obrigando os pesquisadores a desenvolverem técnicas que atuassem de forma exploratória para a análise e calibração multivariável de modelos. A quimiometria busca assim encontrar relações entre variáveis mensuradas e amostras, convertendo essas informações em vetores latentes (Kumar *et al.*, 2014).

Em relação à análise exploratória, busca-se encontrar quais as variáveis que mais afetam determinado processo, bem como as interações entre as mesmas, de forma a determinar as melhores condições de análise.

De maneira geral, os métodos são divididos em métodos:

- a) Supervisionados. Possuem a desvantagem de necessitar informações prévias a respeito das amostras para possibilitar a formação de classes e o desenvolvimento de modelos baseados nas informações contidas nas mesmas. Exemplo: Análise Discriminante com Calibração Multivariável por Mínimos Quadrados Parciais (PLS-DA - *Partial Least Square Discrimination Analysis*)
- b) Não Supervisionados. A separação de classes se baseia em similaridades estatísticas dos dados, sem a necessidade de informação sobre a natureza das amostras. Exemplos: Análise de Componentes Principais (PCA - *Principal Component Analysis*). O objetivo destes métodos é a identificação de agrupamentos naturais entre as amostras e o reconhecimento de tendências e padrões (Brereton, 2007).

A calibração multivariada por sua vez tem o objetivo de estabelecer modelos capazes de relacionar uma elevada gama de medidas químicas, físico-químicas ou espectrais de uma certa amostra para inferir os valores de uma determinada propriedade de interesse. Dentre os principais métodos utilizados, pode-se citar Regressão Multivariável, Regressão de Componentes Principais (PCR - *Principal Component Regression*) e Regressão de Mínimos Quadrados Parciais (PLSR - *Partial Least Squares Regression*) (Wehrens e Springerlink, 2011).

Em se tratando de análises quimiométricas de dados espectrais, duas abordagens podem ser seguidas: a primeira faz uso de todo o espectro e a segunda se baseia na

seleção específica de variáveis com o uso de métodos de otimização para encontrar os elementos espectrais mais correlacionados com a variável de interesse.

2.4.1 *Uso de dados espectrais não segmentados*

Dentre os métodos mais utilizados, são citadas duas técnicas amplamente utilizadas na química analítica, PCA e PLS.

PCA é uma ferramenta quimiométrica que busca evidenciar similaridades ou diferenças em amostras em um determinado conjunto de dados. Este método define novas variáveis, obtidas através da combinação linear dos originais, de forma que a direção do primeiro novo eixo contém a maior variância explicada dos dados originais. Cada nova variável subsequente é ortogonal às variáveis antecedentes, mas sempre na direção que contém a maior variação remanescente. Desta forma, o resultado da análise PCA é um conjunto de vetores ortogonais entre si, chamados de Componentes Principais (PC - *Principal Component*), resultantes do agrupamento de variáveis altamente correlacionadas (Beebe, Pell e Seasholtz, 1998). Resumidamente, a PCA reduz a dimensionalidade do conjunto de dados original, condensando as informações mais relevantes contidas na matriz de dados originais nos primeiros componentes principais (Reddy, 2011).

A ideia por trás do uso de PCA vem do fato de que sistemas de grandes dimensões são compostos por muitas variáveis que agregam pouca ou nenhuma informação útil. Em uma análise detalhada de espectros de alta resolução, por exemplo, comprimentos de onda em vizinhanças próximas são altamente correlacionados e contêm informações similares, que poderiam ser filtradas escolhendo os comprimentos de onda que possuem a maior informação, ou então aqueles que se diferem dos demais.

PCA é uma ferramenta padrão para a análise de dados moderna por ser simples, possuir uma única solução analítica e geralmente levar a uma representação dos dados mais simples de ser interpretada. A desvantagem do método, por sua vez, é que o mesmo não produz como resultado um pequeno grupo de variáveis originais que carregam em si a informação útil do sistema, mas sim um pequeno grupo de PCs (novas variáveis criadas), nos quais todas as variáveis de entrada estão representadas (Shlens, 2009).

A análise PCA, por fim, permite identificar o grau de correlação entre as variáveis de entrada e permite determinar (pelo menos de forma aproximada) quantas das mesmas são necessárias para descrever a informação contida no conjunto de dados, ou o problema em questão, de maneira satisfatória. Assim, se um número qualquer de PCs exprime toda a informação independente em um grande conjunto de dados, é provável que este mesmo número das variáveis originais possa ser usado para modelar o sistema em questão de maneira satisfatória.

A metodologia PLS pode ser vista como um melhoramento das técnicas PCA. Enquanto a PCA trabalha apenas com a matriz de variáveis de entrada (**X**), o PLS utiliza também as informações da matriz de variáveis auxiliares a serem modeladas (**Y**) no procedimento de determinação das variáveis principais do sistema, chamadas de variáveis latentes. Este método aplica transformações lineares nas matrizes **X** e **Y** de

forma a buscar novas direções que maximizam a relação entre as duas matrizes (Dayal e Macgregor, 1997).

A metodologia PLS apresenta algumas vantagens que motivam sua vasta aplicação nas áreas de modelagem quimiométrica de processos, como o fato de ser resistente ao chamado *overfitting*, situação onde modelos estatísticos se ajustam de forma demasiada ao conjunto de calibração, se ajustando a erros de medição e fatores aleatórios presentes no conjunto, mas que não são representativas do processo como um todo. (Land Jr *et al.*, 2011).

Da mesma forma que a PCA, a metodologia PLS possui a desvantagem de não trazer como resposta diretamente quais das variáveis de entrada são as mais indicadas para descrever e modelar o sistema, mas sim quantos vetores latentes são necessários para descrever de forma satisfatória o sistema.

2.4.2 Seleção de Variáveis - Métodos de Otimização Global e Segmentação de Dados Espectrais

A seleção de variáveis pode ser considerada uma das mais importantes etapas em diversos tipos de modelagem, tendo relação direta com a eficiência do resultado. Os métodos anteriormente mencionados possuem capacidade de predição vinculada à seleção de variáveis independentes. Na abordagem sem supervisão, variáveis pouco informativas podem mascarar os dados reais de processo. Em casos supervisionados (procedimentos de classificação e regressão combinados), existe a possibilidade de variáveis pouco correlacionadas (variáveis dependentes) levarem a modelos com poder preditivo insatisfatório (Facchin, 2005).

A seleção de variáveis, então, tem o objetivo de reduzir o número de variáveis independentes selecionadas para apenas aquelas com informação relevante, melhorando a modelagem estatística. A base para esta seleção é feita em termos de desempenho preditivo e a capacidade de interpretação de resultados. Um exemplo simples de seleção de variáveis é o uso de picos em dados de espectroscopia, onde a combinação dos elementos espectrais que apresentam picos pode ser usada para quantificar o(s) analito(s) de interesse.

Usualmente, considera-se que a melhor forma de garantir a capacidade de modelagem de um grupo de variáveis é através do seu teste efetivo. No caso do modelo não apresentar resultados satisfatórios, a solução seria então testar um novo grupo de variáveis. Assumindo uma forma de quantificar o erro de medição, a comunidade científica vem aplicando métodos de otimização na procura dos grupos de variáveis que fornecem a melhor capacidade preditiva para os referidos sistemas (Hemmateenejad *et al.*, 2011; Kumar *et al.*, 2014).

Mesmo com a elevada velocidade computacional das máquinas modernas, casos com quantidades moderadas de variáveis não permitem a exploração de todas as possíveis combinações formadas pela combinação de todos seus elementos.

Métodos do tipo *Stepwise*, que limitam as regiões de busca espectral de forma a viabilizar a busca exaustiva (análise de todas as combinações possíveis do grupo de elementos) podem ser utilizados, entretanto apresentam limitações, uma vez que podem desprezar zonas significativas, fornecendo (na maioria dos casos de grandes conjuntos de dados) soluções formadas por ótimos locais (Roy *et al.*, 2009).

Uma das alternativas para resolver o problema de atração em mínimos locais, é o uso de fatores de busca randômica, onde os conjuntos de variáveis são escolhidos de forma aleatória. No entanto, este método aplicado individualmente possui chances quase nulas de encontrar o ótimo global.

Doravante, a forma mais indicada para busca de elementos espectrais recai em uma combinação das duas técnicas, associando a busca com fatores randômicos e o incremento gradual da informação. Esta combinação pode ser feita por meio de algoritmos de otimização global que combinem soluções de elevada qualidade, mas que possam descartar as mesmas, caso necessário, para sair dos arredores de ótimos locais. Os algoritmos de otimização mais aplicados aos ramos da quimiometria são o Algoritmo Genético (GA - *Genetic Algorithm*) e métodos baseados em exames, como Nuvem de Partículas (*Particle Swarm Optimization Algorithm*) (Gosciniak, 2015) e Colônia de Formigas (*Ant Colony Optimization*) (Mullen *et al.*, 2009).

Dentre os métodos de otimização global mais empregados na seleção de variáveis pode-se citar os Algoritmos Genéticos (GA - *Genetic Algorithm*). A principal ideia por trás dos métodos baseados em GA é tratar a população de possíveis soluções como vetores (cromossomos) formados por valores binários (genes), e através de simulação do que seria a reprodução sexual, combinar os melhores indivíduos. Assim, os indivíduos resultantes são formados por combinações das soluções contidas em seus genitores. A qualidade dos descendentes é medida em uma etapa de avaliação, chamada de aptidão (*fitness*). De forma generalizada, vetores com baixa aptidão terão nenhuma ou pouca probabilidade de reprodução, fazendo com que as gerações futuras apresentem cada vez melhores soluções (seleção natural) (Goldberg, 1989).

Diversos trabalhos envolvendo espectroscopia usam GA para busca de comprimentos de onda mais significantes para calibração de modelos quantitativos, dentre os quais pode ser citado o trabalho de Breitzkreitz (2003), onde determinação total de enxofre em diesel é bucada utilizando NIR e GA para seleção de variáveis.

Uma metodologia alternativa de otimização global possível de ser aplicada para seleção de variáveis é a Otimização do tipo Colônia de Formigas (ACO - *Ant Colony Optimization*). O algoritmo foi concebido com base no comportamento coletivo de formigas que saem em busca de alimento.

Inicialmente, as formigas ocupam todos os possíveis caminhos entre o ninho e a fonte de alimento, e com o decorrer do tempo, todas as formigas ocupam somente o menor caminho entre os dois pontos. Este comportamento é possível com o uso de feromônios, que evaporam com o tempo, e funcionam como chamariz para as demais formigas. As formigas que percorrem o menor caminho retornam ao formigueiro mais

rapidamente, e assim a trilha percorrida por estes indivíduos apresenta maior concentração de feromônio (Goss *et al.*, 1989).

Durante os últimos anos, diversos trabalhos foram publicados com relação ao uso de ACO para seleção de variáveis com bons resultados, dos quais podem ser citados: Allegrini e Olivieri(2011); Hemmateenejad *et al* (2011); Shamsipur *et al.*(2009) e Ranzan *et al.* (2014). Maiores detalhes desta técnica serão apresentados no capítulo de metodologia, uma vez que esta será utilizada como algoritmo de otimização para seleção de variáveis neste trabalho.

Capítulo 3 – Metodologia

A metodologia apresentada neste capítulo é aplicada a todos os dados de espectroscopia por fluorescência 2D. Todas as rotinas de cálculo foram implementadas no software MATLAB® versão 8.0.0.783 (R2012b).

3.1 Tratamento dos Dados

A partir das matrizes de dados associadas aos espectros de fluorescência 2D das amostras, foi realizado o pré-tratamento destes, preparando os mesmos para a seleção de variáveis e ajuste de modelos matemáticos.

Considerando que a intensidade espectral total é dependente de diversos fatores, desde a natureza da amostra até fatores ambientais como temperatura, os espectros foram normalizados de forma que as intensidades de fluorescência possam ser comparadas. O procedimento de normalização foi realizado utilizando a metodologia SNV (*Standard Normal Variate*), método este usualmente aplicado na manipulação de dados de espectroscopia (Wehrens e Springerlink, 2011). Este método essencialmente faz o escalonamento dos dados espectrais de forma que após a normalização cada espectro apresenta média zero e desvio padrão unitário, de acordo com a Equação 2-1.

$$Z_i = (y_i - \bar{y})/s \quad (2-1)$$

Na Equação 2-1, y_i é o valor da variável a ser normalizada, \bar{y} o valor médio e s o desvio padrão.

Cada um dos conjuntos amostrais (Padrões de Componentes Sulfurados em Diesel, Diesel S10 e Diesel HDT - discutidos nos próximos capítulos) é subdividido em dois grupos distintos - calibração e validação. Esta divisão é realizada seguindo a metodologia *Y-rank* (Facchin, 2005). Este método primeiramente ordena os dados de forma crescente

em relação à variável dependente ou de interesse (neste caso específico, concentração de enxofre). Em seguida, as amostras contidas em cada conjunto amostral são separadas em uma dada proporção, por exemplo, de 2 para 1, colocando-se duas amostras no grupo de calibração e uma no grupo de teste. Para uma melhor representatividade dos modelos, amostras com valores máximos e mínimos da variável de interesse são incluídas no respectivo grupo de calibração. A descrição dos subgrupos de calibração e teste é apresentada nos capítulos a seguir.

3.2 Métricas de Qualidade dos Modelos

Na metodologia empregada, foram usadas as seguintes métricas para quantificar a qualidade preditiva dos modelos: Raiz Quadrada do Erro Médio de Predição (RMSEP - *Root-Mean-Square Error of Prediction*) e o Coeficiente de Determinação R^2 . Uma variação do RMSEP, quando aplicado ao conjunto de dados do grupo de calibração, é denominada Raiz Quadrada do Erro Médio de Calibração (RMSEC), também usado como métrica de avaliação dos modelos.

RMSEP é calculado de acordo com a Equação (3-2), onde \hat{y} representa a variável predita pelo modelo, y a variável medida e n o número de amostras contidas no grupo de amostras de teste. Este critério analisa o ajuste do modelo ao conjunto de dados de teste (ou calibração no caso RMSEC), avaliando a precisão dos resultados.

$$RMSEP = \sqrt{\sum_i^n (\hat{y}_i - y_i)^2 / n} \quad (3-2)$$

O segundo critério para avaliação dos modelos é o R^2 , que indica o quanto dados reais se ajustam ao modelo. Este indicador fornece a informação do quanto os dados observados são replicados pelo modelo. O mesmo foi calculado de acordo com a Equação (3-3), onde \bar{y} representa a média do vetor de dados medidos y .

$$R^2 = 1 - \frac{\sum_i^n (y_i - \hat{y}_i)^2}{\sum_i^n (y_i - \bar{y})^2} \quad (3-3)$$

Durante a rotina de otimização para a busca dos subgrupos de elementos espectrais, a função objetivo implementada é baseada na minimização da soma dos resíduos quadráticos (SSR - *Sum of Square Residuals*) entre os dados de calibração e os dados preditos pelos modelos (Equação 3-3). Apesar de esta grandeza ser utilizada no decorrer do procedimento de otimização, seus valores não são utilizados como forma de comparação entre modelos.

$$SSR = \sum_i^n (\hat{y}_i - \bar{y}_i)^2 \quad (3-3)$$

3.3 Desenvolvimento de Modelos e Seleção de Variáveis

Partindo das matrizes de dados de espectroscopia por fluorescência, previamente normalizadas, e com a definição dos sub grupos de amostras destinadas à calibração e validação, é dado início ao procedimento de seleção de variáveis e desenvolvimento de modelos quimiométricos para a predição do conteúdo de enxofre presente nas referidas amostras.

As matrizes de dados foram construídas transformando todas as matrizes de dados de espectroscopia (em duas dimensões) para matrizes linha, onde cada linha corresponde a uma amostra e cada coluna apresenta a intensidade de fluorescência para um par Ex/Em.

3.3.1 *Principal Component Analysis – Análise dos Componentes Principais*

Como avaliação inicial, a rotina PCA é realizada. Esta etapa busca primeiramente avaliar qual é o possível número de variáveis independentes existentes dentro de cada conjunto espectral. Esta estimativa é baseada na quantidade de PCs que carregam consigo a maior quantidade de variância acumulada (superior a 95%). Esta informação pode ser extrapolada de forma provisória para a quantidade de pares de fluorescência (Excitação/Emissão) necessários para o desenvolvimento de um modelo representativo dos dados amostrais. Esta estimativa inicial serve como base para escolha do tamanho ideal dos modelos quimiométricos. Apesar desta estimativa inicial, diversos tamanhos de modelo (quantidade de pares excitação/emissão) são testados e a qualidade dos mesmos, em função do seu tamanho, é comparada com os resultados preliminares descritos.

A metodologia PCA é utilizada, além de fonte de informação qualitativa a respeito dos dados espectrais, no desenvolvimento de gráficos de escores. Estes gráficos, onde são graficadas as coordenadas de cada amostra em função dos seus componentes principais, permite a avaliação da qualidade dos espectros dentro dos conjuntos amostrais. Este tipo de avaliação, referida como Análise de *Clusters*, permite a segmentação dos conjuntos amostrais de acordo com afinidades dos dados espectrais, indicando amostras qualitativamente equivalentes (Cadima, Cerdeira e Minhoto, 2004).

Essa característica apresentada pelo PCA pode ser considerada uma ferramenta de classificação e segmentação de amostras, sendo viável para a construção de sensores capazes de segmentar amostras desconhecidas dentro de grupos previamente caracterizados (Cramer *et al.*, 2009).

3.3.2 *Pure Spectra Chemometric Modeling - Modelagem Quimiométrica com Componentes Espectrais Puros*

A metodologia escolhida para a análise dos dados espectrais e conversão dos mesmos em informação útil de processo foi a Modelagem Quimiométrica com Componentes Espectrais Puros (*Pure Spectral Chemometric Modeling - PSCM*), que combina modelagem utilizando regressão multilinear e a seleção de elementos espectrais por métodos de otimização estocástica (Ranzan *et al.*, 2014).

A metodologia PSCM utiliza modelos de regressão multilinear (MLR) associados ao método de otimização estocástica ACO (*Ant Colony Optimization*). Os elementos espectrais são filtrados, selecionando grupos de componentes espectrais com maior correlação com as variáveis de estado de interesse. Após a seleção deste grupo de elementos, estes são utilizados como variáveis de entrada em modelos multilíneares, calibrados e então submetidos a testes para predição das variáveis de interesse.

A Figura 3-1 apresenta um esquema do algoritmo implementado para a resolução do problema de seleção de grupo de elementos espectrais utilizando ACO. O algoritmo pode ser subdividido em quatro fases: inicialização das variáveis, inicialização da solução, rotina de busca e apresentação dos resultados.

A fase inicial do algoritmo se caracteriza pela inicialização das variáveis necessárias para resolução do problema de otimização:

- a) dados espectrais, dispostos em uma matriz bidimensional, onde cada linha corresponde aos dados espectrais coletados para cada amostra distinta e cada coluna corresponde a um determinado componente espectral (par excitação/emissão);
- b) vetor coluna das variáveis observadas, onde cada célula corresponde ao valor da variável de interesse à respectiva amostra;
- c) escolha do número de ciclos que o algoritmo realizará para busca do ótimo;
- d) escolha do tamanho do "exército de formigas" utilizado na busca;
- e) dimensão do modelo linear, indicando o tamanho do grupo de elementos espectrais a ser buscado;
- f) tipo de modelo utilizado para avaliação do grupo de elementos espectrais - neste caso, linear;
- g) valor inicial da "trilha de feromônios", de forma que todos os elementos inicializem com a mesma quantidade de marcador;
- h) taxa de evaporação de marcador, multiplicador do vetor de feromônios aplicado a cada ciclo.

Na segunda fase é inicializado o vetor de soluções, através da resolução da função objetivo com a seleção aleatória de um conjunto de elementos espectrais. Esta etapa inicializa o vetor de soluções de forma a propiciar futuras comparações dos resultados contidos neste vetor com os próximos resultados obtidos da otimização, gerados na próxima fase.

A terceira fase é o núcleo do algoritmo de otimização. Durante esta etapa é determinada a melhor combinação de elementos espectrais para predizer a variável de

estado observada. Durante o ciclo do algoritmo, o exército de formigas avalia as possíveis soluções na busca pelo grupo de elementos espectrais que forneça o menor valor para a função objetivo. Neste caso, a função objetivo trata da raiz quadrada do somatório da diferença elevada ao quadrado entre os valores observados da variável de estado e o predito pelo grupo de elementos selecionados, para cada amostra (RMSEC). Cada um dos indivíduos do exército de formigas escolhe um grupo de elementos e os submete ao teste da função objetivo. Caso o resultado seja inferior ao previamente armazenado, este então é substituído e o novo melhor resultado toma lugar no vetor de soluções.

A seleção de cada possível candidato, o elemento espectral constituinte do grupo de predição, é baseada em um fator aleatório conjuntamente com o vetor de marcadores (vetor trilha de feromônios).

O fator aleatório garante que o algoritmo de busca não fique retido em mínimos locais. Na prática, o mesmo é implementado através de um "gatilho randômico" gerando valores aleatórios entre zero e um. O valor fornecido pelo gatilho é comparado com o valor de densidade de feromônios acumulada, de forma que o elemento que apresentar densidade de feromônios igual ao valor fornecido pelo gatilho randômico é inserido no grupo de elementos da solução.

As equações utilizadas para o cálculo da densidade relativa de feromônio ρ_{Fni} e densidade acumulada de feromônios C_{Fni} são apresentadas nas Equações (3-5) e (3-6), onde F_i indica a quantidade de marcador associada ao elemento espectral i e N o número de elementos espectrais totais:

$$Fn_i = \tanh\left(\frac{3 * F_i}{\max(F)}\right) \quad (3-4)$$

$$\rho_{Fni} = Fn_i / \sum_{i=1}^N Fn_i \quad (3-5)$$

$$C_{Fni} = \sum_{j=1}^i \rho_{Fnj} \quad (3-6)$$

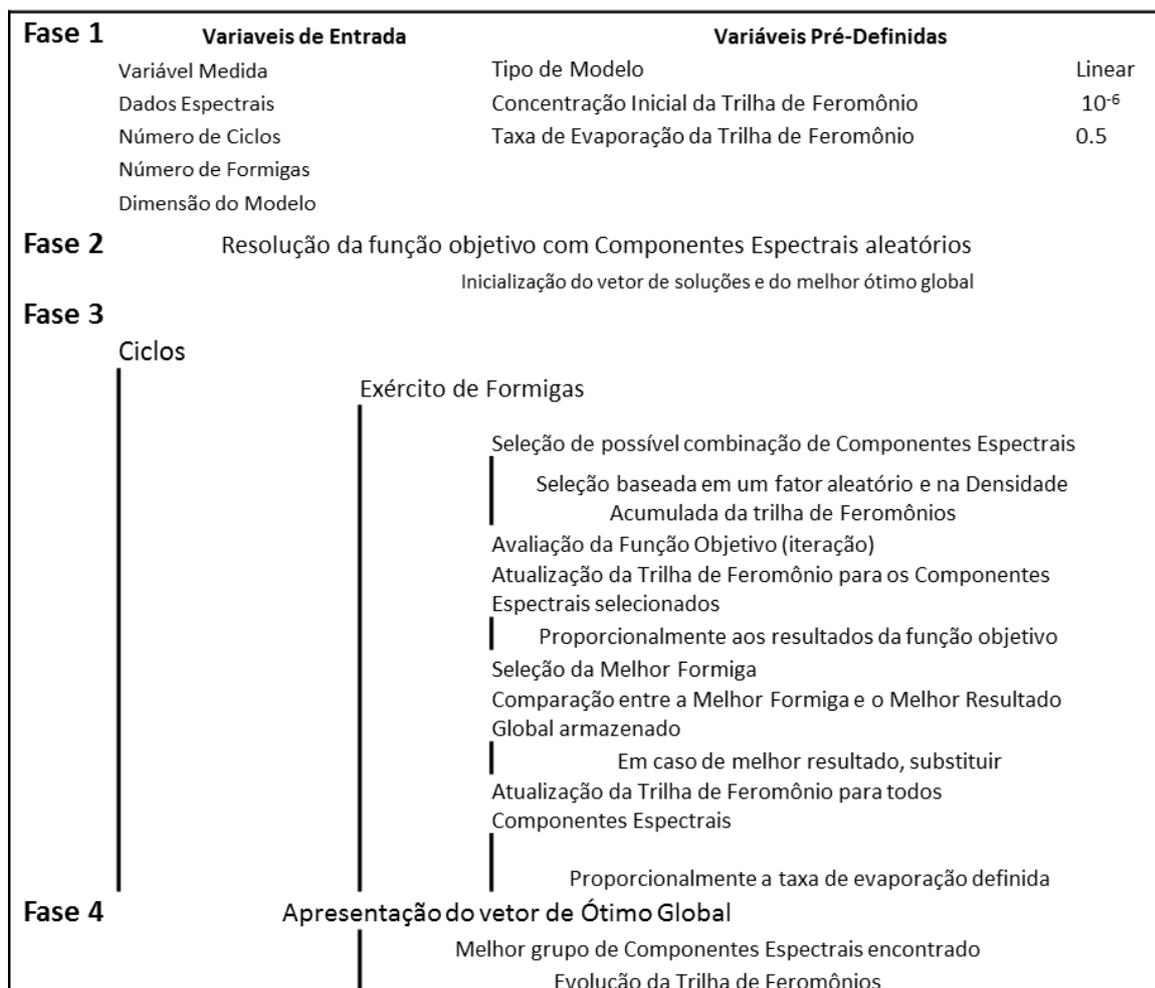


Figura 3-1: Representação esquemática do algoritmo ACO para seleção de grupos de elementos espectrais. Fonte: (Ranzan *et al.*, 2014).

A Figura 3-2 representa a seleção de elementos espectrais da estratégia original de Dorigo e Gambardella(1997).

O fato de o gatilho randômico estar associado à curva C_{Fni} faz com que apenas elementos com elevados valores de F sejam privilegiados em detrimento dos demais, de forma que a probabilidade do algoritmo selecionar elementos com maior concentração de feromônios seja elevada, priorizando a seleção dos elementos espectrais melhores avaliados pelo exército de formigas.

Após cada formiga selecionar e testar um determinado grupo de elementos espectrais, a quantidade de feromônio associada a cada elemento é atualizada de forma inversamente proporcional ao somatório do erro quadrático entre os dados observados e preditos. Quanto maior a raiz quadrada do erro quadrático apresentado pelo modelo, menor o incremento na quantidade de feromônio associada a cada elemento do grupo. A cada fim de ciclo do processo de otimização, todo o vetor de quantidade de marcador

é multiplicado pelo fator que simula a evaporação da trilha de feromônios, penalizando os elementos que não foram selecionados pelo exercito de formigas.

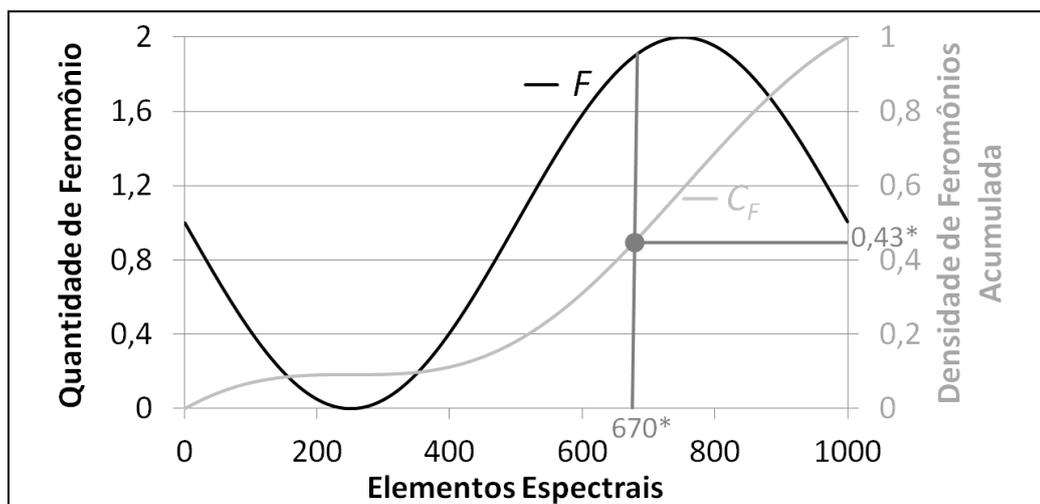


Figura 3-2: Exemplo de seleção de elementos espectrais utilizando a estratégia de Dorigo e Gambardella (1997). F_i corresponde à trilha de feromônios simulada e C_{Fi} à sua respectiva curva de densidade de feromônios acumulada. No exemplo, o gatilho randômico é acionado em 0,43, levando a seleção do respectivo elemento espectral 670. Fonte: (Ranzan *et al.*, 2013)

Ao finalizar o processo de otimização, a trilha de feromônios apresenta forma característica que depende unicamente do tipo de modelo proposto e dos dados de espectroscopia. A forma final desta trilha pode ser considerada uma "assinatura" dos dados de espectroscopia associados à variável de estado, classificando de forma qualitativa as regiões do espectro de fluorescência que tem maior correlação com a variável de estado de interesse.

Na última etapa do processo de seleção de variáveis, são apresentados os resultados da melhor solução obtida durante o processo de otimização. Este conjunto de elementos corresponde à melhor solução testada pelo exército de formigas.

A apresentação final dos resultados traz informações sobre os pares de fluorescência selecionados para cada tamanho de modelo, assim como RMSEC e R^2 do grupo de calibração e seu corresponde RMSEP e R^2 do grupo de teste.

Capítulo 4 – Estudo de Soluções Padrão

Como visto na revisão bibliográfica, são diversos os compostos sulfurados presentes em óleo diesel. Entre tais compostos, alguns grupos como os da família dos dibenzotiofenos são os mais prováveis a resistirem aos processos de hidrotreatamento e são os maiores constituintes sulfurados do óleo diesel final, principalmente em correntes com baixo teor de enxofre como o diesel S10.

Com a intenção de estudar a fluorescência associada aos componentes sulfurados presentes em diesel, quatro diferentes padrões representativos foram adquiridos e avaliados com as técnicas quimiométricas a fim de verificar a viabilidade do uso de espectros de fluorescência em duas dimensões para quantificar a concentração de enxofre em amostras contendo unicamente tais padrões.

Os quatro padrões escolhidos para este estudo foram: (a) Tionafteno 98% (b) Dibenzotiofeno 98% (c) 4,6 - Dimetildibenzotiofeno 97% e (d) 4,6 - Dietildibenzotiofeno 97%, cujas estruturas são mostradas na Figura 4-1. Todos os padrões foram adquiridos da empresa Sigma-Aldrich com CAS: 95-15-8; 132-65-0; 1207-12-1 e 132034-91-4, respectivamente. A escolha destas moléculas foi baseada em avaliações dos trabalhos de (Aburto, 2014) e (Hua, 2003).

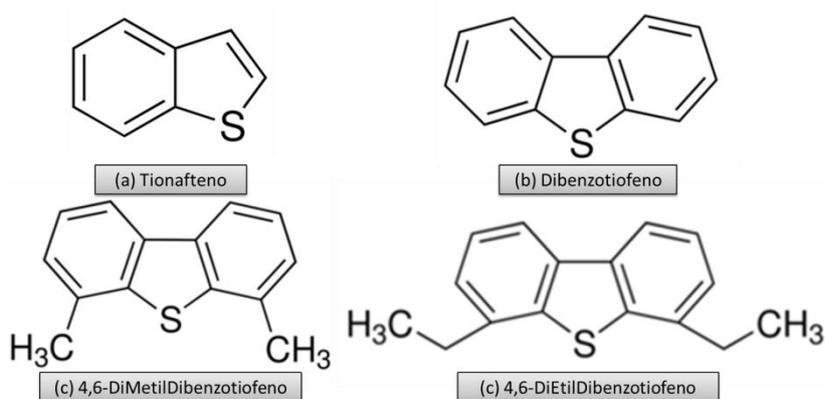


Figura 4-1: Estrutura molecular dos padrões avaliados. Fonte: Sigma-Aldrich

Em se tratando de metodologias não assistidas, PCA foi realizada sobre o conjunto total de grupos de padrões para avaliar a capacidade de segmentar as mesmas com o uso de gráficos de escores. Na metodologia assistida (PSCM), foi realizada a seleção de variáveis para criação de modelos capazes de prever a concentração de enxofre relativa a cada amostra, desenvolvendo modelos específicos para cada um dos quatro padrões utilizando os pares de fluorescência melhores correlacionados à variável de interesse.

4.1 Preparo das Soluções-Padrão

Inicialmente, foi preciso selecionar um solvente com fluorescência baixa ou nula, capaz de solubilizar os quatro compostos escolhidos. Após avaliações da literatura e medidas no espectrofluômetro, o hidrocarboneto escolhido como solvente foi hexano de padrão analítico (marca Mercolab). Os espectros resultantes se mostraram livres de picos, e com pequenas intensidades de fluorescência, comparável ao ruído de medição. Todas as amostras aqui apresentadas e discutidas são soluções compostas pelo padrão adquirido na forma sólida solubilizado com hexano às concentrações desejadas.

A Figura 4-2 apresenta o espectro de fluorescência coletado para o hexano puro. Comparado com a ordem de grandeza da intensidade de fluorescência das soluções-padrão (10^5), a intensidade de fluorescência do hexano não é significativa.

Primeiramente, uma solução inicial com a máxima concentração de enxofre foi preparada dissolvendo os padrões sólidos em hexano, denominada Solução Base. Todas as demais soluções-padrão foram obtidas a partir da diluição com hexano da Solução Base.

A

Tabela **4-1** apresenta as soluções-padrão e a concentração de enxofre pertinente a cada uma. As amostras em negrito e sublinhadas foram selecionadas para compor os grupos de calibração, sendo as demais destinadas para o conjunto de teste.

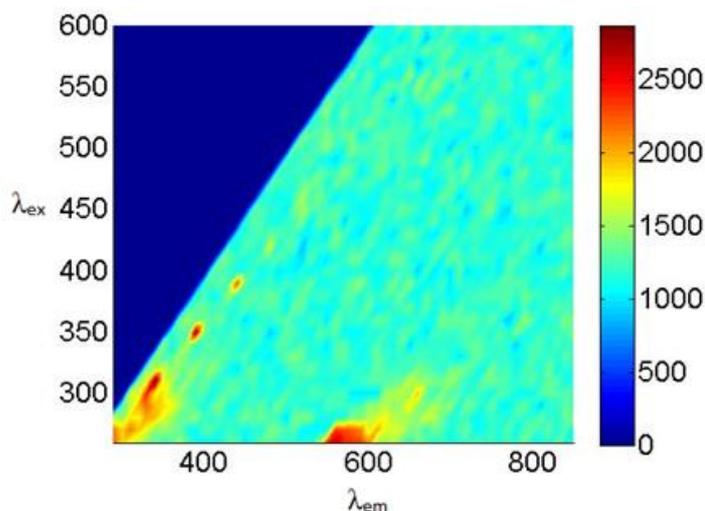


Figura 4-2: Espectro por fluorescência do Hexano.

A concentração das soluções base (C_b) e das demais soluções-padrão (C_a) foram calculadas de acordo com as Equações (4-1) e (4-2) respectivamente. Nestas equações, Mp_i diz respeito à massa de padrão adicionada, Xp_i à porcentagem de enxofre em cada padrão, M_H à massa de hexano adicionada e M_{bi} à massa de solução base adicionada.

$$C_{bi} = \frac{Mp_i * Xp_i}{Mp_i + M_H} \quad (4-1)$$

$$C_{ai} = \frac{C_{bi} * M_{bi}}{M_{bi} + M_H} \quad (4-2)$$

Tabela 4-1: Concentração de enxofre em cada uma das amostras produzidas e a segmentação do conjunto de amostras grupo de teste e calibração (amostras em negrito e sublinhado).

	Tionafteno	Dibenzotiofeno	4,6 - Dimetildibenzotiofeno		4,6 - Dietildibenzotiofeno		Concentração de Enxofre [ppm]	
Soluções Base	<u>524</u>	<u>512</u>	<u>252</u>	<u>202</u>				
Amostras	<u>A1</u>	<u>160</u>	<u>B1</u>	<u>153</u>	<u>C1</u>	<u>151</u>	<u>D1</u>	<u>152</u>
	A2	133	B2	128	C2	125	D2	126
	<u>A3</u>	<u>106</u>	<u>B3</u>	<u>102</u>	<u>C3</u>	<u>100</u>	<u>D3</u>	<u>102</u>
	<u>A4</u>	<u>75</u>	<u>B4</u>	<u>70</u>	<u>C4</u>	<u>70</u>	<u>D4</u>	<u>71</u>
	A5	54	B5	50	C5	50	D5	51
	<u>A6</u>	<u>32</u>	<u>B6</u>	<u>30</u>	<u>C6</u>	<u>30</u>	<u>D6</u>	<u>31</u>
	<u>A7</u>	<u>22</u>	<u>B7</u>	<u>20</u>	<u>C7</u>	<u>20</u>	<u>D7</u>	<u>20</u>
	A8	11	B8	10	C8	10	D8	10
	<u>A9</u>	<u>6</u>	<u>B9</u>	<u>5</u>	<u>C9</u>	<u>5</u>	<u>D9</u>	<u>5</u>

4.2 Coleta dos Espectros de Fluorescência em 2D

Inicialmente todas as soluções-padrão foram ambientadas a uma temperatura média de 25°C, para evitar mudanças no espectro decorrentes da diferença de temperatura entre as amostras.

Os espectros de fluorescência foram medidos com uso do equipamento HORIBA Fluoromax®-4, equipado com lâmpada de xenônio de 150W e faixa de emissão e excitação entre 200nm - 950nm. O equipamento consiste basicamente de uma fonte de excitação, dois monocromadores para seleção de comprimento de onda, um de excitação e outro de emissão e módulo de amostragem com detector de referência e um detector de emissão. Uma cubeta de quartzo com os quatro lados polidos e caminho óptico de 1cm é utilizada como recipiente para acondicionamento da amostra. A Figura 4-3 ilustra os equipamentos utilizados.

Para a coleta dos espectros por fluorescência, as amostras foram excitadas com comprimentos de onda variando entre 260nm e 600nm, utilizando um incremento de 10nm. A emissão por sua vez foi avaliada entre os comprimentos de onda de 290nm a 850nm, com o mesmo incremento. Todas as análises foram realizadas em triplicata, aumentando a confiabilidade dos dados.



Figura 4-3: (a) Espectrofluorômetro Fluoromax-4 e (b) cubeta de quartzo com caminho óptico de 1cm.

Assim, cada espectro de fluorescência conta com 1904 pares de fluorescência correspondentes à intensidade de fluorescência atribuída ao par Excitação/Emissão, distribuídos em uma matriz tridimensional onde os eixos x, y e z são respectivamente: de emissão (λ_{em}), comprimentos de onda de excitação (λ_{ex}), e a intensidade de fluorescência. Uma vez que nenhuma excitação pode levar a uma emissão com maior energia (menor comprimento de onda), todos os espectros são diagonais inferiores,

onde os pares válidos estão situados abaixo da diagonal correspondente ao comprimento de onda de excitação igual ao comprimento de onda de emissão.

A Figura 4-4 traz um exemplo de espectro de fluorescência de duas diferentes amostras de dibenzotiofeno em hexano, onde (a) contém 153 ppm de enxofre e (b) 5 ppm de enxofre, apresentados na forma de curvas de nível suavizadas, onde a terceira dimensão é representada na forma de cores, de acordo com a escala lateral.

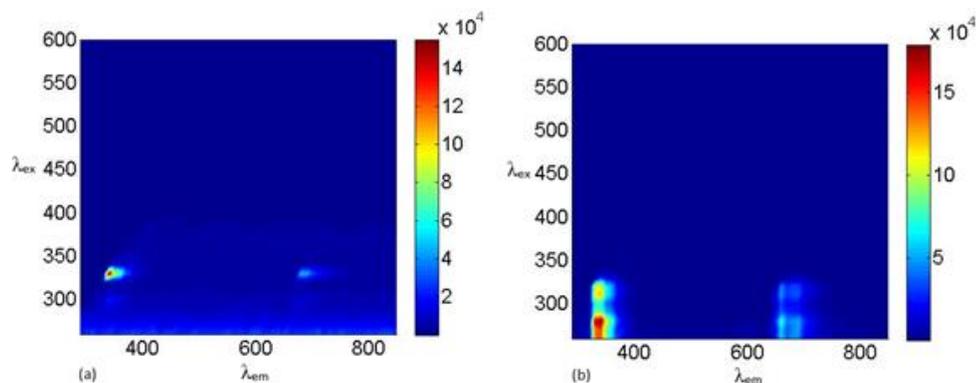


Figura 4-4: Espectros de fluorescência de duas amostras de dibenzotiofeno em hexano. (a) B1 (153 ppm de S) (b) B9 (5 ppm de S).

4.3 Resultados e Discussões

Os resultados das análises dos padrões de enxofre com relação à coleta de dados de espectroscopia por fluorescência 2D serão apresentados em duas etapas. A primeira apresentará os resultados obtidos com a aplicação da metodologia de treinamento não assistida (reconhecimento de padrões baseado apenas nos espectros usando a técnica - PCA), enquanto a segunda parte será focada nos resultados obtidos com a utilização da métrica assistida - PSCM.

4.3.1 Metodologia Não assistida - PCA

Inicialmente, cada padrão foi analisado de forma individual. Sendo assim, somente as dez matrizes de espectroscopia por fluorescência referentes a cada um dos padrões foram utilizadas como dados de entrada da rotina de cálculo. Lembrando que a PCA trata de um método de análise quimiométrica não supervisionada, a informação referente à concentração de enxofre em cada uma das amostras não é adicionada.

Na Figura 4-5 pode ser visto o gráfico de variância acumulada para cada um dos padrões analisados, fornecida pela PCA através dos autovalores da matriz de covariância dos dados de espectroscopia.

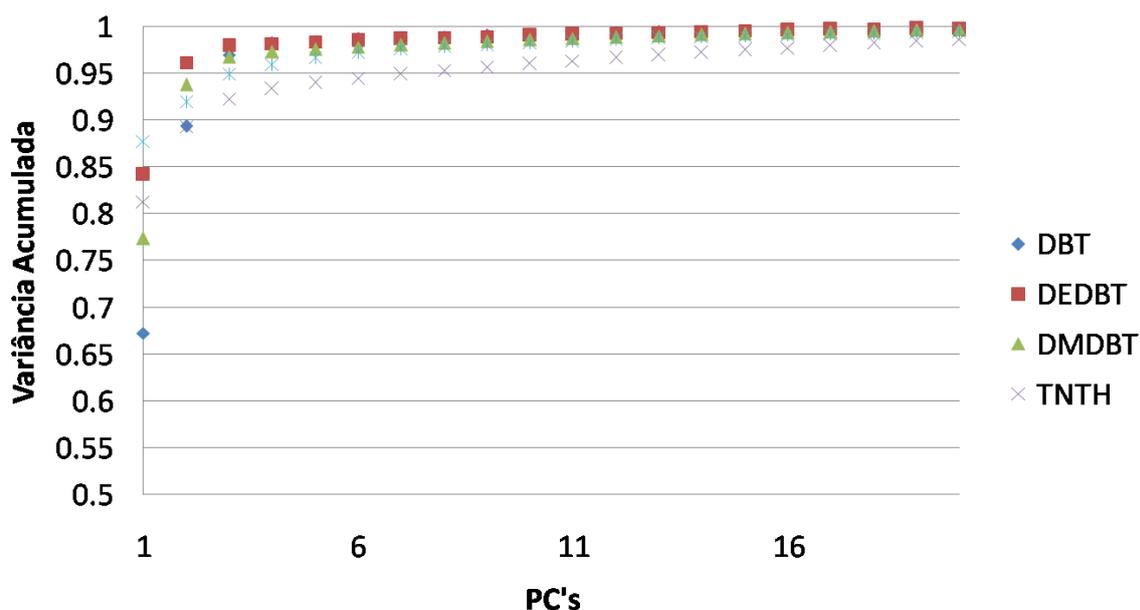


Figura 4-5: Variância Acumulada do grupo de padrões obtida com o aumento do número de componentes principais utilizados para: DBT: dibenzotiofeno; DEDBT: 4,6-dietildibenzotiofeno; DMDBT: 4,6-dimetildibenzotiofeno; TNTH: tionafteno.

Com base na Figura 4-5, pode ser concluído que de forma generalizada os compostos que apresentam a informação contida em seus dados de espectroscopia por fluorescência de forma mais acessível e condensável é: DEDBT, seguido por DMDBT, DBT e por fim TNTH. A Tabela 4-2 apresenta de forma resumida a informação contida na Figura 4-5, salientando o número de componentes principais necessários para expressar mais de 95% da variância contida no conjunto original de dados.

Tabela 4-2: Número mínimo de PCs necessários para expressar mais de 95% da variância contida no conjunto original de dados.

	DBT	DEDBT	DMDBT	TNTH
PCs	3	2	3	7

Assim, podemos concluir que cada conjunto original de dados de DBT e DMDBT possui três variáveis independentes, enquanto DEDBT necessita de apenas dois. Esta informação significa que a matriz de dados de fluorescência destes três elementos pode ser satisfatoriamente comprimida para até três componentes principais, fato que pode ser inferido de forma qualitativa para o número de elementos espectrais necessários para a caracterização destas amostras.

Os resultados apresentados pela análise dos dados de TNTH mostram que apesar de ser necessário um maior número de componentes principais, a PCA foi capaz de

concentrar mais de 95% da variância contida na matriz de dados espectrais em apenas sete elementos espectrais.

A Figura 4-6 apresenta os gráficos de escores. Neste gráfico são apresentadas as componentes de PC1 e PC2 para cada uma das amostras avaliadas individualmente.

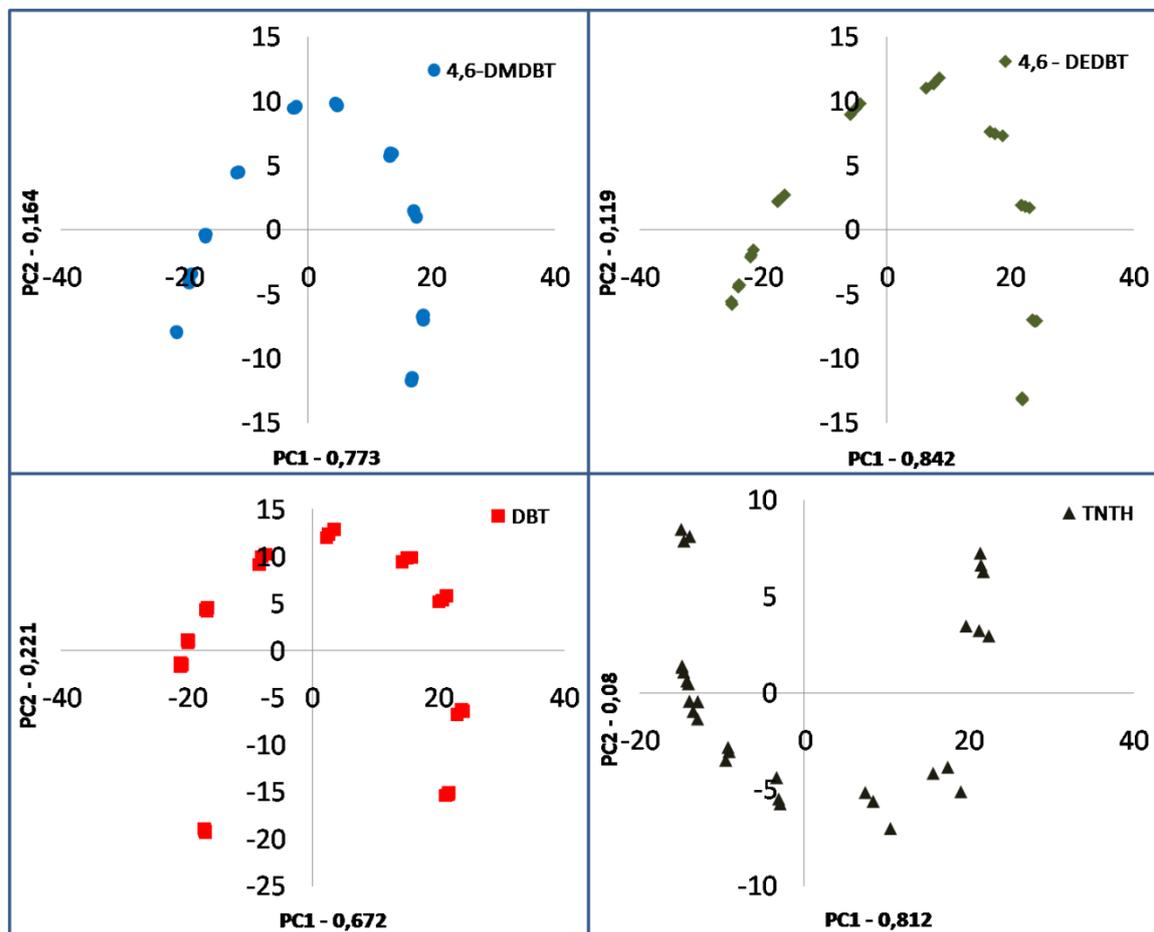


Figura 4-6: Gráfico de Escores dos padrões avaliados individualmente.

O gráfico de escores é usualmente aplicado na análise de representatividade dos dados contidos em uma matriz de dados espectrais. Através deste tipo de análise é possível determinar se os dados amostrais podem ser considerados, de forma qualitativa, equivalentes. A equivalência entre as amostras é visualizada no gráfico de escores, na forma de uma distribuição agrupada das amostras no plano cartesiano.

Os três padrões das famílias dos dibenzotiofenos apresentaram comportamento similar: uma curvatura que difere levemente as soluções-padrão em relação a sua concentração. Sendo a fluorescência uma medida significativamente influenciada pelo meio, um aumento na concentração (e conseqüente aumento no número de moléculas) leva a uma maior absorção interna da fluorescência emitida de uma molécula por outra, assim como a criação de barreiras físicas para emissão de fótons que chegam até o monocromador.

O comportamento apresentado pelas amostras de soluções padrão não pode ser diretamente relacionado com a variação da concentração de enxofre nas amostras avaliadas, uma vez que a variação de enxofre nas amostras segue um comportamento linear que não pode ser identificado nos valores de PC1 versus PC2.

A Figura 4-7 apresenta o gráfico de escores para a PCA dos quatro padrões sendo analisados simultaneamente. Nesta análise, a matriz de dados de entrada contém todas as quarenta matrizes de dados espectrais oriundas das soluções-padrão, possibilitando uma avaliação simultânea das informações contidas nas medidas de espectroscopia por fluorescência 2D. O arranjo da matriz de dados foi construída de forma que cada linha representa uma das 40 amostras e cada coluna apresenta a intensidade de fluorescência dos pares Ex/Em.

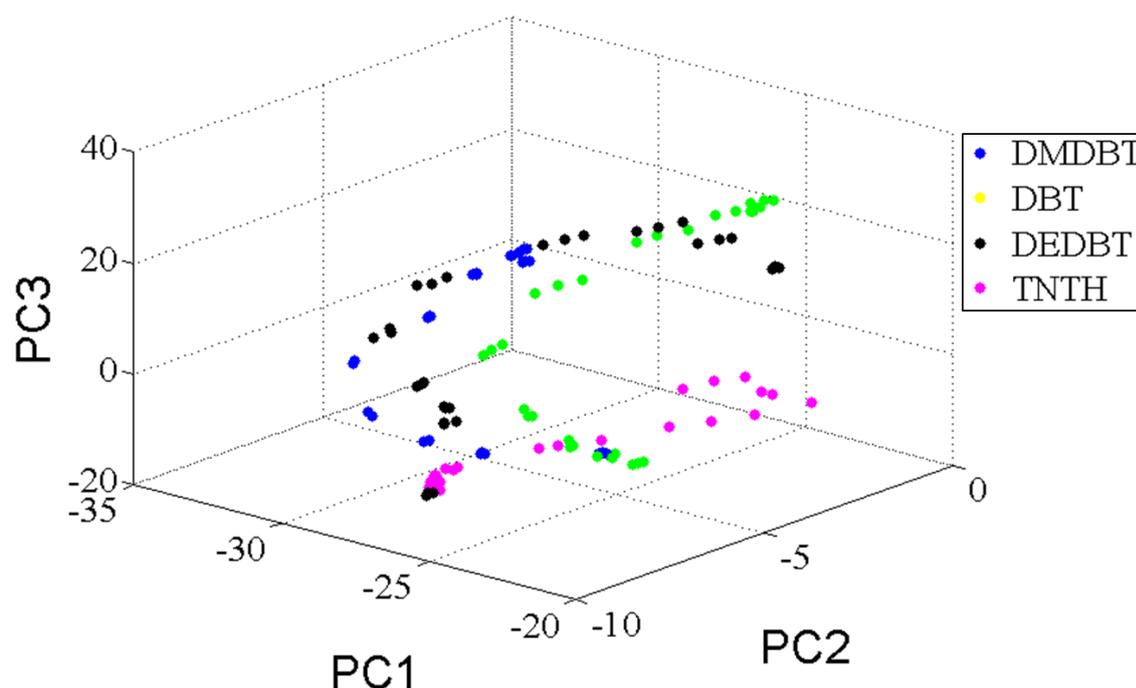


Figura 4-7: Gráfico de Escores da PCA com uso de todos os espectros de soluções-padrão.

Fica visível pela Figura 4-7 que não existe a segmentação em *clusters* dos padrões, indicando que a PCA não encontrou informações contidas nas matrizes de dados de espectroscopia dos padrões capazes de diferenciar o grupo de amostras entre si (i.e., DMDBT, DBT, DEDBT, e TNTH). Entretanto, pode-se verificar a diferenciação até certo grau do efeito das concentrações na localização dos pontos.

Esta similaridade está relacionada com a região espectral de fluorescência de cada padrão, sendo esta semelhante a todos.

As informações contidas nos dados espectrais destes elementos não são suficientes para distingui-los entre si. Esta característica não inviabiliza a utilização de

espectroscopia por fluorescência para determinar concentrações destes elementos em amostras puras dos mesmos, mas inviabiliza a segmentação de cada um destes compostos em correntes contendo misturas destes (como no caso de correntes de Diesel).

Apesar da desvantagem de que todas as moléculas de referência causam o mesmo distúrbio em medidas de espectroscopia por fluorescência, uma vez que todas elas (com exceção do TNTH, que de acordo com a Aburto *et al* (2014), encontra-se em pequenas quantidades em correntes de diesel) fornecem informação equivalentes, a caracterização da concentração de enxofre total em amostras de diesel aparenta ser promissora.

Esse resultado também indica que a quantificação de enxofre total em correntes de ULSD (*Ultra Low Sulfer Diesel*), como o diesel S10, usando espectroscopia por fluorescência deva ser mais factível do que a quantificação de enxofre em correntes de diesel que sofrem hidrotreatamento mais brando.

Isso é devido ao fato que correntes menos tratadas por HDT possuem outros componentes emissores de fluorescência, e um maior número de moléculas que podem causar desativação e interferência do fenômeno de fluorescência, sendo estes removidos durante o processo de HDT, fazendo com que correntes de ULSD apresentem enxofre basicamente devido à presença de moléculas de DBT, DMDBT e DEDBT, que são caracterizadas através das mesmas informações contidas na matriz de fluorescência.

Do ponto de vista prático, o fato destas moléculas apresentarem informações qualitativas equivalentes favorece o projeto e desenvolvimento de sensores baseados em medidas de fluorescência para a caracterização de amostras de ULSD, já que os referidos sensores poderão ser ajustados para medir informações específicas de fluorescência, tornando-os mais precisos e robustos. Essa constatação, entretanto, será verificada no próximo capítulo, onde estudos envolvendo dados de espectroscopia por fluorescência de amostras de Diesel S10 (ULSD) e Diesel HDT serão analisados.

4.3.2 Metodologia Assistida - PSCM

A metodologia PSCM foi aplicada aos dados de espectroscopia por fluorescência 2D das amostras de moléculas padrões contendo enxofre, presentes em correntes de diesel, com dois objetivos principais: (i) determinação das regiões de espectroscopia mais significativas para a caracterização de cada composto de forma individual e (ii) ajuste de modelos multilíneares para a determinação da concentração de enxofre associada a moléculas de DBT, DMDBT, DEDBT e TNTH solubilizadas em hexano, através do uso direto de pares de fluorescência.

A análise da trilha de feromônios visa determinar a região onde a informação contida nas matrizes de fluorescência está localizada. A PCA mostrou que a informação condensada dos dados de espectroscopia por fluorescência de cada padrão é similar, entretanto, a PCA não determina onde esta informação está localizada nos dados espectrais. Com o uso de PSCM é possível determinar se apesar dos padrões fornecerem a mesma informação a respeito da concentração de enxofre, esta informação não está localizada em regiões espectrais distintas.

A Figura 4-8 apresenta as regiões significativas indicadas pela trilha de feromônios para cada um dos quatro compostos padrões. É importante observar que não se está mostrando espectros por fluorescência 2D na Figura 4-8, mas sim o mapeamento dos feromônios. Os resultados apresentados nesta figura são obtidos utilizando cada conjunto amostral de forma integral, sem usar a segmentação em amostras para calibração e teste. A segmentação dos grupos foi descartada nesta etapa, uma vez que esta análise tem caráter meramente qualitativo, não sendo utilizada para ajuste de modelos ou predições.

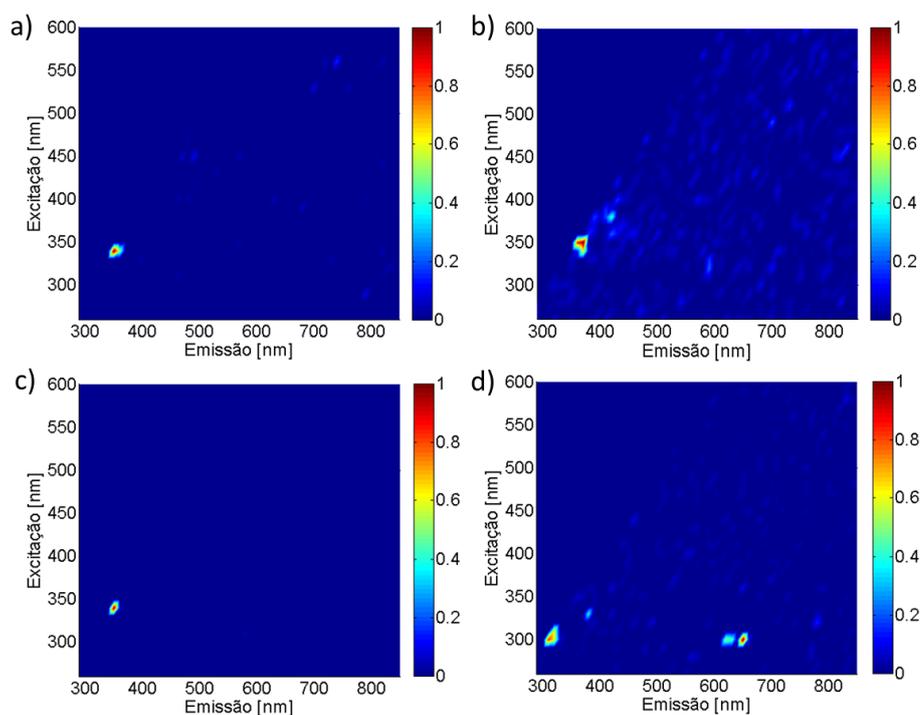


Figura 4-8: Mapeamento de regiões espectrais significativas por análise PSCM baseado nos valores da trilha de feromônios adimensional. (a) Dibenzotiofeno, (b) 4,6 – Dietildibenzotiofeno, (c) 4,6 – Dimetildibenzotiofeno e (d) Tionafteno.

Na Figura 4-8 pode ser visto que a região aos arredores do par Ex340nm/Em360nm é indicada como sendo altamente significativa para a caracterização dos elementos DBT, DEDBT e DMDBD, enquanto que as regiões próximas aos pares Ex310nm/Em320nm e Ex300nm/Em650nm são mais significativas para análise de soluções de TNTH.

Complementando os resultados obtidos com PCA, a proximidade das regiões significativas para as soluções-padrão (principalmente para DBT, DMDBT e DEDBT) são condizentes com a dificuldade em se segmentar as mesmas qualitativamente.

O TNTH apresenta um resultado levemente diferente, com uma variação da região de significância, apresentando um deslocamento da região de Ex340nm/Em360nm (para os demais elementos padrões) para uma região com maior energia associada, tanto de excitação quanto de emissão (Ex310nm/Em320nm), além de apresentar uma segunda região significativa em Ex300nm/Em650nm. Essa característica possibilita a distinção do

TNTH dos demais componentes, uma vez que esta região de emissão em 600nm não é significativa para os demais componentes.

Com base nos resultados iniciais da análise PSCM é feita a avaliação da capacidade de determinação da concentração de enxofre em amostras que contenham unicamente os elementos padrões e hexano.

Nesta etapa, é utilizada a segmentação das soluções-padrão em conjuntos:

(a) calibração e validação para seleção de pares de fluorescência e ajuste de modelos multilíneares e

(b) predição da concentração de enxofre no conjunto de testes.

As Tabelas 4-3 a 4-6 apresentam os resultados obtidos de modelos compostos entre 1 e 8 pares de fluorescência nas etapas de calibração e predição da concentração de enxofre. Nestas tabelas R²c e R²p são os coeficientes de determinação da calibração e da predição, respectivamente.

Tabela 4-3: Seleção de pares por PSCM e métricas dos modelos para DBT.

Número de pares	Pares Selecionados	Dibenzotiofeno			
		RMSEC	R ² c	RMSEP	R ² p
1	Ex340/Em360	23,5148	0,9795	23,4434	0,9735
2	Ex400/Em460 Ex340/Em350	5,5789	0,9988	13,2857	0,9465
3	Ex450/Em720 Ex340/Em350 Ex340/Em360	4,6320	0,9992	7,8636	0,9789
4	Ex260/Em610 Ex340/Em350 Ex530/Em700 Ex580/Em620	3,3603	0,9996	8,1941	0,9805
5	Ex460/Em600 Ex300/Em540 Ex360/Em450 Ex340/Em360 Ex340/Em350	2,5592	0,9998	11,8715	0,9584
6	Ex420/Em710 Ex350/Em820 Ex340/Em360 Ex350/Em740 Ex500/Em690 Ex550/Em710	3,2674	0,9996	14,3040	0,9363
7	Ex270/Em290 Ex290/Em790 Ex360/Em470 Ex500/Em850 Ex580/Em610 Ex300/Em580 Ex340/Em350	2,3494	0,9998	10,8862	0,9544
8	Ex590/Em680 Ex340/Em350 Ex280/Em360 Ex310/Em460 Ex470/Em600 Ex380/Em480 Ex460/Em810 Ex500/Em820	1,8791	0,9999	12,9395	0,9374

Tabela 4-4: Seleção de pares por PSCM e métricas dos modelos para DEDBT.

Número de pares	Pares Selecionados	4,6 - Dietildibenzotiofeno			
		RMSEC	R ² c	RMSEP	R ² p
1	Ex340/Em360	8,6831	0,9836	9,1982	0,9704
2	Ex340/Em350 Ex340/Em360	6,6235	0,9905	6,9674	0,9855
3	Ex370/Em760 Ex340/Em360 Ex340/Em700	4,3376	0,9959	7,2713	0,9784
4	Ex290/Em840 Ex340/Em550 Ex340/Em350 Ex340/Em700	3,0847	0,9979	7,5557	0,9789
5	Ex310/Em830 Ex340/Em350 Ex340/Em700 Ex340/Em360 Ex500/Em550	2,6427	0,9985	7,5041	0,9852
6	Ex400/Em410 Ex340/Em710 Ex340/Em360 Ex320/Em370 Ex290/Em790 Ex530/Em710	3,2438	0,9977	10,1218	0,9628
7	Ex570/Em810 Ex340/Em350 Ex340/Em700 Ex340/Em720 Ex340/Em360 Ex390/Em600 Ex440/Em550	2,4611	0,9987	7,1188	0,9808
8	Ex430/Em580 Ex260/Em300 Ex460/Em780 Ex340/Em360 Ex580/Em690 Ex340/Em350 Ex530/Em820 Ex490/Em800	1,7921	0,9993	8,9294	0,9702

Tabela 4-5: Seleção de pares por PSCM e métricas dos modelos para DMDBT.

Número de pares	Pares Selecionados	4,6 - Dimetildibenzotiofeno			
		RMSEC	R ² c	RMSEP	R ² p
1	Ex340/Em350	6,6251	0,9933	13,0195	0,9371
2	Ex380/Em460 Ex340/Em350	4,6298	0,9967	12,1966	0,9448
3	Ex390/Em470 Ex490/Em520 Ex340/Em350	3,9873	0,9976	13,5651	0,9404
4	Ex350/Em780 Ex340/Em350 Ex600/Em650 Ex550/Em800	3,1644	0,9985	16,7332	0,9078
5	Ex340/Em430 Ex340/Em360 Ex330/Em340 Ex340/Em350 Ex290/Em480	3,3109	0,9983	14,9268	0,9222
6	Ex330/Em450 Ex260/Em530 Ex510/Em590 Ex340/Em350 Ex290/Em580 Ex440/Em820	3,2205	0,9984	16,5320	0,9052
7	Ex500/Em610 Ex420/Em470 Ex330/Em830 Ex340/Em350 Ex340/Em450 Ex340/Em360 Ex510/Em590	2,7862	0,9988	13,7429	0,9255
8	Ex350/Em570 Ex270/Em830 Ex340/Em350 Ex350/Em430 Ex590/Em640 Ex260/Em370 Ex470/Em530 Ex410/Em440	1,9828	0,9994	11,1372	0,9585

Tabela 4-6: Seleção de pares por PSCM e métricas dos modelos para TNTH.

Número de pares	Pares Selecionados	Tionafteno			
		RMSEC	R ² c	RMSEP	R ² p
1	Ex320/Em380	82,9200	0,7639	60,9684	0,9163
2	Ex310/Em320 Ex300/Em310	27,0249	0,9749	41,9170	0,3614
3	Ex590/Em730 Ex310/Em320 Ex300/Em620	18,2187	0,9886	38,9477	0,5557
4	Ex590/Em710 Ex310/Em320 Ex310/Em760 Ex300/Em650	13,2030	0,9940	28,6598	0,7763
5	Ex300/Em580 Ex330/Em380 Ex300/Em320 Ex300/Em660 Ex410/Em820	14,2043	0,9931	36,2320	0,7058
6	Ex560/Em800 Ex300/Em630 Ex480/Em720 Ex270/Em460 Ex410/Em560 Ex320/Em710	9,0363	0,9972	62,8836	0,0490
7	Ex430/Em460 Ex440/Em790 Ex350/Em540 Ex330/Em460 Ex300/Em630 Ex410/Em420 Ex370/Em660	8,5644	0,9975	47,3630	0,6239
8	Ex390/Em720 Ex260/Em350 Ex430/Em750 Ex460/Em570 Ex310/Em320 Ex480/Em840 Ex380/Em700 Ex580/Em700	8,2605	0,9977	100,7889	0,1344

Baseado nos resultados apresentados pelas métricas, os modelos construídos com uso da metodologia PSCM foram capazes de prever satisfatoriamente concentração de enxofre, com uso de espectros por fluorescência, das soluções-padrão. Os compostos DBT, DEDBT e DMDBT apresentaram melhores resultados para predição, e sendo os mesmos compostos sulfurados encontrados em maior concentração no óleo diesel S10, a extrapolação da metodologia para busca de enxofre diretamente sobre amostras de óleo diesel é promissora.

A Figura 4-9 apresenta os pares selecionados para cada um dos oito modelos finais resultantes do processo de seleção de variáveis localizados nas respectivas posições no espectro de fluorescência 2D.

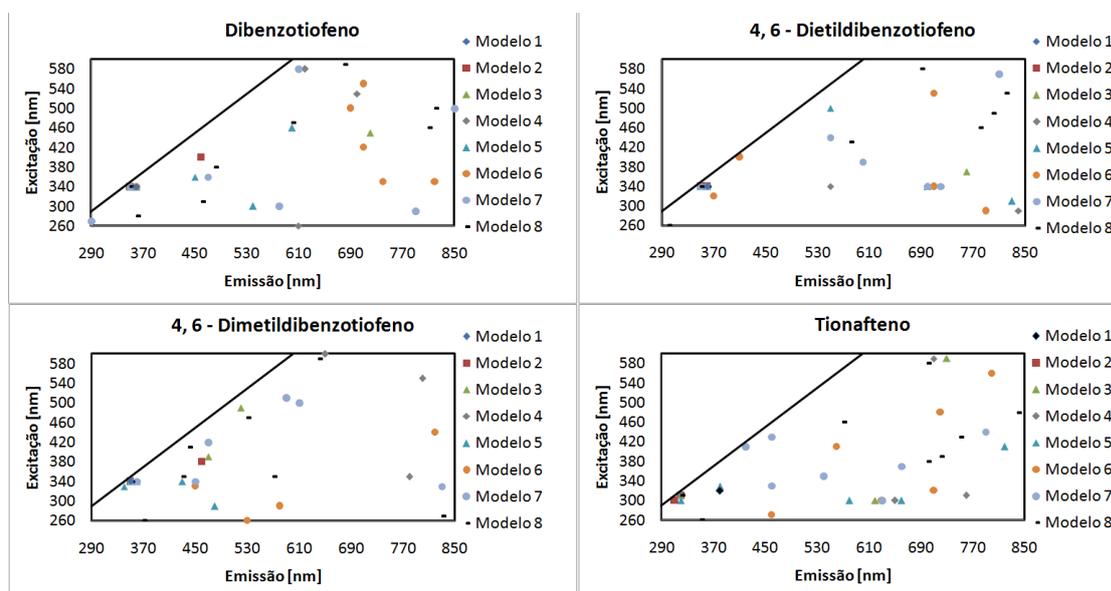


Figura 4-9: Pares de fluorescência selecionados pelos modelos localizados nas regiões do espectro de fluorescência 2D.

A sobreposição (ou proximidade) de pontos indica que diversos modelos selecionaram os mesmos pares para predição de enxofre. Assim, algumas regiões mais significativas podem ser avaliadas como sendo as principais áreas de fluorescência que possuem correlação com a variável de interesse. Comparando a seleção de variáveis com os resultados apresentados na trilha de feromônios, na Figura 4-10, as mesmas regiões destacadas como significativas podem ser visualizadas.

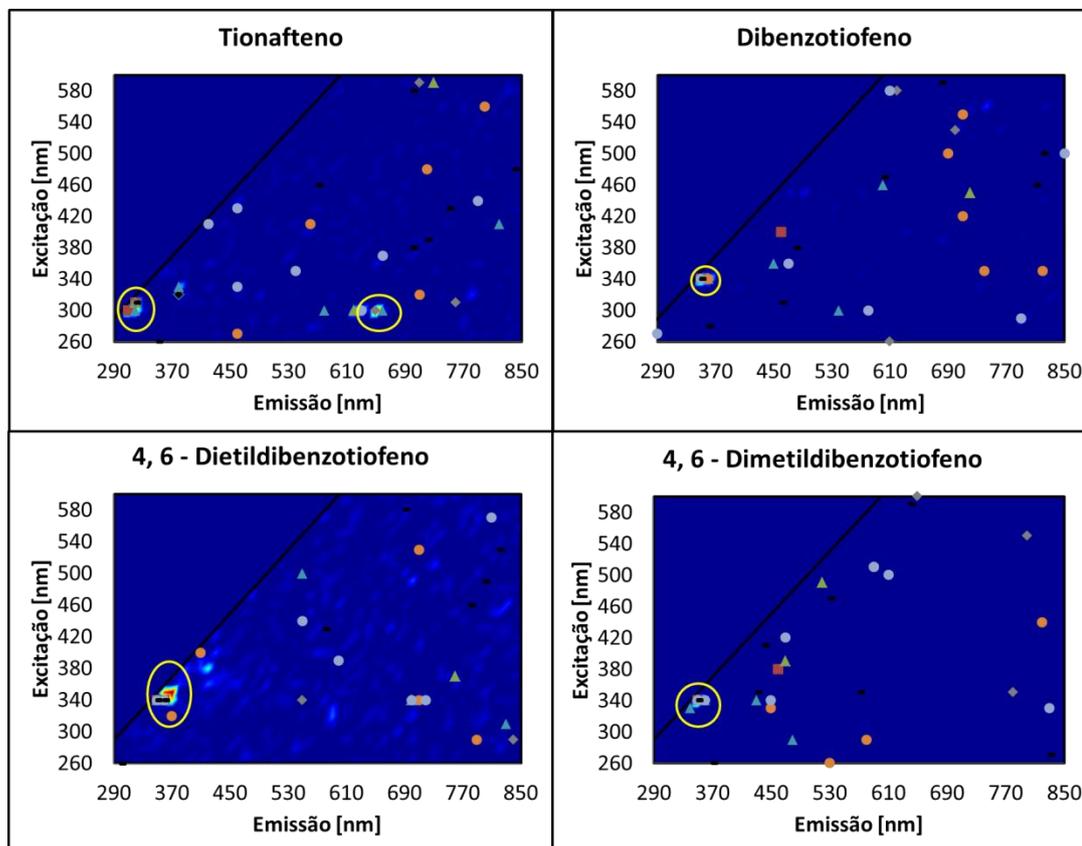


Figura 4-10: Comparação entre pares espectrais selecionados pelos modelos lineares e a trilha de feromônios para cada solução-padrão.

Capítulo 5 – Estudo de amostras de Diesel

Neste capítulo se irão analisar amostras reais de diesel para verificar se a capacidade de detectar compostos com enxofre demonstrada no capítulo anterior podem também ser aplicadas na determinação de enxofre no óleo diesel.

5.1 Materiais

Dois diferentes tipos de diesel foram avaliados neste trabalho. O primeiro deles (doravante denominado Diesel HDT) é constituído de amostras provenientes diretamente da corrente de saída de hidrotreatamento da Refinaria Alberto Pasqualini (REFAP), e dos tanques de armazenagem do mesmo. O segundo tipo de diesel analisado foram amostras de diesel S10, também cedidas pela REFAP, porém, não produzidos no local, sendo a referida refinaria apenas a compradora e certificadora do combustível. Neste capítulo é apresentada a quantidade de enxofre referente a cada amostra, a segmentação dos grupos de calibração e teste, o procedimento realizado na medição dos espectros por fluorescência e as análises realizadas com estes dados.

5.1.1 Diesel HDT

Foram analisadas 51 amostras de diesel coletadas diretamente nas correntes de saída do hidrotreatamento e dos tanques de armazenagem do mesmo. As amostras utilizadas foram coletadas por pessoal especializado, e seguiram o mesmo processo de análise certificadora utilizada como rotina dentro da refinaria.

A coleta das amostras aconteceu uma vez ao dia em um período espaçado de três meses. Assim, as amostras de diesel HDT são oriundas de diferentes cargas de petróleo, o que resulta em amostras com prováveis diferenças em constituição final.

O teste utilizado para certificação de concentração de enxofre das amostras foi o ASTM D-7039, teste padrão regulamentado pela ANP para quantificar concentração de

enxofre para diesel S500 e S10. O teste foi realizado dentro da refinaria no equipamento Sindie®7039 da empresa XOS.

Para a segmentação em grupos de calibração e teste, a metodologia *Y-rank* foi utilizada. A segmentação das amostras pode ser vista pela Tabela 5-1, que apresenta o nome da amostra e sua concentração de enxofre. Amostras destacadas com fundo azul são as selecionadas para o grupo de teste.

Tabela 5-1: Concentração de enxofre nas amostras de Diesel provenientes do processo de hidrotreatamento (HDT). Amostras em fundo azul fazem parte do grupo de testes.

Nome	S (ppm)	Nome	S (ppm)	Nome	S (ppm)
HDT1	73,7	HDT18	97,8	HDT35	104,24
HDT2	73,8	HDT19	97,9	HDT36	104,3
HDT3	75,5	HDT20	98	HDT37	104,4
HDT4	80,2	HDT21	98,2	HDT38	105
HDT5	81,9	HDT22	98,5	HDT39	107,9
HDT6	87,9	HDT23	99,3	HDT40	108,4
HDT7	88,6	HDT24	99,5	HDT41	109
HDT8	89,1	HDT25	99,8	HDT42	109
HDT9	89,5	HDT26	100,8	HDT43	113
HDT10	91,2	HDT27	101	HDT44	113,3
HDT11	92,7	HDT28	101,6	HDT45	114
HDT12	92,8	HDT29	102	HDT46	114
HDT13	93,6	HDT30	102	HDT47	114
HDT14	94	HDT31	102,9	HDT48	115
HDT15	95,5	HDT32	103	HDT49	117
HDT16	96,7	HDT33	104	HDT50	118
HDT17	97,8	HDT34	104	HDT51	138,6

A coleta dos espectros por fluorescência 2D seguiu uma metodologia diferente daquela apresentada no Capítulo 4.

Inicialmente, todas as amostras foram ambientadas a 25°C com uso de banho termostático, visando uniformizar a influência da temperatura sobre os espectros por fluorescência.

Os espectros de fluorescência foram medidos no mesmo espectrofluorômetro HORIBA Fluoromax®-4, apresentado na Figura 4-3 (a). Porém, a coleta foi realizada com o uso de fibra ótica, onde as amostras foram acondicionadas em frascos de vidro e inseridos em uma câmara escura com entrada para a fibra ótica, ilustrada na Figura 5-1.

O uso desta metodologia com fibra ótica e câmara escura foi proposta no trabalho de conclusão de curso de Alves (2012), concluindo que o uso de vidro ao invés de quartzo resulta em uma redução na intensidade de fluorescência e um leve deslocamento da posição dos picos fluorescência.

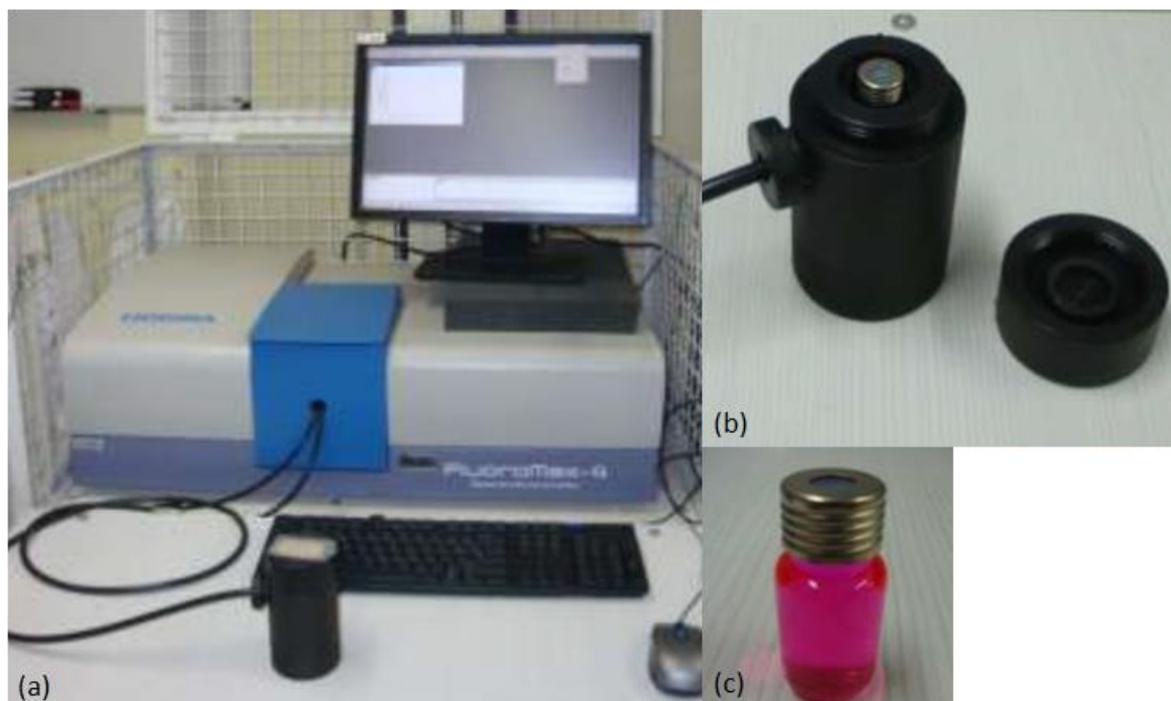


Figura 5-1: Equipamentos utilizados para coleta de espectros por fluorescência do diesel HDT: (a) o equipamento HORIBA Fluoromax®-4 com o módulo para fibra ótica ; (b) a câmara escura e (c) o tipo de frasco utilizado.

As amostras foram excitadas com comprimentos de onda variando entre 260 nm e 600 nm, com incremento de 10 nm. A emissão por sua vez foi avaliada entre os comprimentos de onda de 290 nm a 850 nm, com o mesmo incremento. Todas as análises foram realizadas em triplicata, aumentando a confiabilidade dos dados.

Assim, cada espectro de fluorescência conta com 1904 pares de fluorescência correspondentes à intensidade de fluorescência atribuída ao par Excitação/Emissão, distribuídos em uma matriz tridimensional onde os eixos x, y e z são respectivamente: comprimento de onda de emissão (λ_{em}), comprimento de onda de excitação (λ_{ex}), e a intensidade de fluorescência.

O uso desta metodologia de coleta de espectros com uso de fibra ótica e frasco de vidro foi utilizado pelo fato das amostras de diesel HDT serem coletadas na refinaria diretamente nestes frascos. Assim, o uso dos mesmos evita um maior contato do combustível com a atmosfera e com possíveis contaminantes. Além deste fato, para a construção futura de um possível sensor baseado em espectroscopia por fluorescência, é importante avaliar a capacidade da técnica em ser efetiva em diferentes formas de aquisição dos espectros por fluorescência, abrangendo uma maior gama de estratégias.

A Figura 5-2 apresenta um espectro por fluorescência típico para uma amostra de diesel HDT.

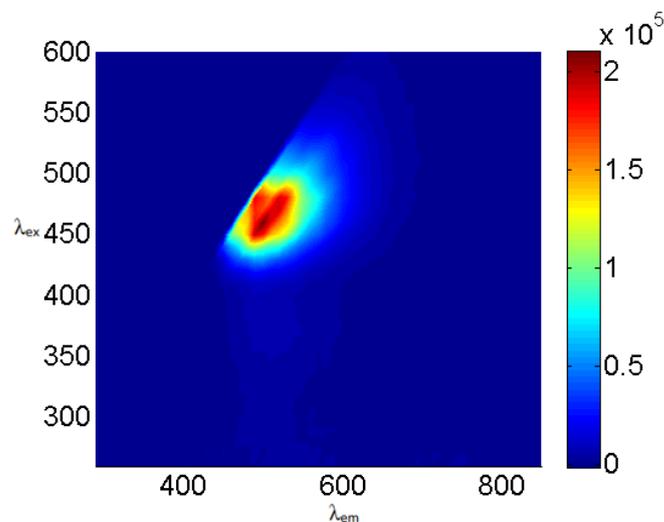


Figura 5-2: Espectro de fluorescência típico de uma amostra de Diesel HDT.

5.1.2 Diesel S10

No total, 11 amostras de diesel S10 tiveram seus espectros de fluorescência coletados para este estudo. As amostras foram divididas em dois grupos: (a) calibração e (b) teste, utilizando a metodologia *Y-rank*. A Tabela 5-2 apresenta a concentração de enxofre referente a cada amostra de diesel S10 e a segmentação dos grupos de teste e calibração (amostras com fundo azul foram separadas para o grupo de teste)

Tabela 5-2: Amostras de diesel S10. Amostras com fundo azul foram selecionadas para grupo de teste.

Amostra	S (ppm)
DS10-1	5,1
DS10-2	5,5
DS10-3	5,5
DS10-4	5,8
DS10-5	5,8
DS10-6	5,9
DS10-7	5,9
DS10-8	6
DS10-9	6
DS10-10	6,2
DS10-11	6,4

O teste utilizado para certificar a concentração de enxofre nas amostras de diesel S10 foi o ASTM D-7039, regulamentado pela ANP para diesel S10. O teste foi realizado dentro da refinaria no equipamento Sindie[®]7039 da empresa XOS.

A metodologia de coleta dos espectros seguiu os mesmos passos daqueles descritos no Capítulo 4.2, utilizando o equipamento HORIBA Fluoromax[®]-4, com uso de cubeta de quartzo como recipiente.

As amostras foram excitadas com comprimento de onda variando entre 260 nm e 600 nm, com incremento de 10 nm. A emissão foi avaliada entre 290 nm e 850 nm, com o mesmo incremento.

A Figura 5-3 apresenta o espectro para a amostra DS10-10, que pode ser vista como modelo de espectro padrão para as amostras de Diesel S10.

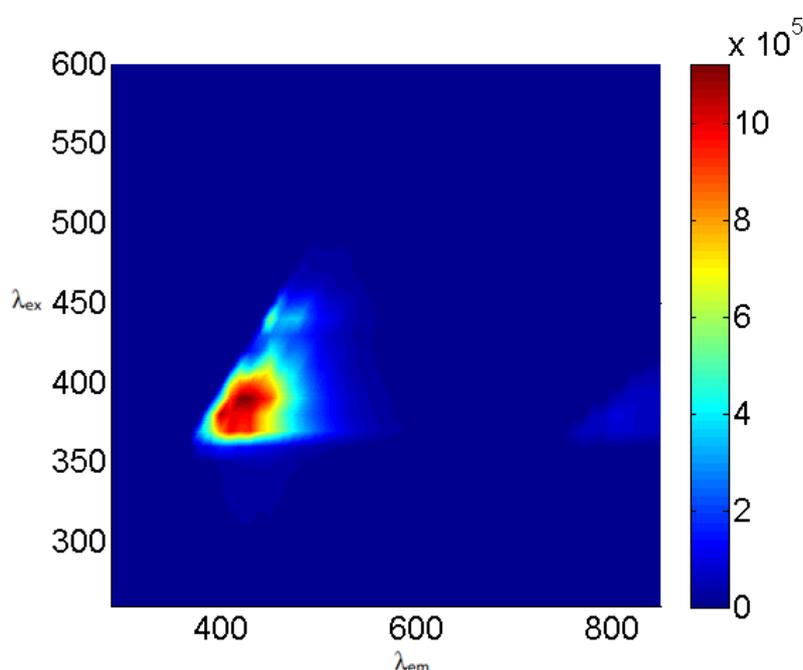


Figura 5-3: Espectro de fluorescência da amostra DS10-10.

5.2 Resultados e Discussões

Os resultados aqui apresentados serão segmentados entre: (i) Metodologias Não Assistidas - com uso de PCA e (ii) Metodologias Assistidas - por meio de PSCM.

5.2.1 Metodologia Não Assistida - PCA

Assim como apresentado no Capítulo 4, acerca de soluções-padrão, o primeiro passo da análise envolvendo o conjunto de dados de amostras de Diesel parte da aplicação da PCA, realizada individualmente sobre os conjuntos de dados espectrais de diesel S10 e diesel HDT.

A Figura 5-4 apresenta o gráfico de variância acumulada fornecida pela PCA de cada classe de amostras de diesel avaliados individualmente. Da mesma forma, a Tabela 5-3 apresenta o número de PCs necessários para expressar mais de 95% da variância acumulada para cada um dos subgrupos.

Tabela 5-3: Número de PCs necessários para descrever 95% da variância acumulada dos dados originais.

	Diesel S10	Diesel HDT
PCs	4	20

Os resultados apresentados pela análise independente de PCA, realizada em cada um dos conjuntos amostrais independentemente, mostra que a informação qualitativa contida na matriz de dados de espectroscopia por fluorescência 2D, é mais facilmente condensável no conjunto de dados oriundos das amostras de Diesel S10, do que naqueles oriundos das amostras de Diesel HDT. Isso pode ser concluído visto que a variância acumulada associada aos primeiros quatro componentes principais do conjunto de Diesel S10 representam mais de 95% da variância contida nos seus dados espectrais, enquanto no caso do Diesel HDT, mais de vinte componentes principais são necessários para representar a mesma quantidade de variância acumulada dos dados. Este resultado fornece indícios de que a quantidade de informação contida em cada um dos conjuntos de dados é significativamente diferente, ou encontra-se distribuída de forma consideravelmente distinta.

A alta remoção de compostos alifáticos emissores de fluorescência que acompanham os processos de hidrotreamento profundo podem ser responsáveis por fazer com que os espectros de fluorescência das amostras de diesel S10 possuam informação qualitativa mais facilmente condensável na forma de PCs, do que aquela apresentada pelo conjunto de dados de Diesel HDT.

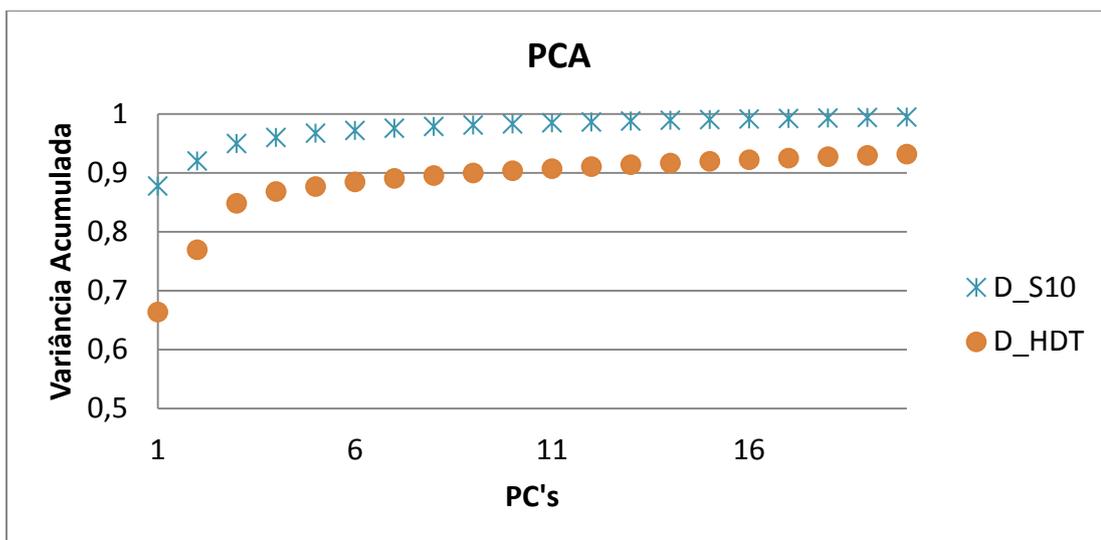


Figura 5-4: Gráfico de Variância Acumulada para análises individuais de diesel S10 (D_S10) e diesel HDT (D_HDT).

A análise de variância acumulada fornece informações qualitativas iniciais a respeito dos dados de espectroscopia de cada conjunto amostras de diesel, entretanto, uma análise mais profunda pode ser feita através dos gráficos de escores para cada conjunto amostral. A Figura 5-5 apresenta a dispersão das amostras de Diesel S10 e HDT baseada nos valores de PC1 e PC2 atribuídos para cada amostra calculadas pelas respectivas PCA. O resultado de ambas análises são apresentados de forma conjunta, para que se possa fazer uma melhor comparação visual da dispersão apresentada pelos conjuntos amostrais no plano cartesiano PC1 x PC2.

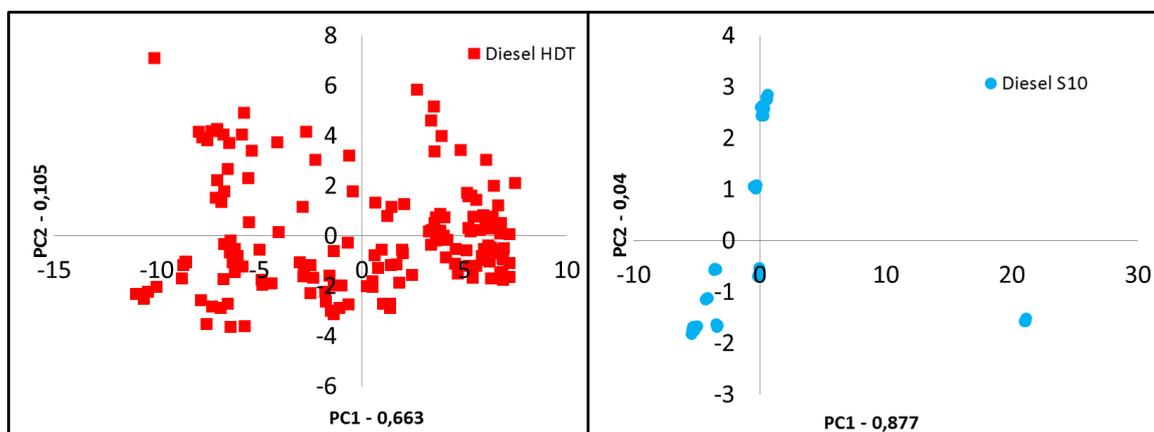


Figura 5-5: Gráfico de Escores do conjunto amostras Diesel HDT e Diesel S10, avaliados de forma individual.

As dispersões apresentadas pelos conjuntos amostrais na Figura 5-5 mostram a clara diferença entre a variabilidade da composição das amostras em cada conjunto. O grupo de amostra de Diesel HDT apresenta uma variação significativa dos pontos em torno da origem do plano cartesiano PC1 x PC2, possivelmente devido a grande diferença na característica das cargas utilizadas na produção desta corrente. Uma vez que a espectroscopia por fluorescência possui sensibilidade elevada, pequenas variações na

composição das amostras são responsáveis por variações significativas dos dados de espectroscopia.

De forma oposta, os resultados de escores apresentados pelas amostras de Diesel S10 assumem distribuição concentrada, mostrando a similaridade apresentada pelas amostras deste grupo. A alta remoção de compostos promovida pelo processo de hidrotreatamento faz com que poucas sejam as moléculas remanescentes capazes de emitir fluorescência, fazendo com que as composições sejam menos influenciadas pelas cargas utilizadas na produção.

A amostra DS10-11 apresentou comportamento diferente das demais amostras de diesel S10, se afastando do agrupamento principal. Para verificar se esta diferença constatada pela PCA pode ser vista no espectro por fluorescência da mesma, a Figura 5-6 traz um comparativo entre o espectro da referida amostra contra outra que aparece dentro do agrupamento (DS10-10).

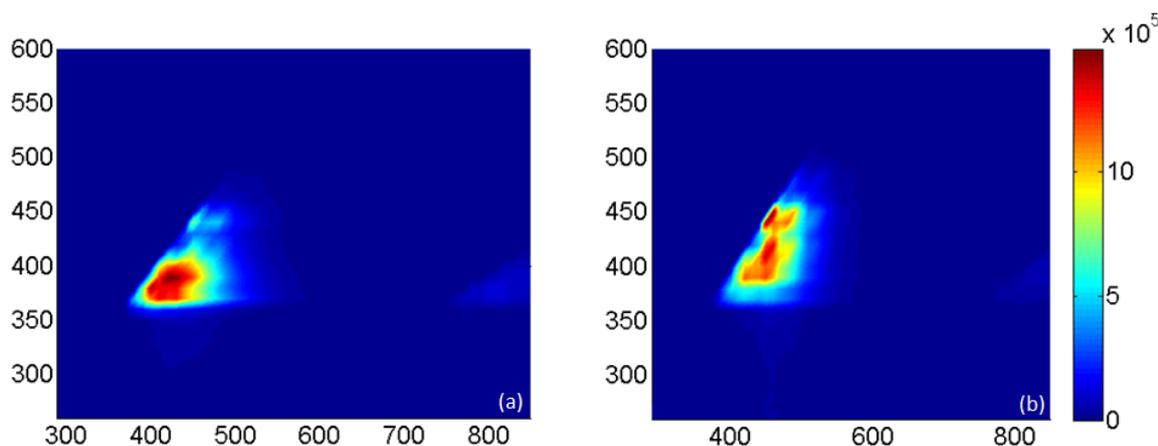


Figura 5-6: Comparativo entre exemplo de espectro por fluorescência de diesel S10 (a) e Amostra DS10-11 (b).

Pela Figura 5-6 é possível constatar visualmente a diferença entre os dois espectros por fluorescência, que levaram a amostra DS10-11 a ficar fora do agrupamento de amostras de diesel S10. Por apresentar informação distribuída de forma diferente das demais amostras de diesel S10, a mesma não será considerada para as demais PCAs realizadas.

Com o objetivo de avaliar de forma qualitativa as características equivalentes entre os conjuntos amostrais, foi aplicada a PCA no conjunto de dados completo, Diesel S10 e HDT, de forma conjunta. Para tanto, ambos os conjuntos de dados de espectroscopia foram adicionados à matriz de dados de entrada da PCA, e os resultados são avaliados na forma dos gráficos de escores. O resultado da PCA dos dados totais é apresentado na Figura 5-7.

Fica claro na Figura 5-7 a formação de dois agrupamentos bem segmentados entre as amostras de diesel HDT e as amostras de diesel S10. Isto indicava que a PCA

encontrou entre os grupos informações suficientemente díspares nos espectros por fluorescência capazes de segmentar e diferenciar os dois grupos de amostras.

Para confirmar que a segmentação decorrente da PCA entre as amostras de diesel HDT e S10 não foi devida apenas à diferença entre a metodologia aplicada à medição dos espectros por fluorescência (uma vez que as medições de diesel HDT foram realizadas com uso de fibra ótica e frascos de vidro, e as de diesel S10 com uso de cubeta de quartzo), quatro amostras de diesel HDT tiveram seus espectros por fluorescência coletados com o uso de cubeta de quartzo, seguindo os mesmos parâmetros aplicados as medidas de diesel S10. Novamente uma PCA foi aplicada a todo o conjunto de dados, gerando um novo gráfico de escores, representado na Figura 5-8. Nesta análise, primeiramente, a PCA foi calibrada utilizando os dois conjuntos de calibração (Diesel S10 e Diesel HDT, apresentado nas seções anteriores) e testada nos conjuntos de teste do Diesel S10, do Diesel HDT e nas novas medidas de Diesel HDT com cubeta. Na Figura 5-8 os pontos em azul são referentes aos dados de calibração, e em preto para os dados de teste, referentes às amostras de diesel. Os pontos S-117, S-103, S-92.7e S-73.8 são referentes aos novos espectros por fluorescência de amostras de diesel HDT com uso de cubeta, com sua concentração de enxofre indicada.

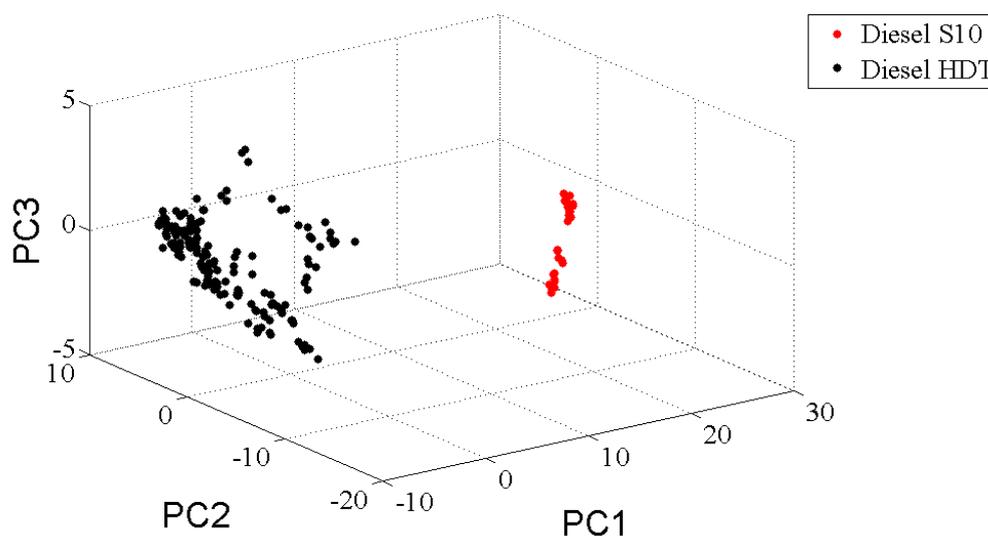


Figura 5-7: Gráfico de Escores do conjunto de amostras Diesel HDT e Diesel S10, avaliados de forma conjunta.

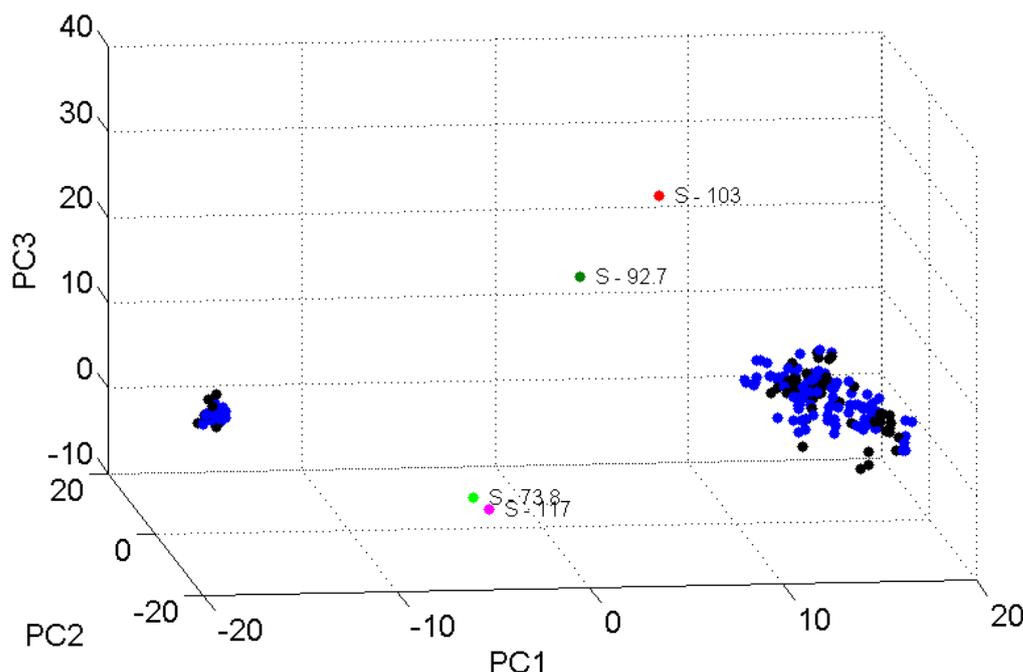


Figura 5-8: Gráfico de Escores do conjunto de amostras de Diesel HDT, Diesel S10 e novas medições de Diesel HDT com cubeta.

De acordo com a Figura 5-8, as novas medições de diesel HDT com cubeta apresentaram diferenças significativas quando comparadas com os espectros por fluorescência medidos com uso de fibra ótica e frasco de vidro. Assim, a PCA foi capaz de segmentar em dois agrupamentos distintos as mesmas amostras de diesel HDT, baseado na metodologia utilizada para medição dos espectros. Ainda assim, as informações contidas nos dados espectrais é suficiente para distinguir os dois conjuntos de diesel HDT e diesel S10, ambos com o uso da mesma metodologia para aquisição dos espectros por fluorescência.

Este resultado comprova que o uso de espectroscopia por fluorescência é capaz de diferenciar e segmentar os dois conjuntos de diesel estudados, e que mesmo a grande sensibilidade da técnica de espectroscopia à metodologia utilizada para medição dos espectros, quando comparados dados espectrométricos adquiridos da mesma forma, é factível a diferenciação entre as amostras de diesel que sofreram processos profundos de hidrotreatamento (S10) daqueles que sofreram processos mais brandos (diesel HDT).

Baseado nestes resultados é de se esperar que a metodologia de análise apresentada neste capítulo seja ainda melhor para segmentação de duas correntes ainda mais diferentes, como no caso entre diesel S10 e diesel S500, vendidos comercialmente, sendo assim, de grande interesse econômico.

5.2.2 Metodologia Assistida - PSCM

Assim como na avaliação das soluções-padrão, a metodologia de análise PSCM foi aplicada com intuito de: (i) determinar as regiões de espectroscopia mais significativas para predição de concentração de enxofre e (ii) o ajuste de modelos lineares baseados em pares de fluorescência para predição da concentração de enxofre em amostras de diesel S10 e diesel HDT.

A análise da trilha de feromônios é utilizada para determinar a região onde a informação contida nas matrizes de fluorescência está localizada. A PCA mostrou que os dados espectrais das amostras de diesel S10 podem ser condensados a um número significativamente menor do que as informações contidas nos espectros do diesel HDT, mas não determina onde esta informação está localizada. Com o uso de PSCM é possível determinar quais os pares e regiões dos espectros estão mais correlacionadas com a variável de interesse, neste caso, a concentração de enxofre.

A Figura 5-9 apresenta as regiões mais significativas indicadas pela trilha de feromônios para os dois grupos de diesel. Os resultados apresentados nesta figura fazem uso de todo o conjunto de dados espectrais, sem usar a segmentação em grupos de calibração e teste, dado o caráter qualitativo desta etapa, não sendo utilizada para ajuste de modelos ou predições.

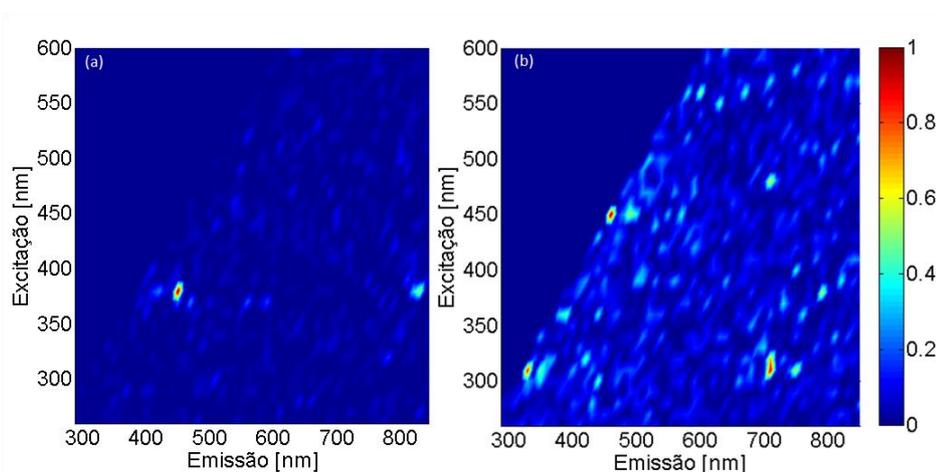


Figura 5-9: Mapeamento de regiões espectrais significativas por análise PSCM baseado nos valores da trilha de feromônios adimensional. (a) Diesel S10 e (b) Diesel HDT.

Na Figura 5-9 (a) pode ser visto que as regiões aos arredores dos pares Ex360/Em450 e Ex360/Em840 são altamente significativas para predição de enxofre em amostras de diesel S10. O diesel HDT, por sua vez, possui diversas regiões significativas, tendo sua informação altamente dispersa.

Complementando os resultados obtidos por PCA, a análise da trilha de feromônios é condizente com o esperado, com diesel S10 apresentando regiões significativas mais concentradas e com diesel HDT contando com varias regiões significativas e alta dispersão de informação dentro dos espectros.

Comparando a trilha de feromônios para diesel S10 com as trilhas apresentadas pelos padrões de componentes sulfurados presentes em óleo diesel (Figura 4-8), é possível perceber que as regiões significativas, apesar de próximas, não são as mesmas. Isto se deve ao fato da alta sensibilidade que acompanha as medidas de espectroscopia por fluorescência, onde o meio amostral e os constituintes causam diversas interferências nas emissões.

A partir dos resultados iniciais da análise PSCM é feita a avaliação da capacidade de determinação da concentração de enxofre nas amostras de diesel. Nesta etapa ambos os grupos de diesel são segmentados em: (a) conjunto de calibração e validação para seleção de pares de fluorescência e ajuste de modelos multilineares e (b) conjunto de teste para predição da concentração de enxofre.

As Tabela 5-4 e Tabela 5-5 apresentam os resultados obtidos para modelos compostos entre 1 e 8 pares de fluorescência nas etapas de calibração e predição da concentração de enxofre. Nestas tabelas R^2_c e R^2_p são os coeficientes de determinação de calibração e da predição, respectivamente.

Tabela 5-4: Pares de fluorescência selecionados por PSCM e valores das métricas de avaliação dos modelos para Diesel HDT.

Número de pares	Pares Selecionados	Diesel HDT			
		RMSEC	R^2_c	RMSEP	R^2_p
1	Ex450/Em460	11,2350	0,2202	8,5479	0,3664
2	Ex410/Em470 Ex450/Em510	8,9848	0,5013	9,9454	0,3154
3	Ex570/Em820 Ex450/Em490 Ex410/Em470	8,6933	0,5331	11,0114	0,1611
4	Ex390/Em450 Ex300/Em350 Ex420/Em460 Ex450/Em480	7,9320	0,6113	8,9740	0,3320
5	Ex450/Em490 Ex330/Em640 Ex410/Em460 Ex410/Em620 Ex320/Em350	7,4113	0,6606	8,9351	0,4106
6	Ex330/Em470 Ex440/Em600 Ex430/Em470 Ex450/Em460 Ex270/Em810 Ex510/Em660	7,0490	0,6930	7,0648	0,5548
7	Ex320/Em500 Ex410/Em520 Ex340/Em550 Ex450/Em460 Ex370/Em380 Ex280/Em550 Ex310/Em350	6,8917	0,7066	9,1334	0,4604
8	Ex350/Em380 Ex430/Em490 Ex490/Em640 Ex450/Em460 Ex380/Em390 Ex470/Em500 Ex410/Em780 Ex410/Em470	6,8273	0,7120	8,4292	0,5039

Tabela 5-5: Pares de fluorescência selecionados por PSCM e valores das métricas de avaliação dos modelos para Diesel S10.

Número de pares	Pares Selecionados	Diesel S10			
		RMSEC	R^2_c	RMSEP	R^2_p
1	Ex390/Em820	0,2332	0,6183	0,3344	0,0658
2	Ex460/Em470 Ex450/Em560	0,1481	0,8461	0,1779	0,5377
3	Ex370/Em380 Ex350/Em480 Ex390/Em820	0,0901	0,9431	0,1999	0,7716
4	Ex470/Em530 Ex390/Em520 Ex440/Em510 Ex410/Em430	0,0897	0,9435	0,1598	0,6905
5	Ex290/Em420 Ex410/Em620 Ex440/Em490 Ex490/Em510 Ex410/Em430	0,0643	0,9710	0,2599	0,5019
6	Ex280/Em510 Ex450/Em560 Ex370/Em400 Ex520/Em530 Ex350/Em470 Ex410/Em840	0,0446	0,9861	0,2027	0,7056
7	Ex360/Em680 Ex560/Em590 Ex350/Em470 Ex400/Em830 Ex480/Em660 Ex550/Em720 Ex470/Em480	0,0485	0,9835	0,2051	0,3240
8	Ex550/Em730 Ex410/Em840 Ex400/Em520 Ex370/Em400 Ex530/Em540 Ex370/Em380 Ex520/Em540 Ex380/Em480	0,0385	0,9896	0,3233	0,7719

Baseado no resultado das métricas apresentadas, os modelos construídos utilizando a metodologia PSCM não foram capazes de prever satisfatoriamente concentração de enxofre nas amostras de diesel HDT. A utilização de modelos multilineares de até 8 pares não apresentou bons resultados de coeficiente de determinação para calibração ou predição. Como visto pela PCA, os espectros por fluorescência de diesel HDT trazem

uma grande gama de informação de difícil condensação, e um alto número de pares de fluorescência são necessários para representar de forma satisfatória os dados contidos nos espectros.

Para diesel S10, os resultados das métricas dos modelos foram muito melhores. Os coeficientes de determinação de calibração foram superiores a 0,9 com o uso de apenas 3 pares. A pequena quantidade de amostras e a faixa estreita de concentrações dificulta o ajuste de modelos, porém, a metodologia empregada se mostrou satisfatória para obtenção dos mesmos, e se faz útil para um estudo mais aprofundado.

Uma vez que a metodologia aqui apresentada para predição de enxofre em óleo diesel não visa ser utilizada como teste certificador para o mesmo, mas sim, servir como medição auxiliar, com capacidade de ser utilizada em sensores online, instalados diretamente sobre correntes de diesel, para propiciar um acompanhamento em tempo real da concentração de enxofre, possibilitar um controle fino dos processos de hidrotreatamento e *blending*, os resultados encontrados com uso de PSCM para ajuste de modelos baseados em espectros por fluorescência para diesel S10 são promissores para a construção de protótipos, uma vez que confirmam a viabilidade da do uso de espectroscopia por fluorescência para predição de enxofre.

A Figura 5-10 apresenta os pares selecionados para cada um dos oito modelos finais resultantes do processo de seleção de variáveis localizados nas respectivas posições no espectro por fluorescência 2D.

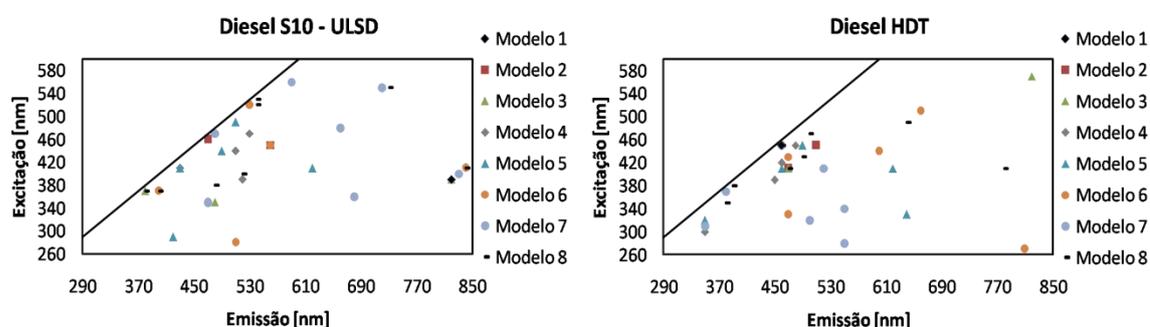


Figura 5-10: Pares de fluorescência selecionados pelos modelos nas respectivas posições no espectro por fluorescência para diesel HDT e diesel S10.

Como pode ser visto na Figura 5-11, que compara os pares selecionados pelos modelos lineares com a trilha de feromônios, os modelos ajustados para diesel HDT fizeram uso de muitos pares concentrados na mesma região, o que não pode ser visto pela trilha de feromônios do mesmo. Isto se deve ao fato de que dado o espalhamento da informação contida nos espectros, os modelos lineares não conseguiram de forma precisa selecionar uma região que trouxesse resultados superiores, fazendo com que diversas regiões do espectro levassem a modelos com erros similares. A região onde estão concentrados a maioria dos pares selecionados pelos melhores modelos lineares é possivelmente a que tem maior correlação linear com a variável de interesse.

Para diesel S10, as regiões significativas, onde a maioria dos modelos selecionou os mesmos pares (ou em pares próximos) segue a mesma orientação daquela vista na trilha de feromônios.

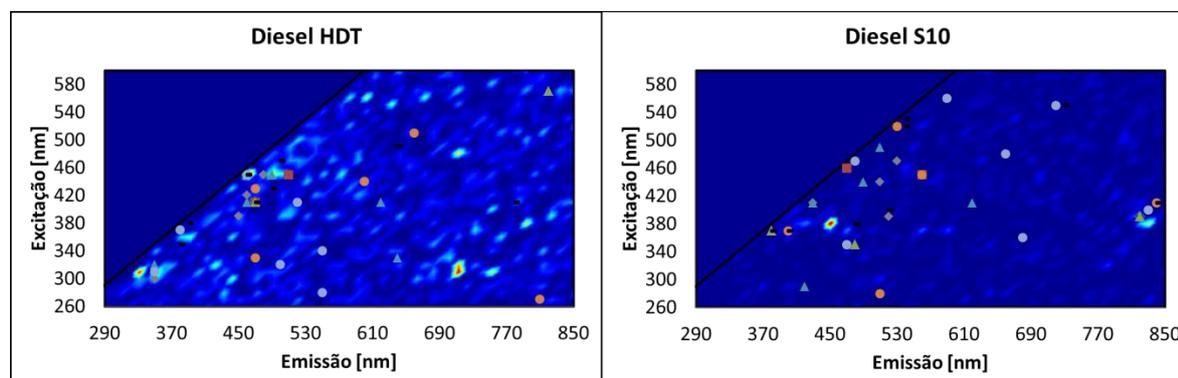


Figura 5-11: Comparação entre os pares selecionados pelos modelos lineares para diesel HDT e diesel S10 e suas respectivas trilhas de feromônios.

A Figura 5-12 apresenta um comparativo entre os pares selecionados pelos modelos lineares de um a quatro pares para predição de enxofre em cada um dos padrões da família DBT, com os pares selecionados para modelos lineares de um a quatro pares para predição de enxofre em diesel S10, para avaliar se é possível uma correlação entre as regiões mais significativas. Os pontos foram plotados sobre as regiões mais significativas para diesel S10.

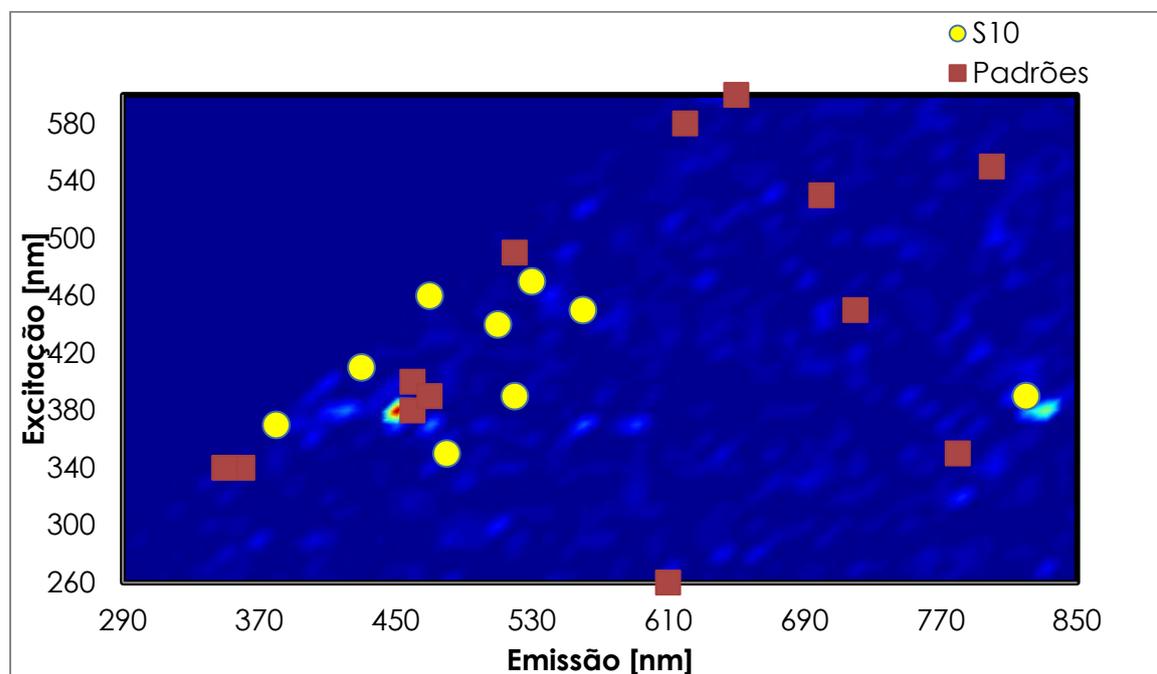


Figura 5-12: Pares de fluorescência selecionados pelos modelos lineares de um a quatro pares para busca de enxofre nos padrões da família DBT e diesel S10, plotados sobre as regiões mais significativas para diesel S10.

Pela análise da Figura 5-12, pode-se perceber que existe uma relativa proximidade entre os pares selecionados para quantificação de enxofre em diesel S10 e nos padrões da família DBT, porém, não houve seleção dos mesmos pares, e em uma análise mais minuciosa não se pode correlacionar diretamente a busca por enxofre nas soluções padrão com a busca em diesel S10.

Capítulo 6 – Conclusões e Sugestões para Trabalhos Futuros

Com uso de PCA foi possível concluir que os conjuntos originais de dados de espectroscopia para DBT, BMDBT e DEDBT apresentam informação que pode ser condensada em até três componentes principais. Assim, a informação contida nestes espectros é localizada em um pequeno grupo de pares de fluorescência. Para o TNTH, por outro lado, um maior número de PCs foi necessário para expressar mais de 95% da variância contida no conjunto original de dados, por estes conterem informação mais dispersa.

Pela análise do gráfico de escores construído para o conjunto total de padrões, não foi possível utilizar a metodologia PCA para segmentar ou diferenciar os compostos padrões estudados. Não há formação de *clusters* e as soluções-padrão apresentaram comportamento muito similar. Isto leva a crer que a informação contida nos dados espectrais das soluções-padrão está concentrada nas mesmas regiões, e assim, podem ser consideradas estatisticamente similares.

Esta característica não inviabiliza a utilização de espectroscopia por fluorescência para determinar concentrações destes elementos em amostras puras dos mesmos, mas inviabiliza a segmentação de cada um destes compostos em correntes contendo misturas destes.

Do ponto de vista prático, o fato destas moléculas apresentarem informações qualitativas equivalentes favorece o projeto e desenvolvimento de sensores baseados em medidas de fluorescência para a caracterização de amostras de diesel S10, já que os referidos sensores poderão ser ajustados para medir informações específicas de fluorescência, tornando-os mais acurados, precisos e robustos.

Quando avaliadas pela metodologia PSCM, as regiões significativas para caracterização de cada uma das soluções-padrão, indicadas pela trilha de feromônios, apresentaram grande similaridade, comprovando as conclusões obtidas a partir da PCA. A proximidade das regiões significativas inviabiliza a segmentação qualitativa das mesmas.

Baseado no resultado das métricas, os modelos multilineares ajustados por PSCM foram capazes de prever a concentração de enxofre nas soluções-padrão de maneira satisfatória, com coeficientes de determinação de calibração maiores que 0,97. A seleção de variáveis resultante da aplicação desta metodologia foi capaz ainda de indicar quais os principais pares de fluorescência correlacionados com a variável de interesse.

Comparando a maior localização dos pares selecionados pelos modelos com as regiões significativas discriminadas pela trilha de feromônios, foi possível observar que as mesmas são equivalentes para todos os padrões.

Com o uso de PCA individualmente nos grupos de diesel S10 e diesel HDT foi possível concluir que a informação contida nos espectros de S10 é mais condensável e que provavelmente esteja localizada em regiões específicas. Para HDT, mais de 20 PCs foram necessários para descrever 95% da variância dos dados originais. Uma vez que o diesel HDT é composto por diversas moléculas e compostos emissores de fluorescência, que são removidos durante processos de hidrotreamento, faz sentido que seus espectros por fluorescência contenham um maior número de variáveis independentes.

Aplicando PCA conjuntamente sobre as amostras de diesel HDT e diesel S10 foi possível, através do gráfico de escores, observar uma nítida segmentação. Assim, os espectros dos dois grupos se mostraram diferentes o suficiente para que a avaliação qualitativa da PCA consiga diferenciar os mesmos. Esta é uma conclusão importante, uma vez que a habilidade em se segmentar diesel que passou por forte hidrotreamento (S10) e outro que sofreu hidrotreamento mais brando (S500, por exemplo) pode ser útil para construção de sensores capazes de rapidamente identificar combustíveis adulterados.

A análise da trilha de feromônios associada a cada um dos grupos foi capaz de identificar as regiões de maior importância para ambos. Como esperado, diesel S10 possui duas regiões de importância bem evidentes. Para diesel HDT, muitas regiões significativas foram identificadas, apresentando uma trilha de feromônios bem dispersa.

Os modelos ajustados por meio de PSCM apresentaram resultados satisfatórios em predição de enxofre para as amostras de diesel S10. Com o uso de três pares de fluorescência foi possível ajustar modelos com coeficiente de determinação relativos aos dados de calibração superiores a 0,94. Para as amostras de diesel HDT, os modelos ajustados se mostraram menos satisfatórios, indicando talvez que modelos lineares possam não ser a melhor opção para quantificar enxofre em amostras de diesel com maior concentração de enxofre.

A comparação entre a seleção de pares para predição de enxofre em diesel S10 e nos padrões da família DBT mostrou que as regiões de importância são próximas, mas não são as mesmas. Esta conclusão é apropriada uma vez que a mistura de diversos outros compostos constituintes do diesel S10 (e não somente dos padrões estudados) afeta diretamente a fluorescência do meio.

Uma vez que a metodologia apresentada neste trabalho não visa servir de teste certificador para concentração de enxofre em óleo diesel, mas sim ser utilizada como uma medida rápida, capaz de ser utilizada para o desenvolvimento de sensores de linha, aumentando a quantidade de informação disponível nas plantas para controle de processos como hidrotreamento e *blending* de correntes, a viabilidade do uso de espectroscopia por fluorescência 2D para predição de enxofre em óleo diesel, na faixa do S10, foi evidenciada como viável, principalmente quando aplicada a baixas concentrações de enxofre, principal objetivo deste trabalho.

Por fim, cabe salientar, que não foi encontrada na literatura nenhuma aplicação ou estudo que utilizasse diretamente espectroscopia de fluorescência em amostras de diesel para quantificação de enxofre, sem a necessidade de pré-tratamento da amostra. Desta forma, este é o primeiro trabalho a ser publicado que demonstra esta possibilidade, especialmente para diesel S10 (ULSD).

Para trabalhos futuros, podem-se citar:

- Construção de um sensor para segmentação de diesel S10 e S500 com uso de PCA e espectros por fluorescência
- Ajuste de modelos não lineares durante a rotina de otimização.
- Construção de um sensor baseado em pares específicos de fluorescência para predição de enxofre em diesel S10.

Referências

ABURTO, P. et al. Quantitative analysis of sulfur in diesel by enzymatic oxidation, steady-state fluorescence, and linear regression analysis. *Energy and Fuels*, v. 28, n. 1, p. 403-408, 2014.

ALLEGRI, F.; OLIVIERI, A. C. A new and efficient variable selection algorithm based on ant colony optimization. Applications to near infrared spectroscopy/partial least-squares analysis. *Analytica Chimica Acta*, v. 699, n. 1, p. 18-25, 8/5/ 2011. ISSN 0003-2670.

ALVES, C. D. V. Uma nova sistemática para análise de enxofre em diesel baseada em fluorescência. TRABALHO DE DIPLOMAÇÃO EM ENGENHARIA QUÍMICA. Porto Alegre: Universidade Federal do Rio Grande do Sul 2012.

ANP. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. RESOLUÇÃO ANP Nº 50, DE 23.12.2013 2013.

ASTM. American Society for Testing and Materials. www.astm.org. 2014

BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. *Chemometrics: A practical guide*. New York: Wiley & Sons 1998.

BEENS, J.; TIJSSEN, R.; BLOMBERG, J. Prediction of comprehensive two-dimensional gas chromatographic separations: A theoretical and practical exercise. *Journal of Chromatography A*, v. 822, n. 2, p. 233-251, 10/2/ 1998. ISSN 0021-9673.

BERNARD, J. et al. Design and characterization of a thermochemical high-performance liquid chromatography flame photometric detector interface for the speciation of sulfur. *The Analyst*, p. 1475-1481, 1994.

BREITKREITZ, M. C. et al. Determination of total sulfur in diesel fuel employing NIR spectroscopy and multivariate calibration. *Analyst*, v. 128, n. 9, p. 1204-1207, 2003. ISSN 0003-2654.

BRERETON, R. *Chemometrics for Pattern Recognition*. Chichester: John Wiley & Sons, 2007.

BUENO, A. F. Desenvolvimento de um analisador de processo por espectroscopia no infravermelho próximo (NIR) para previsão de propriedades de derivados de petróleo. UNICAMP. Tese de Doutorado 2011.

CADIMA, J.; CERDEIRA, J. O.; MINHOTO, M. Computational aspects of algorithms for variable selection in the context of principal components. *Computational Statistics & Data Analysis*, v. 47, n. 2, p. 225-236, 9/1/ 2004. ISSN 0167-9473.

CHANDRA SRIVASTAVA, V. An evaluation of desulfurization technologies for sulfur removal from liquid fuels. *RSC Advances*, v. 2, n. 3, p. 759-783, 2012.

CHAWLA, B.; DI SANZO, F. Determination of sulfur components in light petroleum streams by high-resolution gas chromatography with chemiluminescence detection. *Journal of Chromatography*, v. 589, n. 1-2, p. 271-279, 1992.

CNT. Confederação Nacional do Transporte. Os Impactos da Má Qualidade do Óleo Diesel Brasileiro: Confederação Nacional do Transporte 2012.

CRAMER, J. A. et al. Ultra-low sulfur diesel classification with near-infrared spectroscopy and partial least squares. *Energy and Fuels*, v. 23, n. 2, p. 1132-1133, 2009.

CRONEY, J. C.; JAMESON, D. M.; LEARMONTH, R. P. Fluorescence spectroscopy in biochemistry: teaching basic principles with visual demonstrations. *Biochemistry and Molecular Biology Education*, v. 29, n. 2, p. 60-65, 2001. ISSN 1470-8175.

DAYAL, B. S.; MACGREGOR, J. F. Improved PLS algorithms. *J. Chemom.*, v. 11, p. 73 - 85, 1997.

DORIGO, M.; GAMBARELLA, L. M. Ant Colonies for the Travelling Salesman Problem. *Biosystems*, v. 43, p. 73 - 81, 1997.

FACCHIN, S. Técnicas de análise multivariável aplicadas ao desenvolvimento de analisadores virtuais. 2005. (Dissertação de Mestrado). Departamento de Engenharia Química, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.

GELADI, P. Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Spectrochimica Acta Part B: Atomic Spectroscopy*, v. 58, n. 5, p. 767-782, 5/30/2003. ISSN 0584-8547.

GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning* Boston: Kluwer Academic Publishers, 1989.

GOSCINIAK, I. A new approach to particle swarm optimization algorithm. *Expert Systems with Applications*, v. 42, n. 2, p. 844-854, 2015.

GOSS, S. et al. Self-organized Shortcuts in the Argentine Ant. *Naturwissenschaften*, v. 76, p. 579 - 581, 1989.

HEMMATEENEJAD, B. et al. Building optimal regression tree by ant colony system–genetic algorithm: Application to modeling of melting points. *Analytica Chimica Acta*, v. 704, n. 1–2, p. 57-62, 10/17/ 2011. ISSN 0003-2670.

HUA, R. et al. Determination of sulfur-containing compounds in diesel oils by comprehensive two-dimensional gas chromatography with a sulfur chemiluminescence detector. *Journal of Chromatography A*, v. 1019, n. 1-2, p. 101-109, 2003.

Roy, K. et al. Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques. *European Journal of Medicinal Chemistry*, v. 44, n. 7, p. 2913-2922, 7// 2009. ISSN 0223-5234.

KUMAR, N. et al. Chemometrics tools used in analytical chemistry: An overview. *Talanta*, n. 0, 2014. ISSN 0039-9140.

LAKOWICZ, J. R. Chapter 1 Fluorescence Spectroscopy; Principles and Application to Biological Macromolecules. In: DEENEN, A. N. A. L. L. M. V. (Ed.). *New Comprehensive Biochemistry*: Elsevier, v. Volume 11, Part B, 1985. p.1-26. ISBN 0167-7306.

Principles of fluorescence spectroscopy. 3rd. New York: Springer, 2006. xxvi, 954 p. ISBN 9780387312781.

LAND JR et al. Partial Least Squares Applied to Medical Bioinformatics. *Procedia Computer Science*, v. 6, n. 0, p. 273-278, // 2011. ISSN 1877-0509.

LINDEMANN, C. et al. 2-Dimensional fluorescence spectroscopy for on-line bioprocess monitoring. *Sensors and Actuators B: Chemical*, v. 51, n. 1–3, p. 273-277, 8/31/ 1998. ISSN 0925-4005.

MULLEN, R. J. et al. A review of ant algorithms. *Expert Systems with Applications*, v. 36, n. 6, p. 9608-9617, 8// 2009. ISSN 0957-4174.

MURAKAMI, Y. Analysis of corrosive wear of diesel engines: relationship to sulfate ion concentrations in blowby and crankcase oil. *JSAE Review*, v. 16, n. 1, p. 43-48, 1995.

OZAKI, Y. Near-infrared spectroscopy-its versatility in analytical chemistry. *Analytical Sciences*, v. 28, n. 6, p. 545-563, 2012.

PANTOJA, P. A. et al. Prediction of crude oil properties and chemical composition by means of steady-state and time-resolved fluorescence. *Energy and Fuels*, v. 25, n. 8, p. 3598-3604, 2011.

PASQUINI, C.; BUENO, A. F. Characterization of petroleum using near-infrared spectroscopy: Quantitative modeling for the true boiling point curve and specific gravity. *Fuel*, v. 86, n. 12–13, p. 1927-1934, 8// 2007. ISSN 0016-2361.

RANZAN, C. et al. Caracterizacao de Fermentacoes de glicose com *Saccharomyces cerevisiae* utilizando Espectroscopia Fluorescente 2D e Ant Colony Optimization. Foz do Iguacu, PR, Brasil: 19 Simpósio Nacional de Bioprocessos 2013.

RANZAN,C. et al. Wheat flour characterization using NIR and spectral filter based on Ant Colony Optimization. *Chemometrics and Intelligent Laboratory Systems*, v. 132, n. 0, p. 133-140, 3/15/ 2014. ISSN 0169-7439.

REDDY, T. A. *Applied Data Analysis and Modeling for Energy Engineers and Scientists*. Arizona: Springer, 2011.

RIVEROS, L. et al. Determination of asphaltene and resin content in Venezuelan crude oils by using fluorescence spectroscopy and partial least squares regression. *Energy and Fuels*, v. 20, n. 1, p. 227-230, 2006.

SETTLE, F. *Handbook of Instrumental Techniques for Analytical Chemistry*. Upper Saddle River: Prentice-Hall, 1997.

SHAMSIPUR, M. et al. An efficient variable selection method based on the use of external memory in ant colony optimization. Application to QSAR/QSPR studies. *Analytica Chimica Acta*, v. 646, n. 1-2, p. 39-46, 7/30/ 2009. ISSN 0003-2670.

SHINZAWA, H. et al. Accelerated weathering-induced degradation of poly (lactic acid) fiber studied by near-infrared (NIR) hyper spectral imaging. *Applied Spectroscopy*, v. 66, n. 4, p. 470-474, 2012.

SHLENS, J. *A Tutorial on Principal Component Analysis*. New York: New York University 2009.

SIESLER, H. W. et al. *Near-Infrared Spectroscopy*. Wiley-VCH, 2002.

SKOOG, D. A.; HOLLER, F. J.; CROUCH, S. R. *Principles of instrumental analysis*. 6th. Belmont, CA: Thomson Brooks/Cole, 2007. xv, 1039 p. ISBN 9780495012016 0495012017.

SOTOMAYOR, M. D. P. T. et al. Aplicação e avanços da espectroscopia de luminescência em análises farmacêuticas. *Química Nova*, v. 31, n. 7, p. 1755-1774, 2008.

VALENTI, B. et al. Infrared spectroscopic methods for the discrimination of cows' milk according to the feeding system, cow breed and altitude of the dairy farm. *International Dairy Journal*, v. 32, n. 1, p. 26-32, 2013.

VERGOUW, C. G. et al. No evidence that embryo selection by near-infrared spectroscopy in addition to morphology is able to improve live birth rates: Results from an individual patient data meta-analysis. *Human Reproduction*, v. 29, n. 3, p. 455-461, 2014.

WANG, F. C. Y. et al. Speciation of Sulfur-Containing Compounds in Diesel by Comprehensive Two-Dimensional Gas Chromatography. *Journal of Chromatographic Science*, v. 41, n. 10, p. 519-523, 2003.

WANG, X. et al. Analysis on fluorescence of dual excitable Eu(TTA)3DPBT in toluene solution and PMMA. *Journal of Luminescence*, v. 131, n. 8, p. 1719-1723, 8// 2011. ISSN 0022-2313.

WEHRENS, R.; SPRINGERLINK. *Chemometrics with R Multivariate Data Analysis in the Natural Sciences and Life Sciences. Use R*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg 2011.

WEST, M. et al. 2013 Atomic spectrometry update - A review of advances in X-ray fluorescence spectrometry. *Journal of Analytical Atomic Spectrometry*, v. 28, n. 10, p. 1544-1590, 2013.

WHITFORD, W.; JULIEN, C. *Analytical Technology and PAT*. *BioProcess International*, p. 32-41, 2007.

WILLIAMS, F. Review of present trends in luminescence research. *Journal of Luminescence*, v. 24–25, Part 2, n. 0, p. 929-936, 11// 1981. ISSN 0022-2313.

WILSON, J. H.; ZHANG, C.; KOVACS, J. M. Separating crop species in Northeastern Ontario using hyperspectral data. *Remote Sensing*, v. 6, n. 2, p. 925-945, 2014.

YU, D. et al. Near-infrared fluorescent probe for detection of thiophenols in water samples and living cells. *Analytical Chemistry*, v. 86, n. 17, p. 8835-8841, 2014.