



**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA**

WILLIAN FRANCISCO CORDEIRO DANTAS

**APLICAÇÃO DE MÉTODOS DE RESOLUÇÃO MULTIVARIADA DE CURVAS E
ESPECTROSCOPIA ÓPTICA EM ANÁLISES FORENSES E FOTOQUÍMICA**

**CAMPINAS
2019**

WILLIAN FRANCISCO CORDEIRO DANTAS

**APLICAÇÃO DE MÉTODOS DE RESOLUÇÃO MULTIVARIADA DE CURVAS E
ESPECTROSCOPIA ÓPTICA EM ANÁLISES FORENSES E FOTOQUÍMICA**

**Tese de Doutorado apresentada ao Instituto de
Química da Universidade Estadual de Campinas
como parte dos requisitos exigidos para a obtenção
do título de Doutor em Ciências**

Orientador: Prof. Dr. Ronei Jesus Poppi

**O arquivo digital corresponde à versão final da Tese defendida pelo aluno
Willian Francisco Cordeiro Dantas e orientada pelo Prof. Dr. Ronei Jesus Poppi.**

CAMPINAS

2019

Agência(s) de fomento e nº(s) de processo(s): CAPES

ORCID: <https://orcid.org/0000-0001-8725-0771>

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Química
Camila Barleta Fullin - CRB 8462

D235a Dantas, Willian Francisco Cordeiro, 1989-
Aplicação de métodos de resolução multivariada de curvas e espectroscopia óptica em análises forenses e fotoquímica / Willian Francisco Cordeiro Dantas. – Campinas, SP : [s.n.], 2019.

Orientador: Ronei Jesus Poppi.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Química.

1. Resolução multivariada de curvas. 2. Espectroscopia Raman. 3. Análise forense de imagem. 4. Fotoquímica. I. Poppi, Ronei Jesus, 1961-. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Application of multivariate curve resolution methods and optical spectroscopy in forensic and photochemical analysis **Palavras-chave em inglês:**

Multivariate curve resolution

Raman spectroscopy

Forensic image analysis

Photochemistry

Área de concentração: Química Analítica

Titulação: Doutor em Ciências **Banca**

examinadora:

Ronei Jesus Poppi [Orientador]

Francisco Benedito Teixeira Pessine

José Alberto Fracassi da Silva

Guilherme Post Sabin

Mariana Ramnos de Almeida

Data de defesa: 22-02-2019

Programa de Pós-Graduação: Química

BANCA EXAMINADORA

Prof. Dr. Ronei Jesus Poppi (Orientador)

Prof. Dr. Francisco Benedito Teixeira Pessine (IQ - UNICAMP)

Prof. Dr. José Alberto Fracassi da Silva (IQ - UNICAMP)

Dr. Guilherme Post Sabin (Speclab Holding S.A.)

Profa. Dra. Mariana Ramos de Almeida (Departamento de Química - UFMG)

A Ata da defesa assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Este exemplar corresponde à redação final da Tese de Doutorado defendida pelo aluno **WILLIAN FRANCISCO CORDEIRO DANTAS**, aprovada pela Comissão Julgadora em 22 de fevereiro de 2019.

"The oldest and strongest emotion of mankind is fear, and the oldest and strongest kind of fear is fear of the unknown."

H.P. Lovecraft

Supernatural Horror in Literature,
1927.

"Imagination is more important than knowledge. For knowledge is limited to all we now know and understand, while imagination embraces the entire world, and all there ever will be to know and understand."

Albert Einstein

New York Times, 08 de março de
1931.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer à minha mãe e minha irmã, por todo o apoio e empenho na minha formação acadêmica, e também a meu pai, que não está mais entre nós, mas sempre me apoiou e teve fé de que eu alcançaria meus objetivos.

Aos colegas do LAQQA durante todo o doutorado, Carlos Alberto, Carlos Diego, Carol, Débora, Felipe, Guilherme, Hery, Humberto, Javier, José Augusto, Karen, Laís, Luciana Oliveira, Luciana Terra, Marina, Victor Hugo e Victor Kelis.

Aos amigos feitos durante esses 11 anos de IQ - UNICAMP, Bruna, Bruno, Carol, Douglas, Filipe, Gabrielle, Guilherme, Helena, Jéssica, Laís Helena, Leandro, Luís Gustavo, Marcelo, Mariana, Mônica, Rafael, Ricardo, Tatiana, Thaís e Vanessa.

Aos meus amigos de longa data, Guilherme, Alexandre, Marcelo, Dante, Fabrício, Stefano, Fernanda, Tatiane, Giuliano e Artur.

Aos amigos do CrossFit Aureliz, Ana, Aurélio, Dani, Eliz, Fábio, Fernanda, Fernando, Gabi Thomé, Gabi Wakai, Gabi Zecchini, Geovana, Guilherme, Grazielle, Marcelo, Marcos, Mário, Mayara, Nalva, Rafael, Renata e Thiago, por todos os momentos de treino e descontração dentro e fora do box.

À Professora Teresa Dib Zambon Atvars pela colaboração.

Ao Professor Fabiano Severo Rodembusch do Instituto de Química da Universidade Federal do Rio Grande do Sul pela colaboração.

Ao meu orientador Ronei Jesus Poppi pela oportunidade, amizade, conversas científicas e ensinamentos que foram fundamentais para essa tese.

Aos membros da CPG pelos auxílios nas questões burocráticas.

Ao Instituto de Química da UNICAMP e aos órgãos de fomento FAPESP, ao INCTBio e ao CNPq.

Ao NPC – Campinas por ceder as amostras de cocaína e crack utilizadas nesta tese.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001.

RESUMO

Neste trabalho, foram desenvolvidos métodos analíticos qualitativos e quantitativos utilizando as espectroscopias ópticas Raman e fluorescência em conjunto com métodos de resolução multivariada de curvas. Foram realizadas quatro aplicações, sendo a primeira a criação de modelos de calibração e validação para identificar adulteração em combustíveis utilizando o método de Resolução Multivariada de Curvas com Mínimos Quadrados Alternantes (MCR-ALS) e restrição de correlação. As amostras analisadas eram compostas de misturas binárias e ternárias de diesel, biodiesel e HEFA na presença de seis adulterantes (etanol, cicloexano, gasolina, mistura de gasolina/etanol na proporção 75/25, hexano e querosene). A segunda aplicação esteve relacionada com a análise de 22 amostras de cocaína ou crack apreendidas pela Polícia Militar de Campinas. Foi possível a identificação de seis outros fármacos presentes nas amostras analisadas utilizando a espectroscopia de imagem Raman, Análise de Componentes Independentes pelo algoritmo JADE e o método de Resolução Multivariada de Curvas com Mínimos Quadrados Alternantes. A partir desses resultados foi possível a caracterização e análise semi-quantitativa de todos os compostos presentes nas amostras. A terceira aplicação, propôs a criação de um modelo de calibração e validação para amostras de óleo de gergelim adulteradas com outros seis óleos vegetais (canola, gergelim não comestível, gergelim comestível de uma marca diferente, girassol, milho e soja). Foram obtidas as superfícies de excitação-emissão para cada uma das amostras e os dados foram tratados utilizando o MCR – ALS com restrição de trilinearidade e PARAFAC. Por fim, a última aplicação esteve relacionada com a determinação dos valores de pK_a e pK_a^* para o fotoácido 2-(2'-hidroxifenil)benzotiazol (HBT). Foram obtidas as superfícies de excitação-emissão de 47 amostras de HBT em diferentes valores de H_0 (escala de acidez de Hammett) e pH. Esses dados foram tratados utilizando o MCR – ALS com e sem a restrição de trilinearidade. Com base nos escores do MCR – ALS foi possível determinar os valores de pK_a e pK_a^* para o HBT.

ABSTRACT

In this study, qualitative and quantitative analytical methods were developed using Raman and fluorescence spectroscopies in conjunction to multivariate curve resolution methods. Four applications were performed, being the first one the creation of calibration and validation models to identify adulteration in fuels using the multivariate curve resolution with alternating least squares (MCR – ALS) and correlation constraint. The samples analyzed were composed by binary and ternary mixtures of diesel, biodiesel and HEFA in the presence of six adulterants (ethanol, cyclohexane, gasoline, blend of gasoline/ethanol, hexane and kerosene). The second application was related to the analysis of 22 samples of cocaine or crack, seized by the Military Police of Campinas. It was possible to identify six other drugs presented in the samples using Raman image spectroscopy, independent component analysis by JADE algorithm and the multivariate curve resolution with alternating least squares. From these results it was possible the characterization and semi-quantitative analysis of all the compounds present in the samples. In the third application, it was proposed the creation of a calibration and validation model for samples of sesame oil adulterated with six other vegetable oils (canola, inedible sesame, edible sesame, sunflower, corn and soybean). Excitation-emission surfaces were obtained for each of the samples and the data were treated using the MCR-ALS with restriction of trilinearity and PARAFAC. Finally, the last application was related to the determination of the values of pK_a and pK_a^* for the 2-(2'-hidroxyphenyl)benzothiazole (HBT) photoacid. It was obtained 47 excitation-emission matrices of HBT at different acid concentrations in H_0 (Hammett acidity) and pH scales. These data were treated using the MCR - ALS with and without the trilinearity constraint. Based on the MCR - ALS scores it was possible to determine the pK_a and pK_a^* values for HBT.

LISTA DE FIGURAS

- Figura 1. Representação esquemática da obtenção dos dois tipos de espalhamentos observados: (a) representação gráfica; (b) níveis energéticos que podem ser acessados a partir da incidência da radiação monocromática. S_0 e S_1 são os estados fundamental e excitado da molécula, V é chamado de estado virtual. As cores verde, azul e vermelho indicam os espalhamentos Rayleigh, Raman Stokes e Raman anti-Stokes, respectivamente.27
- Figura 2. Modos vibracionais que são observados na espectroscopia de espalhamento Raman. (a) Estrutura molecular do éster que pode ser classificado como biodiesel. Modos vibracionais presentes nos ésteres: (b) estiramento simétrico da carbonila (estrutura em vermelho na Figura 2a), (c) estiramentos e deformações presentes na estrutura do CH_2 (estrutura em roxo na Figura 2a), (d) estiramentos e deformações presentes na estrutura $C = C$ (estrutura em verde na Figura 2a); (e) estiramentos e deformações presentes na estrutura do CH_3 (estrutura em roxo na Figura 2a)^{27,28}. .29
- Figura 3. Diagrama de Jablonski, no qual é possível observar os processos de absorção e emissões radiativas e não radiativas. Os estados S_0 , S_1 , S_2 e T_1 representam os estados fundamental, primeiro estado eletrônico excitado, segundo estado eletrônico excitado e primeiro estado tripleto excitado, respectivamente^{18,31}.30
- Figura 4. Representação da obtenção das matrizes de excitação-emissão. (a) Espectros de excitação, em verde, e emissão, em rosa, da amostra; (b) vista bidimensional dos dados referentes à superfície EEM. A linha rosa representa os comprimentos de onda de emissão e a linha verde os comprimentos de onda de excitação.33
- Figura 5. Representação da estrutura de dados tridimensional formada pela união das matrizes de excitação-emissão de várias amostras.34
- Figura 6. Representação dos dois modos de se obter uma imagem química: (a) modo mapeamento ponto-a-ponto; (b) modo mapeamento em linha; (c) modo *imaging* para o LCTF. As coordenadas x e y representam as coordenadas espaciais da imagem e λ representa a coordenada espectral.36
- Figura 7. Representação esquemática da transformação de um conjunto de espectros numa matriz de dados que poderá ser utilizada em quimiometria. Cada linha

representa o espectro de uma amostra e cada coluna são as variáveis (como, por exemplo, números de onda).....	38
Figura 8. Exemplo da aplicação do Filtro de Savitzky-Golay. A linha azul representa um sinal periódico com ruído e a linha vermelha representa o mesmo sinal só que depois do tratamento com o filtro.	39
Figura 9. Representação matricial do modelo bilinear do MCR, onde as matrizes D , C , ST e E possuem dimensões $m \times n$, $m \times k$, $k \times n$ e $m \times n$, respectivamente.	41
Figura 10. Arranjos dos dados em matrizes aumentadas: (a) aumentada no sentido de C , mais amostras são adicionadas; (b) aumentada no sentido de S , adição de mais uma análise para a mesma amostra; (c) aumentada em C e S , mais amostras e análises adicionadas.	43
Figura 11. Reação de produção do HEFA. Em azul e roxo está representada a rota de desoxigenação, em vermelho a rota de hidrotreamento e em verde geração dos alcanos.....	51
Figura 12. Representação esquemática dos dados que são utilizados em cada uma das calibrações: (a) ordem zero; (b) primeira ordem; (c) segunda ordem.....	53
Figura 13. Esquema do funcionamento da restrição de correlação que é feito pelo <i>toolbox</i> MCR – ALS GUI 2.0.....	54
Figura 14. Espectros Raman dos três combustíveis puros: (a) diesel; (b) biodiesel; (c) HEFA.....	60
Figura 15. Espectros Raman obtidos experimentalmente para os conjuntos de calibração e validação.....	61
Figura 16. Primeira derivada dos espectros apresentados na Figura 15.	61
Figura 17. Curvas de calibração para os combustíveis: (a) diesel; (b) biodiesel; (c) HEFA. Amostras de calibração e validação estão representadas em azul e preto, respectivamente. A reta em vermelho é o ajuste linear para as amostras de calibração.	62
Figura 18. Comparação entre as derivadas dos espectros dos combustíveis puros (em azul) e dos recuperados pelo MCR (em vermelho), (a) diesel – correlação de 0,9993, (b) biodiesel – correlação de 0,9933 e (c) HEFA – correlação de 0,9977.	64
Figura 19. Representação esquemática dos dois conjuntos de dados propostos para a identificação de adulterações em amostras de diesel B10: (a) matrizes com somente	

um adulterante; (b) matriz aumentada com todas as amostras adulteradas. D = diesel, B = biodiesel e H = HEFA.	65
Figura 20. Comparação entre as derivadas dos espectros dos combustíveis puros (em azul) e recuperados pelo MCR (em vermelho): (a) cicloexano; (b) etanol; (c) gasolina; (d) hexano; (e) querosene; (f) mistura de gasolina/etanol.	68
Figura 21. Derivada do espectro do cicloexano (a) antes e (b) depois da remoção da região de 750 a 850 cm^{-1}	69
Figura 22. Identificação dos <i>pixels</i> anômalos e remoção de spikes: (a) espectros obtidos experimentalmente; (b) gráfico de T^2 Hotelling contra Resíduos Q; (c) <i>pixel</i> em vermelho é um <i>pixel</i> anômalo e terá seu espectro substituído pelos espectros dos <i>pixels</i> em verde; (d) conjunto de dados após remoção dos <i>pixels</i> anômalos; (e) conjunto de dados após remoção dos <i>spikes</i>	75
Figura 23. (a) Sinal da ICA referente à fenacetina; (b) conjunto de dados de ruído aleatório gerado para ser adicionado ao sinal da fenacetina; (c) espectros referentes à imagem Raman simulada da fenacetina.	77
Figura 24. Desdobramento e concatenação dos dados das imagens hiperespectrais e construção da matriz de dados a partir da concatenação das matrizes da amostra e dos compostos puros.	78
Figura 25. Resultado do ICA por Blocos para uma amostra (Amostra 1). As imagens de (a) até (h) simbolizam os resultados para a extração de 2 a 9 sinais, respectivamente. Em vermelho estão apresentadas as maiores correlações obtidas para as componentes extraídas de cada um dos dois blocos.	79
Figura 26. Resultado dos 20 cálculos do ICA por blocos obtidos para a Amostra 1.	80
Figura 27. Resultado da determinação do número de ICs com base no critério de Durbin-Watson. A coloração azul indica alta correlação entre o número de componentes e as variáveis. Já a coloração vermelha indica a observação de ruído nas ICs extraídas.	80
Figura 28. Análise dos valores singulares da matriz aumentada da Amostra 1.	81
Figura 29. Arranjo da matriz aumenta utilizada para a decomposição do MCR – ALS. A matriz D apresenta retângulos vermelhos e verdes que estão relacionados com os espectros da amostra e dos fármacos/drogas, respectivamente. A matriz C apresenta os resultados dos escores do MCR para a amostra (em vermelho) e dos compostos	

puros (em amarelo e azul). Em amarelo e azul estão representados os escores referentes a ausência e presença dos compostos puros, respectivamente.	82
Figura 30. Matriz de escores obtida após a otimização da matriz D com os métodos de resolução de curvas: (a) matriz C completa; (b) escores a serem normalizados, cada linha representa a imagem de um componente puro (retângulo verde); (c) escores a serem utilizados para a média de cada composto na amostra. Os valores de 0 e 1 representam os escores esperados para ausência e presença de um composto químico na matriz C	83
Figura 31. Sinal obtido com o JADE para a Amostra 3, onde é possível observar que esse sinal não apresenta nenhuma banda que poderia ser relacionada com algum tipo de informação química. Esse sinal está relacionado com a variação de linha de base dos espectros da Amostra 3.....	85
Figura 32. Comparação entre os espectros de referência (em azul) e os sinais (em vermelho) obtidos pela ICA – JADE: (a) cafeína; (b) lidocaína; (c) cocaína.....	86
Figura 33. Imagens químicas da Amostra 3: (a) cafeína; (b) lidocaína; (c) cocaína. As cores azul e vermelho simbolizam um gradiente de proporções químicas desses compostos da sua ausência até a presença nos <i>pixels</i> analisados.	87
Figura 34. (a) Espectro da fenacetina encontrado na literatura ⁹² ; (b) sinal extraído pelo JADE referente à fenacetina.	89
Figura 35. Espectro médio das amostras que apresentaram somente um componente químico após a ICA: (a) Amostra 22 – cocaína; (b) Amostra 19 – crack.....	90
Figura 36. Comparação entre os espectros Raman de referência (em azul) e recuperados pelo MCR – ALS (em vermelho) para a Amostra 1: (a) benzocaína – correlação de 0,9999; (b) cafeína – correlação de 0,9800; (c) levamisol – correlação de 0,9991; (d) lidocaína – correlação de 0,9870; (e) teofilina – correlação de 0,9999; (f) fenacetina – correlação de 0,9954; (g) cocaína – correlação de 0,9992; (h) crack – correlação de 0,9996.....	93
Figura 37. Imagens químicas da Amostra 1: (a) benzocaína; (b) cafeína; (c) levamisol; (d) lidocaína; (e) teofilina; (f) fenacetina; (g) cocaína; (g) crack.	94
Figura 38. Valores dos “limites de identificação” em termos de escores para os compostos analisados.....	95
Figura 39. Cálculo dos escores dos compostos presentes nas amostras para: (a) benzocaína; (b) cafeína; (c) levamisol; (d) lidocaína; (e) teofilina; (f) fenacetina; (g)	

cocaína; (h) crack. A reta em vermelho representa a média acrescida de dois desvios padrão para cada um desses compostos.....	98
Figura 40. Imagens químicas para os compostos identificados para a Amostra 1: (a) cafeína; (b) lidocaína.....	99
Figura 41. Imagens químicas para os compostos identificados para a Amostra 5: (a) cafeína; (b) lidocaína; (c) fenacetina.	99
Figura 42. Imagens químicas para os compostos identificados para a Amostra 16: (a) benzocaína; (b) levamisol; (c) cocaína; (d) crack.....	100
Figura 43. Construção da matriz de dados a partir da concatenação das superfícies de EEM das amostras: (a) MCR – ALS; (b) PARAFAC.....	106
Figura 44. Representação dos espalhamentos Rayleigh e Raman em superfícies EEM.	107
Figura 45. Representação esquemática da análise <i>split-half</i> , sendo que a nomenclatura para essa análise é ‘S ₄ C ₆ T ₃ ’. S simboliza o número de <i>splits</i> (pedaços) em que a matriz de dados foi separada, C é o número de combinações criadas e T o número de testes para as combinações propostas.....	108
Figura 46. Determinação do número de fatores utilizando a análise <i>split-half</i> . Em (a) até (e) para cinco fatores e de (f) até i para quatro fatores. as curvas tracejadas e contínuas estão relacionadas com os espectros de excitação e emissão, respectivamente.....	109
Figura 47. Decomposição da matriz pelo PARAFAC em pesos e escores.	110
Figura 48. Decomposição da matriz de dados com o MCR – ALS com a restrição de trilinearidade.....	111
Figura 49. Superfícies EEM normalizadas dos óleos puros: (a) canola; (b) gergelim não comestível; (c) gergelim comestível marca B; (d) girassol; (e) milho; (f) soja; (g) gergelim comestível marca A.	112
Figura 50. Determinação do número de fatores para o conjunto de amostras adulteradas com óleo de milho utilizando (a) SVD e (b) até (e) análise <i>split-half</i> para quatro fatores.	114
Figura 51. Resultados do MCR – ALS e PARAFAC para as amostras adulteradas com óleo de milho. Espectros de excitação (azul) e emissão (vermelho) para: (a) tocoferóis e fenóis; (b) compostos oxidativos; (c) pigmentos de clorofila; (d) possivelmente relacionado com os diferentes tocoferóis e fenóis presentes.....	116

Figura 52. Resultados do MCR – ALS e PARAFAC para as amostras adulteradas com óleo de soja. Espectros de excitação (azul) o e emissão (vermelho) para: (a) tocoferóis e fenóis; (b) compostos oxidativos; (c) pigmentos de clorofila; (d) possivelmente relacionado com os diferentes tocoferóis e fenóis presentes.	117
Figura 53. Curvas de calibração para as amostras de óleo de gergelim adulterado com óleo de milho. (a) óleo de gergelim pelo MCR-ALS, (b) óleo gergelim pelo PARAFAC, (c) óleo de milho pelo MCR-ALS, (d) óleo de milho pelo PARAFAC.	119
Figura 54. Curvas de calibração para as amostras de óleo de gergelim adulterado com óleo de soja. (a) óleo gergelim	120
Figura 55. Ciclo fotoquímico do processo de ESIPT a 2-(2'-hidroxifenil)benzotiazol e seus enóis derivados.	126
Figura 56. Superfícies de EEM em diferentes concentrações ácidas: (a) H_0 -3,78; (b) pH 5,00; (c) pH 11,50; (d) matriz de dados completa, no eixo y está representado o perfil de emissão e no eixo x é representada a multiplicação entre as 47 amostras e todos os comprimentos de onda de excitação.	128
Figura 57. Equilíbrios químicos reportados na literatura para o HBT e seus derivados no estado fundamental.	129
Figura 58. Espectros de excitação (em vermelho) e emissão (em azul) recuperados pelo MCR – ALS para as três estruturas químicas apresentadas na Figura 55: (a) HBT protonado; (b) HBT neutro; (c) HBT desprotonado.	130
Figura 59. Escores do MCR – ALS em função de H_0 e pH. As linhas em azul, verde e vermelho representam as estruturas do HBT protonado, HBT neutro e HBT desprotonado, respectivamente.	131
Figura 60. Escores do MCR – ALS em função de H_0 e pH para o conjunto de dados não trilineares. As linhas em azul, verde e vermelho representam as estruturas do HBT protonado, HBT neutro e HBT desprotonado, respectivamente (Figura 58).	133
Figura 61. Espectros de emissão recuperados pelo MCR – ALS para o conjunto de dados não trilineares. As linhas em azul, verde e vermelho representam as estruturas do HBT protonado, HBT neutro e HBT desprotonado, respectivamente (Figura 58).	134

LISTA DE TABELAS

Tabela 1. Porcentagem dos volumes de diesel, biodiesel e HEFA adicionados em cada uma das amostras, sendo que o volume final de 20,00 mL.	56
Tabela 2. Porcentagem dos volumes de diesel, biodiesel, HEFA e adulterante adicionados em cada uma das amostras, sendo que o volume final de 22,00 mL....	57
Tabela 3. Atribuição de bandas do espectro Raman para diesel e biodiesel.	60
Tabela 4. Valores de R^2 , LOF, RMSEC e RMSEP para os modelos propostos.....	63
Tabela 5. Resultados de RMSEP e correlação espectral para as misturas binárias e ternárias das amostras adulteradas. Também estão especificados os valores de correlação espectral entre os espectros puros e recuperados pelo MCR.	66
Tabela 6. Identificação dos compostos presentes nas amostras de cocaína/crack com base nos resultados da ICA – JADE.	88
Tabela 7. Atribuição das bandas mais características para cocaína e crack.	90
Tabela 8. Comparação entre os “limites de identificação” com e sem as imagens simuladas. Esses limites estão apresentados em porcentagem de escores do MCR.	96
Tabela 9. Identificação dos compostos presentes nas amostras de cocaína/crack após a análise do “limite de identificação”.	101
Tabela 10. Porcentagem massa/massa de cada um dos óleos utilizados para construção das curvas de calibração e validação.	105
Tabela 11. Resultados para as curvas de calibração de todos os conjuntos de amostras para os dois algoritmos.....	122
Tabela 12. Valores dos comprimentos de onda de excitação e emissão e pK_a calculados pelo ajuste sigmoide dos escores do MCR – ALS (pK_a e pK_a^*) e estimado pela Equação de Förster.	132
Tabela 13. Comparação entre os resultados obtidos com os dados trilineares, bilineares e encontrados na literatura ¹¹⁰	133

LISTA DE ABREVIATURAS

Abreviatura	Descrição em Inglês	Descrição em Português
BSS	<i>Blind Source Separation</i>	Separação Cega de Sinais
CCD	<i>Charge Coupled Device</i>	Dispositivo de Carga Acoplada
EEM	<i>Excitation-Emission Matrix</i>	Matriz de Excitação-Emissão
ESIPT	<i>Excited State Intramolecular Proton Transfer</i>	Transferência de Próton Intramolecular no Estado Excitado
ESPT	<i>Excited State Proton Transfer</i>	Transferência de Próton no Estado Excitado
FAME	<i>Fatty Acid Methyl Ester</i>	Éster Metílico de Ácido Graxo
HBT	<i>2-(2'-hydroxyphenyl)benzothiazole</i>	2-(2'-hidroxifenil)benzotiazol
HEFA	<i>Hydroprocessed Esters and Fatty Acids</i>	Ésteres e Ácidos Graxos Hidroprocessados
HVO	<i>Hydrotreated Vegetable Oil</i>	Óleo Vegetal Hidrotratado
IC	<i>Independent Component</i>	Componente Independente
ICA	<i>Independent Component Analysis</i>	Análise de Componentes Independentes
JADE	<i>Joint Approximate Diagonalization of Eigenmatrices</i>	Diagonalização Aproximada de Automatrizes
LCFT	<i>Liquid Crystal Tunable Filters</i>	Filtros de Cristal Líquido
LOF	<i>Lack of Fit</i>	Falta de Ajuste
MCR – ALS	<i>Multivariate Curve Resolution with Alternating Least Squares</i>	Resolução Multivariada de Curvas com Mínimos Quadrados Alternantes
MSC	<i>Multiplicative Scatter Correction</i>	Correção de Espalhamento Multiplicativo
PARAFAC	<i>Parallel Factors Analysis</i>	Análise de Fatores em Paralelo
PARALIND	<i>Parallel Profiles with Linear Dependencies</i>	Perfis Paralelos com Dependência Linear
PC	<i>Principal Component</i>	Componente Principal

PCA	<i>Principal Component Analysis</i>	Análise de Componentes Principais
PCR	<i>Principal Component Regression</i>	Regressão de Componentes Principais
PLS	<i>Partial Least Squares</i>	Mínimos Quadrados Parciais
R	<i>Correlation Coefficient</i>	Coefficiente de Correlação
R²	<i>Coefficient of Determination</i>	Coefficiente de Determinação
RMSE	<i>Root Mean Square Error</i>	Raiz Quadrada do Erro Quadrático Médio
RMSEC	<i>Root Mean Square Error of Calibration</i>	Raiz Quadrada do Erro Quadrático Médio de Calibração
RMSEP	<i>Root Mean Square Error of Prediction</i>	Raiz Quadrada do Erro Quadrático Médio de Previsão
SNV	<i>Standard Normal Variate</i>	Varição Normal Padrão
SVD	<i>Singular Value Decomposition</i>	Decomposição em Valores Singulares

SUMÁRIO

Capítulo 1 – Motivação.....	21
1.1. Análises Forenses	22
1.2. Fotoquímica e Fotofísica	23
1.3. Análise Multivariada de Dados	24
Capítulo 2 – Introdução	25
2.1. Espectroscopia	26
2.1.1. Espectroscopia Raman	26
2.1.2. Espectroscopia UV/Vis – Fluorescência.....	29
2.1.2.1. Matriz de Excitação-Emissão de Fluorescência (EEM) ^{33,34}	33
2.1.3. Imagens Hiperespectrais.....	34
2.2. Quimiometria	37
2.2.1. Pré-Processamento de dados.....	37
2.2.1.1. Construção da Matriz de Dados	37
2.2.1.2. Remoção de Radiação Espúria dos Espectros Raman	38
2.2.1.3. Alisamento e Derivada.....	38
2.2.1.4. Normalização.....	39
2.2.2. Resolução Multivariada de Curvas com Mínimos Quadrados Alternantes - MCR – ALS ⁴⁷	40
2.2.3. Análise de Componentes Independentes – ICA.....	43
2.2.3.1. Joint Approximate Diagonalization of Eigenmatrices (JADE) ⁵³ ..	44
2.2.4. Análise de Fatores em Paralelo – PARAFAC	46
Capítulo 3 – Objetivos Gerais	48
Capítulo 4 – Análise de Diesel e Biocombustíveis	50
4.1. Introdução Teórica e Motivação	51
4.1.1. Combustíveis Renováveis.....	51
4.1.2. Calibração Multivariada.....	52
4.1.2.1. MCR-ALS com restrição de correlação	54

4.1.2.2. Avaliação dos modelos de calibração.....	55
4.2. Objetivos.....	55
4.3. Procedimento Experimental.....	55
4.3.1. Preparo de Amostra	55
4.3.2. Equipamento Utilizado	58
4.3.3. Tratamento de Dados.....	58
4.4. Resultados e Discussão	59
4.4.1. Modelos de Calibração e Amostras sem Adulteração.....	59
4.4.2. Amostras com Adulteração	64
4.5. Conclusões Parciais	69
Capítulo 5 – Caracterização de Amostras de Cocaína e Crack	71
5.1. Introdução e Motivação	72
5.2. Objetivos.....	72
5.3. Procedimento Experimental.....	73
5.3.1. Amostras	73
5.3.2. Equipamento Utilizado	73
5.3.3. Tratamento de Dados.....	74
5.3.3.1. Identificação de Pixels Anômalos e Remoção de Spikes	74
5.3.3.2. Simulação das Imagens Químicas	76
5.3.3.3. Construção da Matriz de Dados	77
5.3.3.4. Pré-processamento das Imagens Hiperespectrais	78
5.3.3.5. Determinação do Número de Componentes	78
5.3.3.6. Métodos de Resolução de Curvas.....	81
5.3.3.7. Limite de Identificação	82
5.4. Resultados e Discussão	84
5.4.1. Caracterização das amostras com JADE – ICA.....	84
5.4.2. Caracterização das amostras com MCR – ALS e aplicação do limite de identificação.....	91
5.5. Conclusões Parciais	101
Capítulo 6 – Adulteração de Óleo de Gergelim.....	103

6.1. Introdução teórica e Motivação.....	104
6.2. Objetivos.....	104
6.3. Procedimento Experimental.....	104
6.3.1. Preparo de Amostra	104
6.3.2. Equipamento Utilizado	105
6.3.3. Tratamento de Dados.....	105
6.3.3.1. Construção da Matriz de Dados	105
6.3.3.2. Pré-Processamentos	106
6.3.3.3. Determinação do Número de Fatores.....	107
6.3.3.4. Métodos de Resolução de Curvas.....	110
6.4. Resultados e Discussão	111
6.5. Conclusões Parciais	122
Capítulo 7 – Determinação das Constantes de Desprotonação do 2-(2'-hidroxifenil)benzotiazol.....	124
7.1. Introdução teórica e Motivação.....	125
7.2. Objetivos.....	126
7.3. Procedimento Experimental.....	127
7.3.1. Preparo de Amostra	127
7.3.2. Equipamento Utilizado	127
7.3.3. Tratamento de Dados.....	127
7.4. Resultados e Discussão	127
7.5. Conclusões Parciais	134
Capítulo 8 – Conclusões Gerais.....	135
Referências Bibliográficas.....	137
Anexos	143
Anexo 1 – Autorização para análise das amostras de cocaína e crack	144
Anexo 2 – Identificação e substituição de Outliers	145
Anexo 3 – Modos de vibração dos aditivos da cocaína.....	149
Anexo 4 – Soluções de Ácido Clorídrico e Soluções Tampão	152

CAPÍTULO 1
MOTIVAÇÃO

1.1. Análises Forenses

As investigações criminais originaram-se no ano de 1810 em Paris, França. Um criminoso local, Eugène François Vidocq, trocou sua sentença de prisão para se tornar um informante da polícia. Suas contribuições deram início a um novo olhar para as investigações criminais¹.

Anos mais tarde, em 1863, o cientista alemão Christian Schönbein descobriu a oxidação da hemoglobina com peróxido de hidrogênio, sendo essa a primeira técnica para detecção de sangue em evidências criminais^{1,2}. Essa nova atividade policial foi consagrada na literatura por causa de autores como Sir Arthur Conan Doyle com as histórias de Sherlock Holmes (1887). No início do século XX foi publicado o livro *Forensic Chemistry* de Alfred Lucas (1921)². Ele exibiu um panorama geral de como as análises químicas poderiam auxiliar na elucidação dos fatos ocorridos durante o crime sob escrutínio. Hoje em dia há diversos exemplares de livros dedicados a esse tema.

Com o passar dos anos foram obtidos avanços tecnológicos que possibilitaram o desenvolvimento de melhores instrumentos analíticos. Esses avanços proporcionaram a obtenção de uma quantidade maior de informação sobre a amostra em estudo. Assim, a identificação e, até mesmo, quantificação de substâncias que podem estar associadas a crimes de adulterações em medicamentos, alimentos, combustíveis e documentos, tornaram-se mais precisas e confiáveis. Houve também a possibilidade de análises de amostras oriundas de resíduos de disparo de arma de fogo (resíduo residuográfico)³, resíduos de explosivos^{4,5}, resíduos de incêndio, identificação de impressões digitais, tintas, drogas⁶ e identificação de fluidos biológicos (sêmen, sangue e urina)^{7,8}.

Por definição análises forenses são todos os tipos de análises que podem ser aplicadas na solução de problemas associados de alguma forma às autoridades judiciárias². É importante destacar que o perito criminal é responsável pela interpretação dos resultados obtidos, sendo que suas conclusões serão utilizadas para definir se a pessoa suspeita é culpada ou inocente. Por esse motivo, o perito deve estar ciente de que seus resultados devem ser verdadeiros e inquestionáveis. Consequentemente, para que isso seja possível, ele deve estar munido de técnicas capazes de comprovar as suspeitas dos investigadores.

Existem diversas técnicas instrumentais que podem ser utilizadas na análise das evidências coletadas em uma cena de crime. Podem ser citadas as técnicas de espectrometria de massas, eletroforese capilar, as cromatografias líquida e gasosa, além das técnicas espectroscópicas como Raman, infravermelho médio, infravermelho próximo, fluorimetria de raios X, ultravioleta/visível (UV-Vis) e espectrometria atômica (emissão e absorção)⁹.

É extensa a gama de técnicas instrumentais que podem ser utilizadas, sendo que a limitação não é qual técnica utilizar, mas sim o quanto de amostra está disponível para ser analisada. Por causa disso, é imprescindível o uso de técnicas que possam evitar a perda ou destruição da evidência. Com base nessa limitação é importante utilizar técnicas espectroscópicas, pois elas são capazes de manter a integridade física e obter informações relevantes e conclusivas sobre as amostras¹⁰.

1.2. Fotoquímica e Fotofísica

A espectroscopia é a área que estuda os efeitos que ocorrem a partir da interação da radiação eletromagnética com a matéria. Podem ser obtidos, por exemplo, espectros que possuem informações estruturais sobre a molécula em estudo ou observar processos fotoquímicos e fotofísicos. Os processos fotoquímicos envolvem a transformação da matéria devido a absorção de radiação. Já a fotofísica é o estudo dos processos de absorção e emissão de radiação e transferência de energia, observando as diferentes escalas de tempo e diferentes condições experimentais, com finalidade de entender o comportamento dos materiais¹¹.

A relevância da fotoquímica também reside em suas variadas aplicações em ciência e tecnologia. A fotoquímica orgânica sintética forneceu vários métodos para fabricação de muitos produtos químicos que não poderiam ser produzidos por reações sem a presença da radiação eletromagnética. Além disso, esses métodos possuem como vantagem maior eficiência e seletividade. Podem ser citados como exemplos viáveis do ponto de vista industrial as sínteses fotoquímicas: (1) síntese da vitamina D a partir de certos esteróis isolados; (2) síntese de cubanos, os quais são agentes antivirais; (3) síntese industrial de caprolactama, como monômero de Nylon 6; (4) fabricação de solventes e inseticidas por fotocloração; (5) síntese de antioxidantes por fotosulfonação, etc¹².

1.3. Análise Multivariada de Dados

O desenvolvimento de técnicas instrumentais proporcionou a obtenção de enormes quantidades de dados a partir de uma única medida. Porém, quando há grande quantidade de informação se torna praticamente impossível seu estudo sem o uso de ferramentas computacionais. Essas informações podem estar relacionadas com propriedades físico-químicas, comprimentos de onda em espectrometria, tempos de retenção em cromatografia, valores de potencial em voltametria, dentre outras respostas analíticas.

Visando o estudo de grande quantidade de dados, definiu-se a Quimiometria como as operações associadas ao tratamento matemático e à interpretação de dados químicos¹³. Esses tratamentos estão relacionados com a aplicação de métodos de estatística multivariada e pré-processamento de sinais em química. Esses métodos podem ser divididos em: classificação multivariada, calibração multivariada, resolução de curvas e aprendizado de máquina.

Em especial, os métodos de resolução de curvas proporcionam resultados que podem ser diretamente comparados com algumas observações experimentais, como: espectros, voltamogramas ou cromatogramas dos componentes químicos presentes nas amostras. Além disso, eles fornecem informações quantitativas dos componentes presentes no sistema estudado. Dentro dos métodos para resolução de curvas podem ser citadas a resolução multivariada de curvas com mínimos quadrados alternantes (*Multivariate Curve Resolution with Alternating Least Squares - MCR – ALS*)¹⁴, análise de fatores paralelos (*Parallel Factor Analysis - PARAFAC*)¹⁵ e análise de componentes independentes (*Independent Component Analysis - ICA*)¹⁶.

CAPÍTULO 2
INTRODUÇÃO

2.1. Espectroscopia

2.1.1. Espectroscopia Raman

O efeito Raman recebeu esse nome em homenagem ao físico indiano *Sir Chandrasekhra Venkata Raman*, que em 1928 observou e descreveu o efeito de espalhamento inelástico da luz em líquidos. Concomitantemente, os físicos soviéticos Grigory Samuilovich Landsberg e Leonid Mandelstam observaram o mesmo espalhamento em cristais¹⁷. Por fim, o trabalho de Raman representava um estudo mais detalhado e com isso ele foi laureado com o Prêmio Nobel de física em 1930. No entanto, o físico teórico Adolf Smekal propôs matematicamente, em 1923, que seria possível a observação desse tipo de espalhamento inelástico da radiação¹⁷.

A interação de radiação monocromática com a matéria pode gerar os efeitos de absorção e espalhamento de luz, Figura 1a, desde que a luz incidente possua energia e frequências adequada para que esses efeitos sejam observados. A absorção de radiação UV/Vis ocorre quando a energia do fóton incidente é igual a diferença de energia entre os níveis eletrônicos fundamental (S_0) e excitado (S_1) da molécula¹⁸. No caso da espectroscopia vibracional, a absorção da radiação ocorre quando a energia do fóton incidente é igual a diferença de energia entre os níveis vibracionais fundamental (ν_0) e excitado (ν_i com $i \neq 0$) do estado eletrônico fundamental¹⁹.

A radiação incidente pode ser espalhada de forma elástica ou inelástica. O espalhamento elástico ocorre quando a radiação espalhada possui a mesma frequência da radiação incidente. O espalhamento elástico da luz também é chamado de espalhamento Rayleigh. Já o espalhamento inelástico ocorre quando a radiação espalhada possui frequência deslocada em comparação à frequência da radiação incidente. O espalhamento inelástico da radiação é chamado de espalhamento Raman¹⁹.

A incidência da radiação com energia menor que a necessária para a absorção gera um estado virtual (V) na molécula, Figura 1b. Esse estado é chamado de virtual simplesmente por não ser solução da equação de Schroedinger. Logo em seguida, a molécula volta para o estado fundamental; no entanto, ela pode voltar para um estado vibracional diferente do fundamental. A diferença de energia entre os estados vibracionais inicial (ν_i) e final (ν_f) irá definir qual espalhamento será obtido¹⁹. Quando ν_i for menor do que ν_f será observado o espalhamento Raman Stokes, e

quando ν_i for maior do que ν_f será observado o espalhamento Raman anti-Stokes. Essa nomenclatura foi dada em homenagem ao matemático e físico irlandês George Gabriel Stokes, pioneiro nos estudos de fluorescência¹⁹.

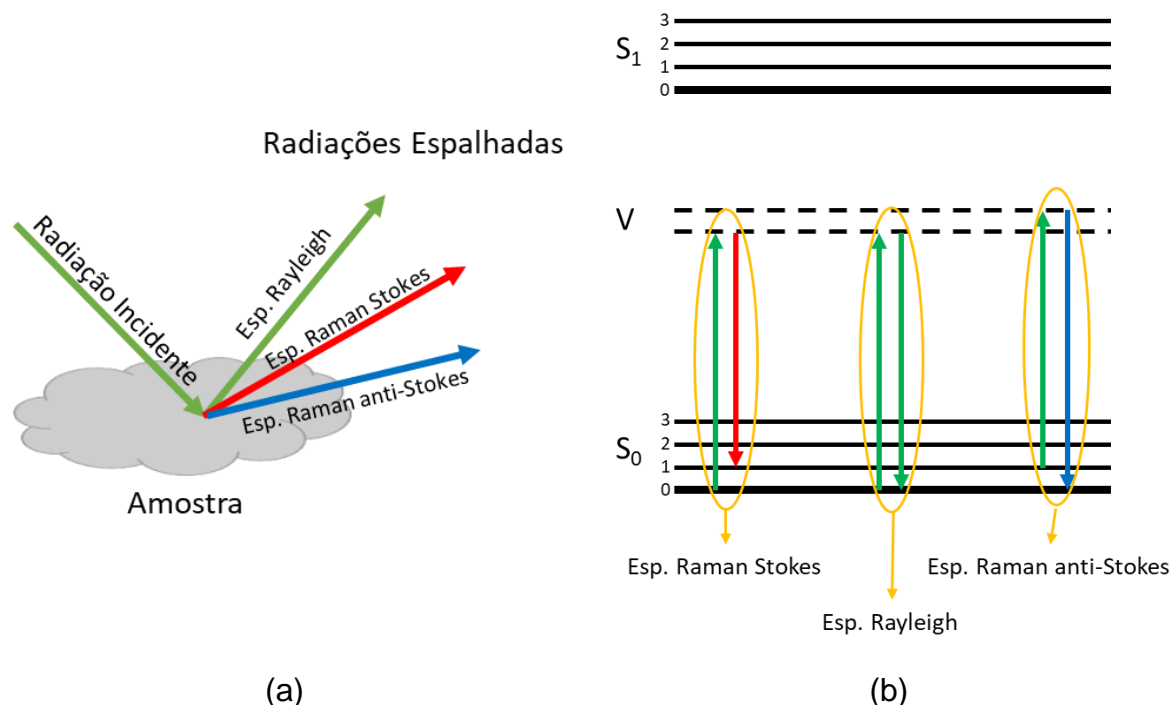


Figura 1. Representação esquemática da obtenção dos dois tipos de espalhamentos observados: (a) representação gráfica; (b) níveis energéticos que podem ser acessados a partir da incidência da radiação monocromática. S_0 e S_1 são os estados fundamental e excitado da molécula, V é chamado de estado virtual. As cores verde, azul e vermelho indicam os espalhamentos Rayleigh, Raman Stokes e Raman anti-Stokes, respectivamente.

Os ramos Stokes e anti-Stokes possuem bandas com o mesmo número de onda, mas com deslocamentos diferentes em relação ao espalhamento Rayleigh, e elas possuem diferentes intensidades. As intensidades estão relacionadas com a distribuição das moléculas no estado fundamental e excitado, sendo que é seguida a distribuição de Boltzmann. Nessa distribuição os níveis menos energéticos são ocupados pelas moléculas, ou seja, quanto menor for a energia do nível vibracional haverá maior população associada a ele. Com isso, o nível fundamental vibracional possui muito mais moléculas que o primeiro nível vibracional excitado do estado eletrônico fundamental. Por consequência, mais moléculas presentes nesse nível irão interagir com a radiação incidente gerando espectros mais intensos para o ramo Stokes que as moléculas que gerarão o ramo anti-Stokes. Por causa disso, o ramo

Stokes é utilizado para aplicações analíticas, como identificação de adulterações em leite^{20,21} e combustíveis²², identificação de drogas lícitas²³ e ilícitas^{24,25}, explosivos²⁶, dentre outras.

Por ser uma espectroscopia vibracional, o espectro Raman é analisado de forma similar à espectroscopia no infravermelho médio, onde são observados os modos vibracionais correspondentes à estiramentos, deformações e torções. A Figura 2 apresenta como exemplo, todos os tipos de modos vibracionais que podem ser observados na espectroscopia Raman para os ésteres como os encontrados no biodiesel.

Esses modos vibracionais são classificados de acordo com o efeito que eles causam na estrutura molecular. Os estiramentos estão relacionados com o movimento rítmico ao longo do eixo que conecta os átomos, fazendo com que haja aumento e diminuição da distância interatômica. Além disso, esses estiramentos são divididos em simétricos e assimétricos. Também existem as deformações, que fazem com que haja alterações nos ângulos das ligações químicas e são divididas em tesoura, *wagging*, *twisting* e *rocking*.

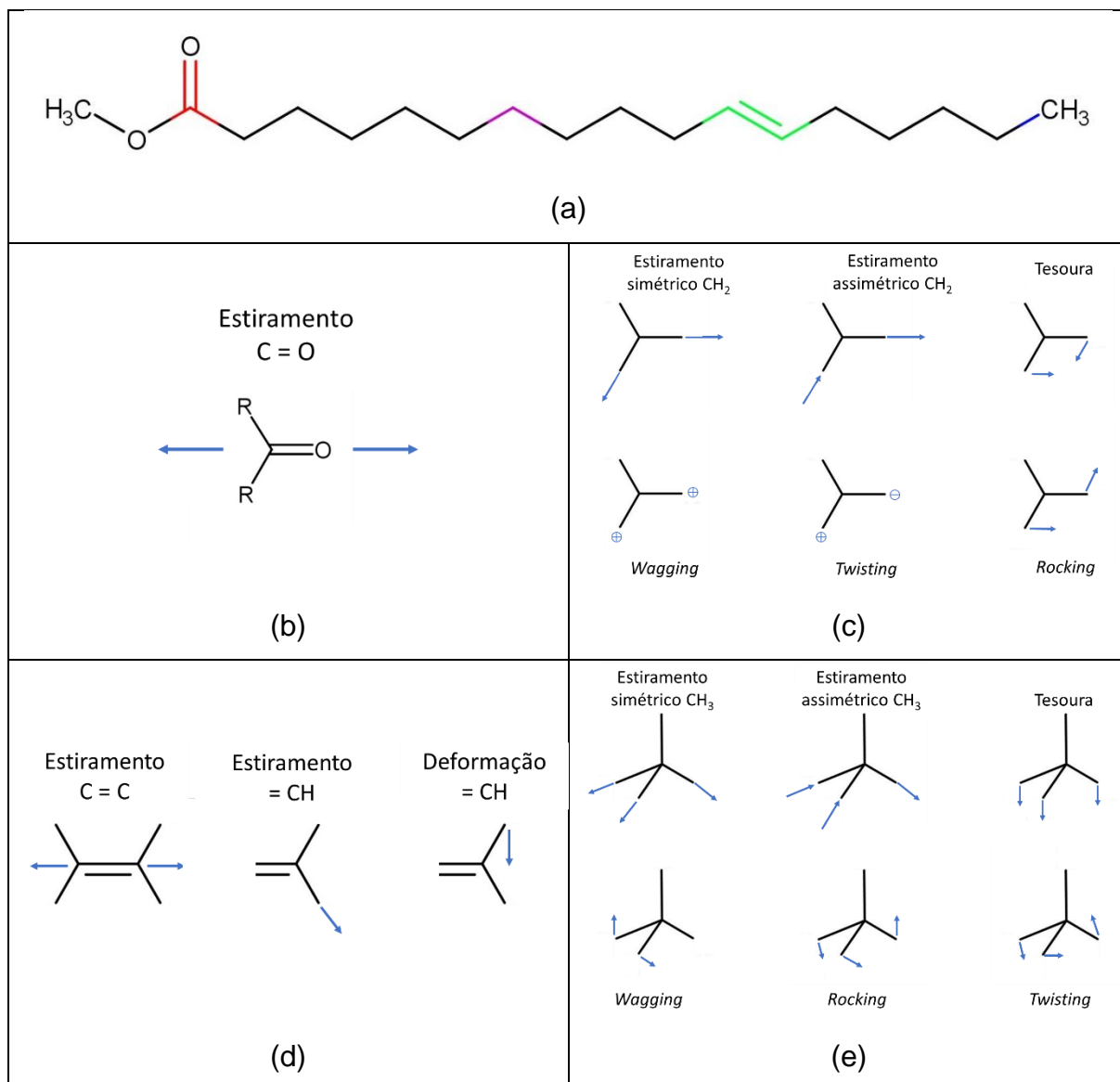


Figura 2. Modos vibracionais que são observados na espectroscopia de espalhamento Raman. (a) Estrutura molecular do éster que pode ser classificado como biodiesel. Modos vibracionais presentes nos ésteres: (b) estiramento simétrico da carbonila (estrutura em vermelho na Figura 2a), (c) estiramentos e deformações presentes na estrutura do CH₂ (estrutura em roxo na Figura 2a), (d) estiramentos e deformações presentes na estrutura C = C (estrutura em verde na Figura 2a); (e) estiramentos e deformações presentes na estrutura do CH₃ (estrutura em roxo na Figura 2a)^{27,28}.

2.1.2. Espectroscopia UV/Vis – Fluorescência

Fluorescência é um fenômeno de emissão de radiação que pertence à luminescência. A palavra luminescência vem do latim *lumen* que significa luz e foi

primeiramente descrita pelo físico alemão Eilhardt Wiedemann em 1888. Wiedemann a descreveu como “todos os fenômenos de luz que não são apenas condicionados pelo aumento da temperatura”, sendo assim oposta à incandescência. Esse antagonismo leva dizer que a luminescência é uma luz fria e a incandescência uma luz quente²⁹. A partir dessa descrição entende-se que os processos luminescentes estão diretamente ligados à emissão de radiação. Mais especificamente, a luminescência é um fenômeno que ocorre em moléculas que estão no estado excitado e que perdem energia de forma radiativa, voltando para o estado fundamental.

Os fenômenos de absorção e emissão de radiação foram estudados mais profundamente pelo físico polonês Alexander Jablonski. Em 1933, ele descreveu que as moléculas orgânicas poderiam ser representadas por três níveis energéticos, dois estados singletos (S_0 e S_1) e um estado tripleto (T_1)³⁰. A Figura 3 apresenta o diagrama de Jablonski (atualizado) com os processos de absorção e emissões radiativas (fluorescência e fosforescência) e não radiativas (relaxação vibracional, conversão interna, conversão externa e cruzamento intersistemas).

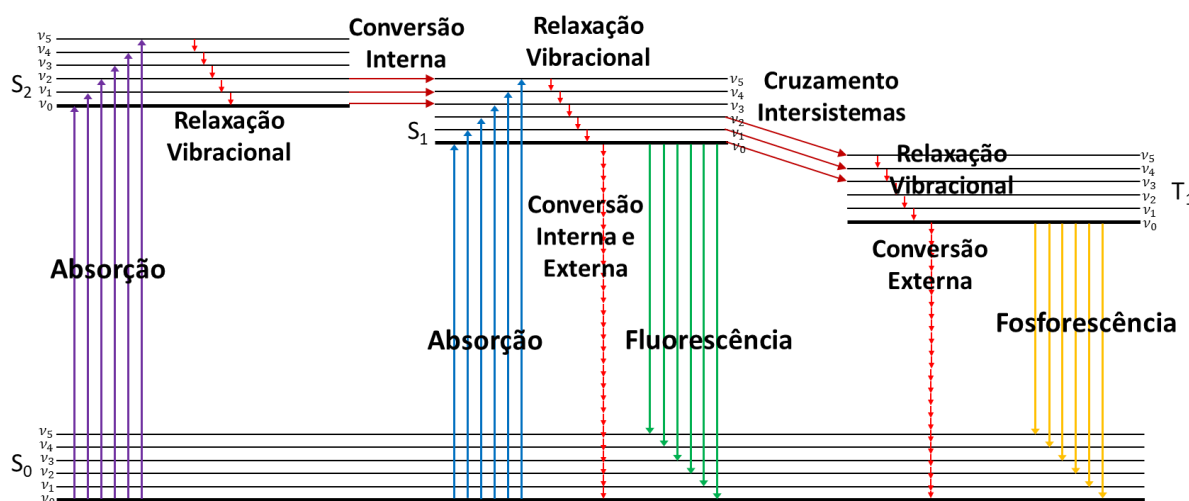


Figura 3. Diagrama de Jablonski, no qual é possível observar os processos de absorção e emissões radiativas e não radiativas. Os estados S_0 , S_1 , S_2 e T_1 representam os estados fundamental, primeiro estado eletrônico excitado, segundo estado eletrônico excitado e primeiro estado tripleto excitado, respectivamente^{18,31}.

Na Figura 3 foram apresentados os processos observados na região de UV/Vis^{18,31}. Esses processos são divididos em:

- *Absorção*: ocorre quando a radiação incidente possui o mesmo comprimento de onda da diferença entre os níveis eletrônicos fundamental (S_0) e excitado (S_1 e S_2). Essas transições estão representadas pelas setas violeta e azul na Figura 3;
- *Relaxação Vibracional*: transferência de energia entre uma molécula no estado vibracional excitado para o solvente, sendo que após essa transferência a molécula irá para o nível vibracional de menor energia do estado eletrônico em que ela se encontra. Durante esse processo não é observada emissão dessa radiação, sendo assim chamado de processo não radiativo. Entretanto, pode ser observado leve aumento da temperatura do solvente. Além disso, uma consequência direta da relaxação vibracional é a banda de fluorescência possuir maior comprimento de onda que o da radiação absorvida (essa diferença é chamada de deslocamento Stokes). Essa transição está representada pelas setas vermelhas na Figura 3;
- *Conversão Interna*: é a transição entre dois estados vibracionais de estados eletrônicos distintos, mas, necessariamente, de mesma multiplicidade de spin (transição entre dois estados singletos ou tripletos). Esse processo pode levar ao fenômeno de pré-dissociação, ou seja, quando a molécula foi para um nível vibracional que possui uma energia vibracional grande o suficiente favorecendo a quebra da ligação química. Essa transição está representada pelas setas vermelhas na Figura 3;
- *Conversão Externa*: desativação do estado excitado a partir da transferência de energia entre o estado excitado e o solvente do meio; essa transferência de energia se dá pela colisão entre essas moléculas. Essa transição está representada pelas setas vermelhas na Figura 3;
- *Fluorescência*: é um fenômeno observado quando uma molécula emite energia na faixa de 380 a 800 nm, região do visível. Essa emissão ocorre quando uma molécula decai do estado excitado S_1 para o estado fundamental S_0 , como demonstrado pelas setas em verde na Figura 3;
- *Cruzamento Intersistemas*: é um processo em que há um cruzamento entre estados eletrônicos de multiplicidades diferentes, conforme demonstrado no estado singleto para um estado de tripleto (S_1 a T_1). Durante essa transição

ocorre mudança de spin molecular, de emparelhado para não emparelhado. Essa transição está representada pelas setas vermelhas na Figura 3;

- *Fosforescência*: é um fenômeno observado quando a molécula emite radiação na faixa do visível, entretanto, seu comprimento de onda deve ser maior que o observado para a fluorescência. Essa emissão ocorre quando uma molécula decai do estado excitado T_1 para o estado fundamental S_0 , como demonstrado pelas setas em amarelo na Figura 3.

Alguns fatores como a presença de duplas ligações conjugadas (sistemas π conjugados), rigidez estrutural, temperatura, pH, solvente e a presença de átomos pesados podem determinar quais sistemas uma molécula poderá percorrer no estado excitado.

Com base nesses fatores e nos processos radiativos e não radiativos que influenciam a fluorescência é possível a obtenção de seu rendimento quântico (ϕ), Equação 1, sendo uma razão entre o número de fótons emitidos por fluorescência e o número de fótons absorvidos. Além disso, a Equação 1 depende das constantes de velocidade de todos os processos que podem ocorrer após a absorção da radiação.

$$\phi = \frac{k_f}{k_f + k_i + k_{ce} + k_{ci} + k_{pd} + k_d} \quad \text{Equação 1}$$

onde k_f , k_i , k_{ce} , k_{ci} , k_{pd} e k_d são as constantes de velocidade da fluorescência, cruzamento intersistemas, conversão externa, conversão interna, pré-dissociação e dissociação, respectivamente.

Por fim, é possível obter uma relação matemática entre a intensidade de fluorescência da amostra com a concentração do fluoróforo na solução. A intensidade de fluorescência (F) é proporcional ao produto de uma constante de eficiência quântica de fluorescência (K') vezes a diferença entre a intensidade da radiação incidente (P_0) e a intensidade após essa radiação passar pela amostra percorrendo um caminho óptico (b), Equação 2.

$$F = K'(P_0 - P) \quad \text{Equação 2}$$

Utilizando a Lei de Lambert-Beer e a expansão de MacLaurin é possível determinar que F será proporcional à concentração (c), Equação 3³², onde K representa a expressão $2,3K'\epsilon bP_0$.

$$F = Kc \quad \text{Equação 3}$$

2.1.2.1. Matriz de Excitação-Emissão de Fluorescência (EEM)^{33,34}

Nas medidas for fluorescência molecular podem ser obtidas as matrizes de excitação-emissão de fluorescência, a partir dos espectros de excitação e emissão de uma amostra. Os espectros de excitação são produzidos pela medida da intensidade de luminescência da amostra, onde o comprimento de onda de emissão é mantido fixo e varia-se o comprimento de onda de excitação (Figura 4a espectro em verde). Já os espectros de emissão são obtidos de forma contrária, mede-se a intensidade da luminescência de uma amostra mantendo-se fixo o comprimento de onda de excitação e varia-se o comprimento de onda de emissão (Figura 4a espectro em rosa).

Ao final são obtidas as matrizes de excitação-emissão de fluorescência, essa matriz possui suas dimensões de comprimentos de onda independentes. Cada linha corresponde à um comprimento de onda de excitação (linha verde na Figura 4b) e suas colunas são os comprimentos de onda de emissão (linha rosa na Figura 4b). Esses dados podem ser projetados como um arranjo bidimensional, onde é possível observar todas as bandas de excitação e emissão dos fluoróforos presentes.

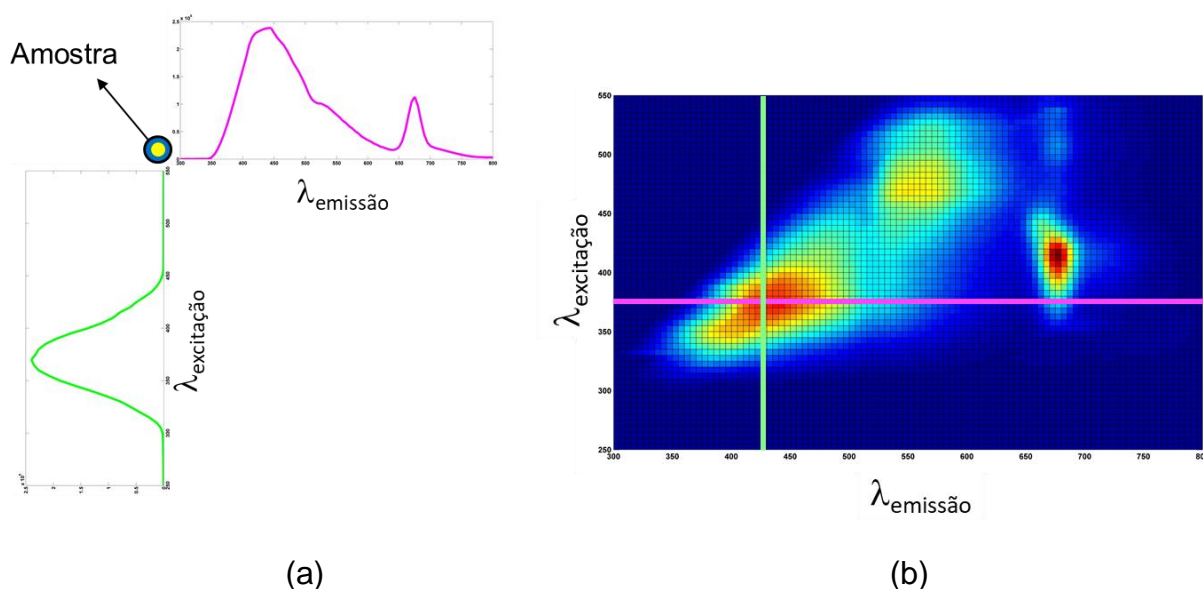


Figura 4. Representação da obtenção das matrizes de excitação-emissão. (a) Espectros de excitação, em verde, e emissão, em rosa, da amostra; (b) vista bidimensional dos dados referentes à superfície EEM. A linha rosa representa os

comprimentos de onda de emissão e a linha verde os comprimentos de onda de excitação.

A junção destes espectros para várias amostras gera um arranjo tridimensional de fluorescência total, com todas as bandas de excitação e emissão dos fluoróforos presentes, como mostrado na Figura 5. Para amostras onde existe mais de um fluoróforo presente (na matriz da amostra ou em diferentes analitos), a dificuldade de interpretação direta desta superfície é maior, pois podem ocorrer grandes superposições espectrais.

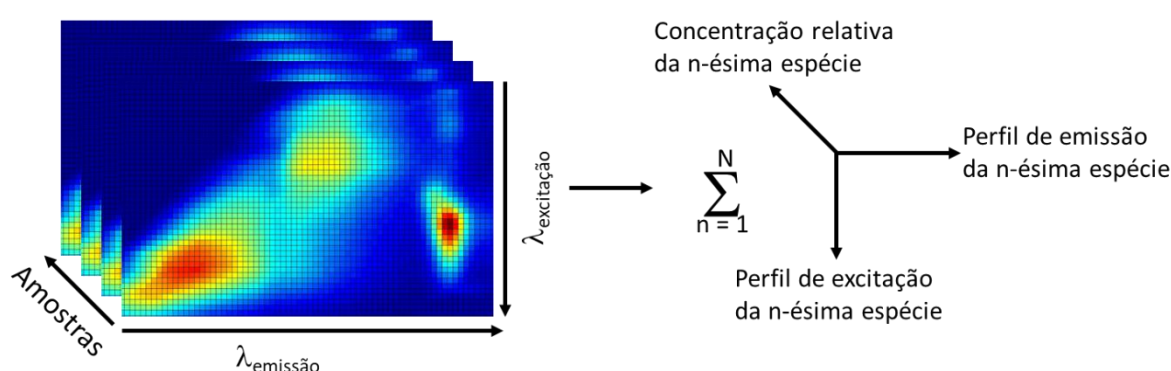


Figura 5. Representação da estrutura de dados tridimensional formada pela união das matrizes de excitação-emissão de várias amostras.

2.1.3. Imagens Hiperespectrais

Essa técnica foi desenvolvida em 1985 por Alexander Goetz *et al*³⁵ e desde então vários avanços foram realizados para que seja possível quantificar³⁶ e identificar³⁷ quais são as regiões em que os analitos podem ser encontrados. Nessa técnica, foi preciso acoplar um microscópio no caminho da radiação, antes desta ter contato com a amostra, possibilitando a escolha de uma região específica para ser analisada.

Os avanços mais importantes para a coleta de dados estão relacionados aos diferentes modos de serem obtidas as imagens espectroscópicas. Esses modos são separados em duas classes, *imaging* e mapeamento^{38,39}. A Figura 6 representa esquematicamente a obtenção das imagens para esses modos³⁸.

No modo mapeamento os dados podem ser coletados de duas formas, a primeira é realizando uma varredura da área selecionada com um tamanho de *pixel* definido pelo analista. O equipamento irá realizar a coleta dos dados ponto-a-ponto (*pixel-a-pixel*, Figura 6a) até que tenha obtido uma imagem completa da área selecionada.

O segundo método é chamado de mapeamento em linha (Figura 6b); neste tipo de medida uma linha inteira de cada amostra é iluminada e com isso são obtidos os espectros nessa única direção da amostra. Essa metodologia apresenta um tempo de coleta de dados menor que o da metodologia ponto-a-ponto, mas as imagens obtidas são mais limitadas, já que é muito difícil manter uma iluminação igual em todos os pontos que serão analisados. Essa diferença na iluminação de todas as partes da amostra resultará na quantidade de informação que é obtida em cada um dos *pixels* analisados. Quanto menor for a quantidade de *pixels* analisados, em uma mesma área, é esperada uma imagem com menor resolução espacial.

Já para o *imaging*, toda a área a ser analisada é iluminada ao mesmo tempo e é obtida uma imagem para cada comprimento de onda. Como no caso do mapeamento em linha, os métodos de imageamento são mais rápidos na obtenção dos resultados, quando comparados com a metodologia ponto-a-ponto.

O *imaging* pode ser realizado com filtros de cristal líquido (LCTF, do inglês *Liquid Crystal Tunable Filters*), Figura 6c. Essa técnica é capaz de transmitir e/ou bloquear comprimentos de onda específicos que são selecionados pelo analista⁴⁰, obtendo espectros com melhores resoluções espectrais.

Ao final da coleta é obtido um arranjo de dados com três dimensões (x , y e λ), onde os valores de x e y correspondem às coordenadas espaciais de cada um dos *pixels* e λ é a dimensão que corresponde ao número de onda do espectro. O valor que aparece em cada elemento da matriz é o valor da intensidade do espectro para cada número de onda nas coordenadas xy ²⁶.

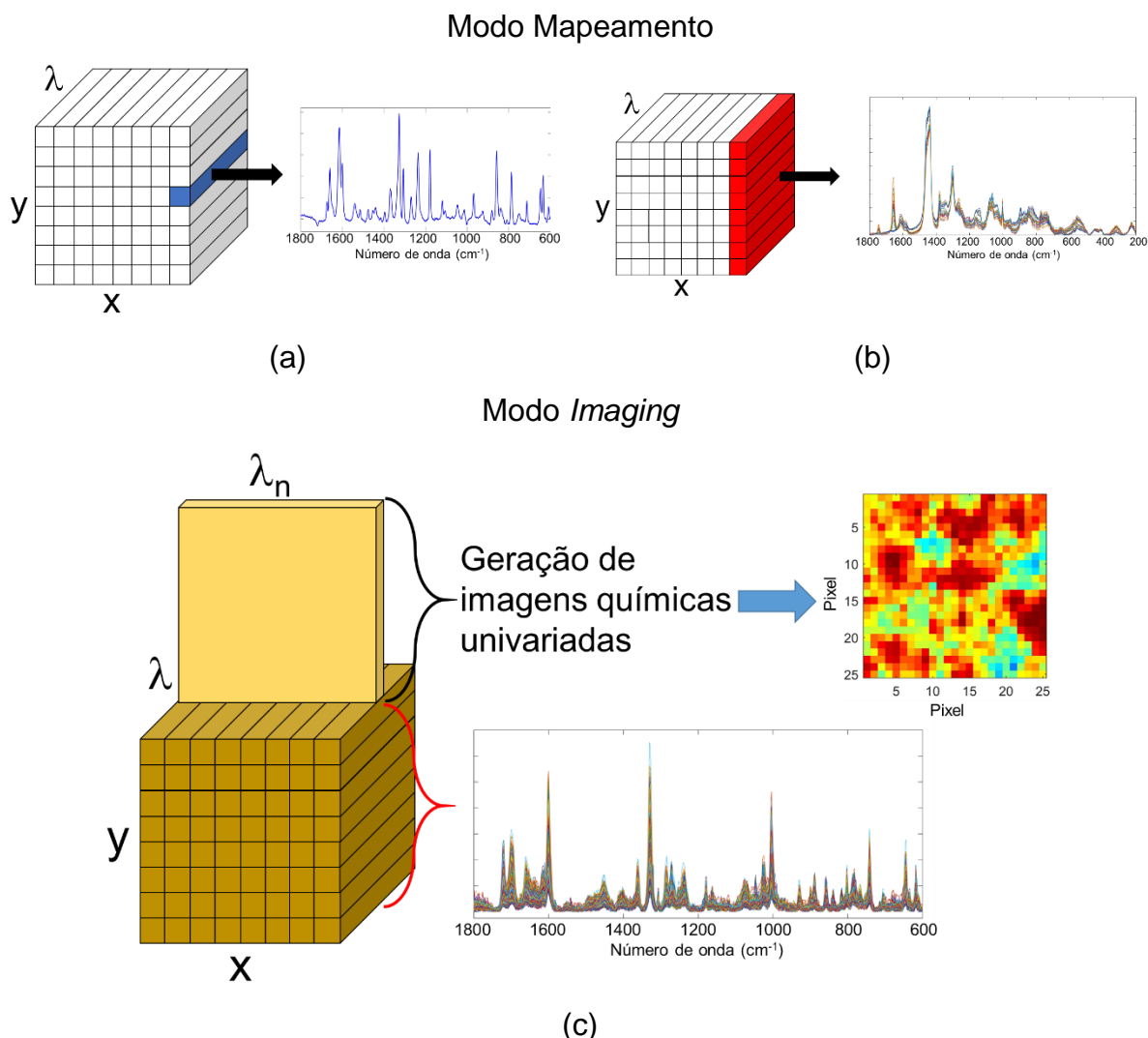


Figura 6. Representação dos dois modos de se obter uma imagem química: (a) modo mapeamento ponto-a-ponto; (b) modo mapeamento em linha; (c) modo *imaging* para o LCTF. As coordenadas x e y representam as coordenadas espaciais da imagem e λ representa a coordenada espectral.

Muitos dos artigos publicados atualmente não atentam mais para a primeira definição sobre as imagens espectroscópicas^{2,11}, dada a partir da quantidade de informações espectrais que as formavam. Elas eram definidas em três categorias: imagens multiespectrais (espectros com menos que 10 pontos espectrais), imagens hiperespectrais (espectros com pontos espectrais entre 10 e 1000) e imagens ultraespectrais (espectros com mais de 1000 pontos espectrais). Esses pontos

espectrais estão relacionados com a resolução espectral dos dados coletados. Como, por exemplo, para uma análise com resolução espectral de 8 cm^{-1} na faixa espectral de 200 a 240 cm^{-1} há seis pontos espectrais coletados ($200, 208, 216, 224, 232$ e 240 cm^{-1}), sendo assim classificada como uma imagem multiespectral.

A obtenção das imagens espectrais, também chamadas de imagens químicas, pode ser realizada de forma univariada, a partir da observação dos espectros dos analitos puros. Nesse procedimento são observados quais os comprimentos de onda (λ_n) que somente são observados no analito de interesse. Com o valor conhecido do comprimento de onda, utilizam-se os valores de intensidade presentes em seu plano xy para construir uma imagem de diferentes cores que irão indicar a presença ou ausência dessa banda específica, Figura 6c. Normalmente, a cor vermelha está associada com a presença do analito escolhido e a cor azul está associada com sua ausência.

Muitas vezes não é possível encontrar um comprimento de onda que seja específico para o analito de interesse. Nesse caso imagens químicas ainda podem ser obtidas com auxílio de métodos quimiométricos de resolução de curvas, onde os espectros das misturas são decompostos nos espectros dos compostos puros presentes.

A espectroscopia de imagem possibilita adquirir informações espaciais das amostras, como as obtidas utilizando MALDI, DESI, raios-X, etc⁴¹. É possível identificar em quais regiões da amostra estão presentes os analitos de interesse e quando possível determinar sua concentração.

2.2. Quimiometria

2.2.1. Pré-Processamento de dados

2.2.1.1. Construção da Matriz de Dados

Em quimiometria os dados são organizados na forma de uma matriz, onde suas linhas e colunas representam as amostras e variáveis, respectivamente. Essas variáveis podem ser os tempos de retenção (cromatografia), números de onda (espectroscopia), comprimentos de onda (espectroscopia), razão massa/carga (espectrometria de massas), dentre outros. A Figura 7 apresenta um exemplo de como são as matrizes de dados que serão analisadas pelos diferentes métodos quimiométricos.

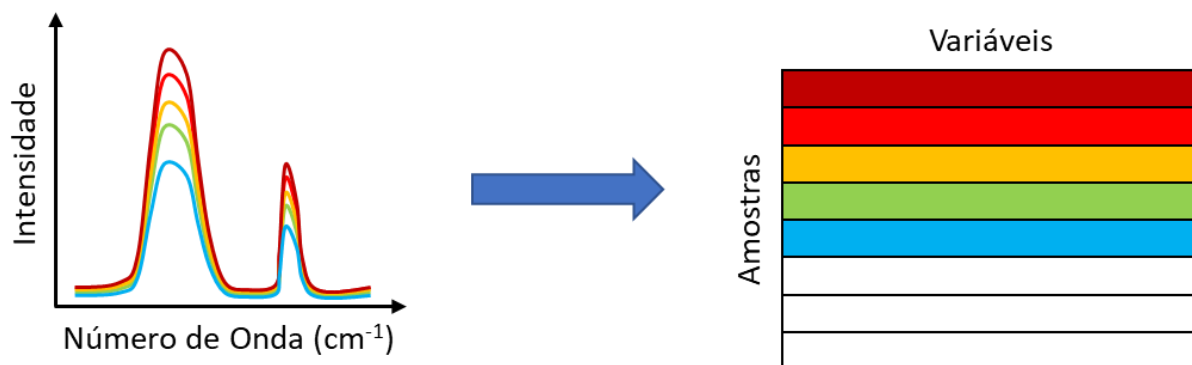


Figura 7. Representação esquemática da transformação de um conjunto de espectros numa matriz de dados que poderá ser utilizada em quimiometria. Cada linha representa o espectro de uma amostra e cada coluna são as variáveis (como, por exemplo, números de onda).

2.2.1.2. Remoção de Radiação Espúria dos Espectros Raman

Os dados de espectroscopia Raman podem apresentar sinais que não estão relacionados com a estrutura molecular da amostra. Esses sinais são totalmente aleatórios e imprevisíveis e estão muitas vezes relacionados com os raios cósmicos. Esses sinais são comumente chamados de *spikes* e são bandas muito finas e intensas. Esses sinais aparecem devido ao uso de detectores tipo CCD (*Charge Coupled Device*) para a região do visível na maioria dos espectrômetros utilizados em equipamentos Raman. Como esses sinais não tem relação com as amostras e podem interferir em posterior análise eles devem ser removidos do conjunto de dados⁴².

2.2.1.3. Alisamento e Derivada

Dependendo da resolução espectral escolhida para obtenção dos espectros é possível observar sinais ruidosos. O alisamento dos dados proporciona uma suavização do ruído presente nos dados experimentais. O filtro mais comum, chamado de filtro Savitzky-Golay, foi desenvolvido em 1964 pelos pesquisadores Abraham Savitzky e Marcel Golay, ambos funcionários da empresa Perkin Elmer⁴³.

Essa técnica utiliza um intervalo constante de pontos (janela de dados) em que é utilizado um polinômio de grau N (a escolha do analista) que ajusta os pontos para aproximar os resultados. Essa janela move-se ao longo de todo o espectro, possibilitando a suavização de espectros que possuem componentes de baixa frequência e com relação sinal/ruído maior que 3.

A Figura 8 representa um conjunto de dados com ruído antes (linha azul) e após (linha vermelha) um tratamento de alisamento utilizando o filtro Savitzky-Golay com uma janela de nove pontos e com um polinômio de segundo grau.

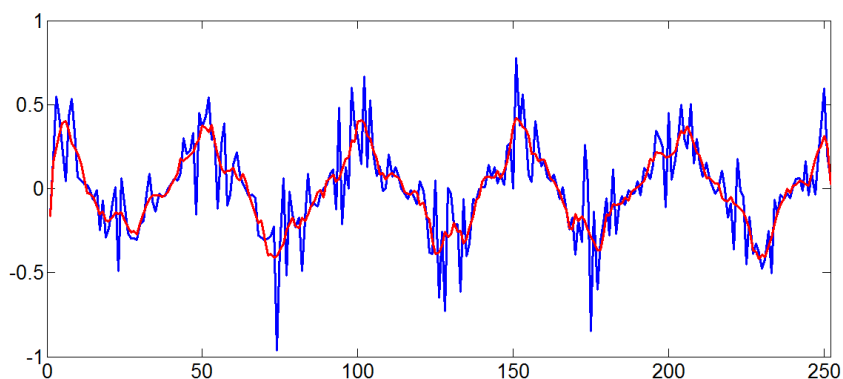


Figura 8. Exemplo da aplicação do Filtro de Savitzky-Golay. A linha azul representa um sinal periódico com ruído e a linha vermelha representa o mesmo sinal só que depois do tratamento com o filtro.

Além de realizar o alisamento dos dados, o filtro Savitzky-Golay também pode ser utilizado para realizar a derivada dos espectros. Para esse procedimento também é preciso definir uma janela de dados e ajusta-se um polinômio de grau N . A função é então derivada e utiliza-se os dados da função derivada como os novos pontos do espectro. Também é possível definir a ordem da diferenciação que será realizada.

A derivação dos dados é utilizada na maioria das vezes para corrigir variações na linha de base dos espectros e também podem ajudar na identificação de bandas características para cada componente da amostra, proporcionando assim melhor resolução dos espectros obtidos. A primeira aplicação de derivadas nos tratamentos de dados é datada de 1983 utilizando a segunda derivada de espectros no infravermelho de proteínas em água por Susi *et al*⁴⁴.

2.2.1.4. Normalização

Em experimentos de espectroscopia é comum obter espectros com diferentes intensidades das bandas. Essas diferenças são causadas a partir de diferentes caminhos ópticos, efeitos de espalhamento, variação na intensidade da fonte ou do detector ou outros efeitos de natureza instrumental. Causas químicas ou físicas na amostra também podem comprometer a obtenção de um espectro no equipamento, como exemplo, a posição da amostra em relação ao detector. Essas

alterações modificam a intensidade das bandas obtidas durante o experimento e isso pode levar resultados não totalmente corretos.

Por isso, faz-se necessário o uso de algum tipo de processamento nos espectros obtidos para que essas diferenças sejam minimizadas durante a análise dos dados. Há diversos tipos de processamentos que podem ser realizados durante o tratamento de dados, dentre elas podem ser citadas, a normalização pelo valor máximo de intensidade do espectro, SNV (*Standard Normal Variate*), MSC (*Multiplicative Scatter Correction*) dentre outras^{45,46}. Os espectros contêm informações químicas, parte dessa informação pode ser perdida durante a normalização, porém a informação estrutural permanece a mesma.

2.2.2. Resolução Multivariada de Curvas com Mínimos Quadrados Alternantes - MCR – ALS⁴⁷

O método de resolução multivariada de curvas (MCR – *Multivariate Curve Resolution*) tem como foco a obtenção de informações relevantes dos componentes puros de um sistema a partir de um modelo bilinear que decompõe a matriz de dados experimentais, \mathbf{D} , nos produtos das matrizes \mathbf{C} e \mathbf{S}^T , como demonstrado pela Equação 4 e pela Figura 9. A matriz \mathbf{C} contém informações sobre a concentração de cada um dos componentes da mistura e a matriz \mathbf{S}^T contém as informações dos espectros de cada um dos componentes. O número de linhas da matriz \mathbf{S}^T corresponde ao número de componentes do sistema (que pode ser estimado pelo cálculo do SVD – análise de valores singulares⁴⁵). O número de colunas da matriz \mathbf{C} também representa o número de componentes do sistema estudado. A matriz \mathbf{E} , por sua vez, refere-se aos resíduos, ou seja, as informações contidas nessa matriz não são explicadas pelo modelo proposto pelo MCR.

$$\mathbf{D} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \quad \text{Equação 4}$$

O diagrama ilustra a equação $\mathbf{D} = \mathbf{C}\mathbf{S}^T + \mathbf{E}$ usando representações visuais de matrizes:

- Uma caixa quadrada à esquerda contém o símbolo \mathbf{D} e, abaixo dela, o rótulo $m \times n$.
- Igualdade (=) segue.
- Uma caixa vertical retangular contém o símbolo \mathbf{C} e, abaixo dela, o rótulo $m \times k$.
- Um símbolo \times está entre as caixas de \mathbf{C} e \mathbf{S}^T .
- Uma caixa horizontal retangular contém o símbolo \mathbf{S}^T e, abaixo dela, o rótulo $k \times n$.
- Um símbolo $+$ está entre as caixas de $\mathbf{C}\mathbf{S}^T$ e \mathbf{E} .
- Uma caixa quadrada à direita contém o símbolo \mathbf{E} e, abaixo dela, o rótulo $m \times n$.

Figura 9. Representação matricial do modelo bilinear do MCR, onde as matrizes \mathbf{D} , \mathbf{C} , \mathbf{S}^T e \mathbf{E} possuem dimensões $m \times n$, $m \times k$, $k \times n$ e $m \times n$, respectivamente.

O algoritmo do método de resolução multivariada de curvas com mínimos quadrados alternantes (MCR – ALS) soluciona o modelo básico bilinear utilizando restrições que melhoram a interpretação dos perfis das matrizes \mathbf{C} e \mathbf{S}^T . As restrições, que podem ser aplicadas tanto à concentração quanto aos espectros, possuem papel importante ao ajudar a definir uma solução com sentido químico e não apenas matemático. Não negatividade no espectro e concentração, unimodalidade, balanço de massa, são alguns exemplos de restrições.

Uma das restrições que pode ser utilizada a fim de construir modelos de calibração é a restrição de correlação. Nesse tipo de restrição são construídos modelos de calibração univariados internos a partir das concentrações calculadas pelo algoritmo do MCR e os valores reais de concentração. Com base nesse modelo proposto são utilizadas amostras de validação para determinar se o modelo proposto é adequado.

O algoritmo dos mínimos quadrados alternantes (*Alternating Least Squares* - ALS) é um modelo iterativo de MCR que resolve a Equação 4 visando a melhora dos resultados dos perfis de concentração e espectros obtidos em cada etapa da otimização. Ao final da otimização é esperado que tenha sido obtido o melhor ajuste para o conjunto de dados iniciais. Abaixo estão listadas as etapas que são realizadas pelo algoritmo^{45,48}:

- Detecção do número de componentes;
- Estimativas iniciais de \mathbf{C} ou \mathbf{S}^T para o ALS;
- Estimativas dos perfis de concentração (se tem estimativa inicial de \mathbf{S}) ou espectrais (se tem estimativa inicial de \mathbf{C}). Para a estimativa dos perfis dos espectros e das concentrações são utilizadas as Equações 5 e 6, respectivamente;

$$\mathbf{S} = \mathbf{C}^+ \mathbf{D} \quad \text{Equação 5}$$

$$\mathbf{C} = \mathbf{D}(\mathbf{S}^T)^+ \quad \text{Equação 6}$$

- Processo de otimização é iniciado, resolvendo por iteração as duas equações levando em consideração as restrições selecionadas;

- Analisar se a convergência de \mathbf{CS}^T foi satisfatória. Em caso negativo o processo deve ser iniciado novamente até se obter convergência satisfatória.

Uma vez estimadas as matrizes \mathbf{C} e \mathbf{S} nas etapas indicadas pelas Equações 5 e 6, seu produto pode ser calculado (\mathbf{CS}^T) e uma nova matriz, \mathbf{D}^* , é reconstruída a cada iteração. Essas etapas são repetidas até que algum critério de convergência estabelecido previamente seja alcançado, como um número máximo de iterações a ser realizado ou caso a variação entre o resultado da iteração atual e aquele da iteração anterior seja menor que um valor pré-estabelecido. É importante ressaltar que, ainda que a convergência seja alcançada nessa etapa iterativa, a solução obtida não é necessariamente a melhor resposta para o conjunto de dados estudado, pois há sempre o risco de se alcançar um mínimo local apesar das restrições pré-definidas.

Dessa comparação da matriz de dados originais \mathbf{D} com essa nova matriz \mathbf{D}^* podem ser calculados dois importantes parâmetros, que fornecem informações sobre a qualidade do modelo construído: a porcentagem de falta de ajuste (LOF – *Lack of Fit*) e a porcentagem de variância explicada (R^2). O primeiro desses parâmetros é calculado conforme descrito na Equação 7 e compara cada termo da nova matriz ($e_{ij} = c_{ij}s_{ij}^T - d_{ij}$, onde i e j estão relacionados com a i -ésima linha e a j -ésima coluna da matriz \mathbf{D}) com o termo correspondente da matriz original (d_{ij}). Como o desejado é que essas matrizes sejam as mais semelhantes, os valores dos erros (e) devem ser pequenos e o valor de %LOF deverá estar mais próximo a zero^{49,50}.

$$\%LOF = 100x \sqrt{\frac{\sum e_{ij}^2}{\sum d_{ij}^2}} \quad \text{Equação 7}$$

O segundo parâmetro segue o mesmo raciocínio, comparando as duas matrizes, mas dessa vez descrevendo o quanto da informação original está sendo modelada pelo MCR, ou seja, o quanto da matriz original \mathbf{D} está contida na nova matriz \mathbf{D}^* . A forma de calcular a porcentagem da variância explicada é descrita na Equação 8 e o ideal é que 100% da informação original seja expressa pelo modelo sendo, portanto, desejado um valor para R^2 próximo a esse.^{49,50}

$$R^2 = 100x \left(1 - \frac{\sum e_{ij}^2}{\sum d_{ij}^2} \right) \quad \text{Equação 8}$$

A matriz de dados **D** pode ser construída a partir da concatenação de diversas matrizes de dados. Essas matrizes podem ser construídas de três maneiras diferentes, a primeira é realizada colocando uma matriz abaixo da outra, e assim aumenta o número de linhas da matriz **D**, nesse caso aumenta-se o número de amostras presentes no conjunto de dados (Figura 10a). No segundo caso, coloca-se uma matriz ao lado da outra, aumentando o número de colunas da matriz **D**, aumenta o número de variáveis presentes. A concatenação de duas ou mais matrizes de dados da mesma amostra, mas obtidas com técnicas diferentes é chamada de fusão de dados (Figura 10b). E no terceiro caso, são aumentados os números de linhas e o número de colunas da matriz **D** (Figura 10c).

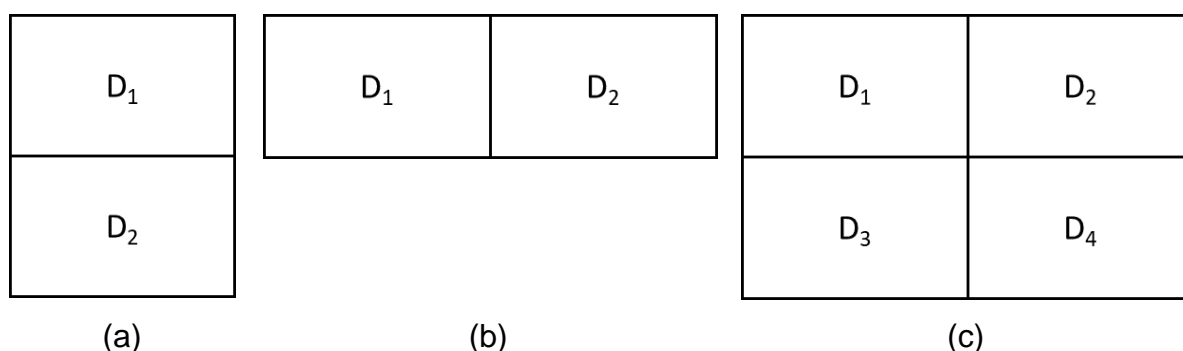


Figura 10. Arranjos dos dados em matrizes aumentadas: (a) aumentada no sentido de **C**, mais amostras são adicionadas; (b) aumentada no sentido de **S**, adição de mais uma análise para a mesma amostra; (c) aumentada em **C** e **S**, mais amostras e análises adicionadas.

2.2.3. Análise de Componentes Independentes – ICA

A análise de componentes independentes originou-se a partir da análise cega de sinais (BSS – *Blind Source Separation*)^{51,52}, em que não é necessário o conhecimento prévio das componentes presentes no sistema estudado.

Do ponto de vista matemático a ICA pode ser descrita da seguinte forma:

$$\mathbf{X} = \mathbf{AS} + \mathbf{E} \quad \text{Equação 9}$$

onde a matriz \mathbf{X} corresponde ao conjunto de dados experimentais, \mathbf{A} é a matriz referente à matriz de proporções dos componentes “puros” desconhecidos e a matriz \mathbf{S} é a matriz de sinais “puros” que correspondem aos espectros dos componentes presentes na amostra. Como no caso do MCR, se a matriz \mathbf{X} possui dimensões iguais à $i \times j$, as matrizes \mathbf{A} e \mathbf{S} terão as dimensões $i \times c$ e $c \times j$, respectivamente. A dimensão c corresponde ao número de componentes independentes que foram identificados na decomposição dos sinais.

Para o cálculo da matriz \mathbf{S} é essencial o desenvolvimento de uma matriz de sinais que torne os ICs estatisticamente independentes entre si, o que não ocorre no caso do MCR. Essa diferença é explicada pelo fato de que ICA é baseada na Teoria Central do Limite⁵², que prova que para que essas ICs sejam estatisticamente independentes há necessidade de que a matriz de resíduos siga uma distribuição gaussiana. A determinação da matriz \mathbf{S} é realizada da seguinte forma

$$\mathbf{S} = \mathbf{W}\mathbf{X} = \mathbf{A}^{-1}\mathbf{A}\mathbf{S} \quad \text{Equação 10}$$

onde \mathbf{W} é chamada de matriz de pesagem (sendo que $\mathbf{W} = \mathbf{A}^{-1}$).

2.2.3.1. *Joint Approximate Diagonalization of Eigenmatrices (JADE)*⁵³

O algoritmo JADE é uma aproximação algébrica da ICA, em que é realizada uma diagonalização para otimização dos cumulantes de segunda e quarta ordens. Os cumulantes mais importantes em análise de sinais são os de primeira até quarta ordem, e eles estão relacionados com a média, variância, medida de assimetria e curtose, respectivamente.

A primeira etapa da ICA envolve a realização de uma PCA com a matriz \mathbf{X} centrada na média, sendo que o objetivo dessa etapa é o branqueamento e redução do número de linhas da matriz \mathbf{X} . O branqueamento é um pré-processamento onde a matriz \mathbf{X} é transformada de tal forma que as suas colunas se tornam ortogonais. Vários métodos já foram propostos na literatura para determinar o número ideal de ICs a serem extraídos de um dado conjunto de dados, mas a maioria deles depende do conhecimento prévio dos dados. Nesse sentido, Boveresse *et. al*⁵⁴ propuseram dois métodos baseados em dados experimentais, chamados de ICA por Blocos e o critério Durbin-Watson⁵⁵.

2.2.3.1.1. *ICA por Blocos (ICA-by-Blocs)*

Diferentemente da SVD essa técnica realiza uma estimativa do número de ICs com base na correlação entre as ICs que são obtidas a partir dos blocos que são gerados pelo algoritmo. Esse cálculo pode ser sumarizado da seguinte forma:

- Escolha do número de blocos em que os dados serão divididos; normalmente os dados são divididos em dois blocos. A construção de cada um dos blocos é realizada com a divisão aleatória das amostras nesses dois blocos selecionados;
- Em seguida, é escolhido o número máximo de ICs que serão extraídos de cada um dos blocos (normalmente são utilizados 20 ICs);
- Após a escolha do número de ICs e do número de blocos o algoritmo irá realizar a extração das 20 ICs dos dados. Primeiramente serão extraídas duas ICs, e será calculada a correlação entre elas. Uma correlação alta entre duas ICs significa que essa IC estimada corresponde a uma IC verdadeira. Esse procedimento é realizado para todas as componentes que são estimadas por esse algoritmo;
- A partir de todos os valores de correlação obtidos entre todas as ICs pode ser obtido um gráfico que relaciona o número de ICs com as correlações que foram obtidas. É esperado que com aumento do número de ICs extraídas haja diminuição na correlação obtida para as ICs posteriores.
- Esse procedimento pode ser realizado diversas vezes para o mesmo conjunto de dados (múltiplo ICA por blocos), para isso devem ser construídos diferentes blocos para que não sejam obtidos resultados iguais. Essa etapa é a mais demorada, pois a obtenção dessas ICs não é uma etapa trivial e por isso deve ser realizada diversas vezes para que os resultados obtidos sejam reprodutíveis. Uma determinação equivocada do número de componentes pode gerar resultados que contenham uma quantidade de ruído muito grande (no caso de serem obtidos mais ICs do que o necessário). Por conta disso, há a necessidade de realizar diversas vezes a determinação do número de ICs.

2.2.3.1.2. *Critério de Durbin-Watson*

O critério de Durbin-Watson (Equação 11) trata-se de um critério que mede a relação sinal/ruído das ICs extraídas. Primeiramente, calcula-se a matriz de

concentrações utilizando de um até o máximo de componentes independentes possíveis. Posteriormente, utiliza-se a matriz residual de cada componente para obter uma estimativa do valor de Durbin-Watson (DW). A análise dos valores de DW pode revelar o número adequado de fatores a serem extraídos no modelo.

$$DW = \frac{\sum_{i=2}^n (s(i) - s(i-1))^2}{\sum_{i=1}^n s(i)^2} \quad \text{Equação 11}$$

Os valores de $s(i)$ representam os elementos i do espectro s na matriz de dados \mathbf{X} , ou seja, são os valores da intensidade de uma variável (número de onda) dos espectros da matriz \mathbf{X} . O valor de DW tende a zero quando não há ruído no sinal extraído pela ICA.

2.2.4. Análise de Fatores em Paralelo – PARAFAC

O algoritmo do PARAFAC é um modelo iterativo que visa a decomposição de um arranjo tridimensional $\underline{\mathbf{X}}$ em três matrizes diferentes (\mathbf{A} , \mathbf{B} e \mathbf{C}), como representado pela Equação 12. Essas matrizes são divididas em dois conjuntos, o primeiro possui as informações das respostas instrumentais medidas (matrizes \mathbf{A} e \mathbf{B}), o segundo possui informações sobre cada uma das amostras (matriz \mathbf{C}). Para encontrar o melhor resultado da combinação entre essas matrizes o modelo iterativo procura minimizar a soma dos quadrados dos resíduos do modelo, $\underline{\mathbf{E}}$. As matrizes \mathbf{A} , \mathbf{B} e \mathbf{C} sempre terão o mesmo número de fatores, o que não acontece para o modelo de Tucker¹³. Com isso, o PARAFAC é considerado uma generalização do modelo de Tucker quando as três matrizes possuem o mesmo número de fatores.

$$\underline{\mathbf{X}} = \mathbf{A}(\mathbf{C}|\otimes|\mathbf{B})^T + \underline{\mathbf{E}} \quad \text{Equação 12}$$

Esse modelo iterativo também utiliza o ALS como no caso do MCR. Também assim como no MCR é preciso haver alguma estimativa inicial para que o algoritmo encontre a solução para $\underline{\mathbf{X}}$. Essa etapa é realizada com a estimativa de \mathbf{B} e \mathbf{C} , onde \mathbf{A} pode ser obtido por mínimos quadrados, Equação 13^{56,57}.

$$\mathbf{X} = \mathbf{A}(\mathbf{C}\odot\mathbf{B})^T \quad \text{Equação 13}$$

onde $(C \odot B)$ é a matriz obtida pelo produto de Khatri-Rao de B e C . Os seguintes passos demonstram como o algoritmo do PARAFAC funciona^{13,56}:

- Determinar o número de fatores presentes no conjunto de dados⁵⁸;
- Estimar as matrizes B e C ;
- Estimar A utilizando a Equação 14,

$$\mathbf{Z}^T = (\mathbf{C} \odot \mathbf{B})^T \quad \text{Equação 14}$$

$$\min_A \|\mathbf{X}^{I \times J \times K} - \mathbf{AZ}^T\|^2 \quad \text{Equação 15}$$

onde I , J e K são as dimensões das matrizes A , B e C , respectivamente. Sendo que todas elas possuem o mesmo número de fatores igual a F ;

- Estimar B como feito para a estimativa de A , $\min_B \|\mathbf{X}^{J \times K I} - \mathbf{BZ}^T\|^2$;
- Estimar C como feito para a estimativa de A e B , $\min_C \|\mathbf{X}^{K \times I J} - \mathbf{CZ}^T\|^2$;
- Observar a convergência, caso haja convergência desejada o processo é finalizado, caso não haja convergência o processo é retornado para a estimativa de A .

As matrizes obtidas ao final desse processo possuem informações sobre as duas dimensões de resposta obtidas experimentalmente (matrizes A e B) e sobre as amostras analisadas (matriz C). A partir do conhecimento dos constituintes do sistema é possível sua comparação com os resultados obtidos para essas matrizes. Para o caso da construção de curvas de calibração a matriz C possui as informações relacionadas às proporções de cada um dos fatores em todas as amostras analisadas.

**CAPÍTULO 3 –
OBJETIVOS GERAIS**

De forma geral esta Tese visou a aplicação de técnicas espectroscópicas ópticas, como as espectroscopias Raman e de fluorescência, aliadas com métodos de resolução de curvas para a análise de amostras de interesse forense e fotoquímico. Essas aplicações incluem a construção de modelos de calibração para análise de amostras adulteradas de diesel/biodiesel/HEFA utilizando a restrição de correlação do MCR – ALS; caracterização de amostras de cocaína utilizando imagens hiperespectrais com ICA – JADE e MCR – ALS; construção de modelos de calibração para a análise de amostras adulteradas de óleo de gergelim utilizando a espectroscopia de fluorescência, MCR – ALS e PARAFAC e determinação das constantes de desprotonação de um fotoácido (2-(2'-hidroxifenil)benzotiazol) com a espectroscopia de fluorescência e MCR – ALS.

**CAPÍTULO 4 –
ANÁLISE DE DIESEL
E BIOCOMBUSTÍVEIS**

4.1. Introdução Teórica e Motivação

4.1.1. Combustíveis Renováveis

Atualmente, há um aumento pela procura de combustíveis alternativos ou pela implementação dos existentes. Isso está relacionado diretamente com o aumento da demanda de fontes alternativas que minimizem a emissão de gases poluentes. O biodiesel⁵⁹ é o combustível alternativo mais utilizado no Brasil. Ele é obtido a partir da conversão de óleos vegetais ou animais em éster metílico de ácido graxo (FAME – *Fatty Acid Methyl Ester*). Essa reação é realizada em baixas temperaturas para favorecer a reação de transesterificação em catálise homogênea. O biodiesel de soja, utilizado nesse estudo, é uma mistura de ésteres metílicos derivados de triglicerídeos com ácidos graxos de cadeias carbônicas variando de C16 até C18 e com um número de insaturações variando de 0 até 3.

Outro combustível que vem sendo utilizado como substitutivo do diesel é o HEFA (Ésteres e Ácidos Graxos Hidroprocessados)^{60,61}, classificado como óleo vegetal hidrotratado (HVO – *Hydrotreated Vegetable Oil*). Esses óleos são produzidos através da conversão de triglicerídeos em alcanos. Essa transformação ocorre em duas etapas: a primeira é a conversão de triglicerídeos em alcanos utilizando catálise de desoxigenação seguida de uma etapa de saturação das insaturações. A segunda etapa é a catálise de isomerização e hidrotratamento, sendo ao final obtidos alcanos isômeros e na fração apropriada para ser considerado diesel (Figura 11). O HEFA é considerado um diesel renovável por ser uma mistura de n-parafinas e i-parafinas com cadeias carbônicas variando de C15 até C18. Além disso, ele é livre de compostos aromáticos, oxigenados, nitrogenados e sulfonados e possuem alto número de cetanos.

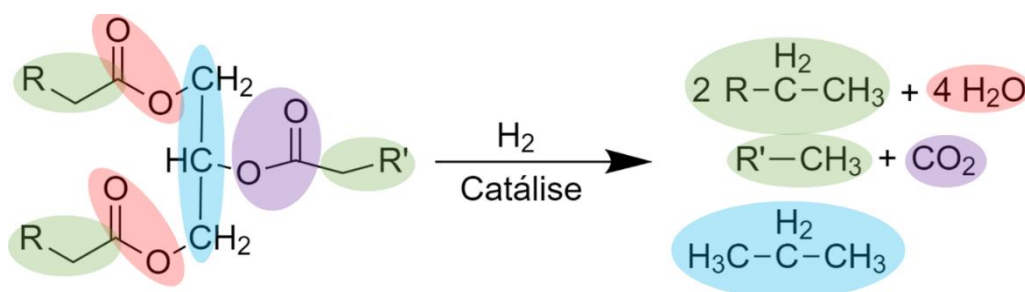


Figura 11. Reação de produção do HEFA. Em azul e roxo está representada a rota de desoxigenação, em vermelho a rota de hidrotratamento e em verde geração dos alcanos.

O diesel é uma mistura de hidrocarbonetos saturados com cadeias carbônicas de C10 a C18, formado principalmente de hidrocarbonetos naftênicos e hidrocarbonetos aromáticos. A quantidade relativa desses compostos no diesel pode variar devido às características de petróleo bruto e aos processos de refinação utilizados⁶².

O mercado dos combustíveis renováveis está aumentando a cada ano e com isso começa a ser observada a presença de combustíveis adulterados no mercado⁶³. No Brasil, de 2008 até 2013, aumentou a quantidade de biodiesel presente no diesel, passando de 2% (diesel B2) para 5% (diesel B5)⁶⁴. Há alguns estudos governamentais que indicam que até 2020 será obrigatório o uso do diesel com 10% de biodiesel em sua composição (diesel B10)⁶⁵.

Por causa desses fatores acredita-se que haverá aumento na adulteração desses combustíveis. Estimasse que em 2020 o Brasil irá consumir em torno de 70 bilhões de litros de diesel por ano⁶⁵. Os adulterantes podem causar problemas como menor desempenho do motor, no controle de partida do motor, aquecimento do motor, aceleração e consumo de combustível; também aumenta a emissão de material particulado, hidrocarbonetos e gases de exaustão^{64,66}. A adulteração de diesel tem sido reportada na literatura, sendo encontrados trabalhos com os adulterantes etanol, ciclohexano, gasolina, mistura de gasolina/etanol na proporção 75/25, hexano e querosene⁶⁷⁻⁶⁹.

4.1.2. Calibração Multivariada

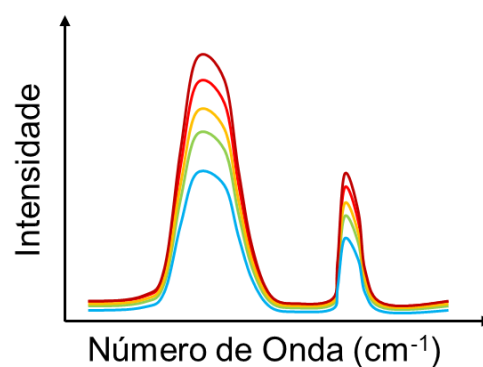
Os métodos de calibração são classificados de acordo com a complexidade ou dimensionalidade dos dados. Com isso, eles são divididos em ordem zero, primeira e segunda ordem⁷⁰. Para as calibrações de ordem zero é utilizado apenas um único valor de medida experimental por amostra, como por exemplo, a construção de uma curva de calibração utilizando valores de absorbância para um conjunto de amostra (Figura 12a). Como a medida da absorbância pode ser realizada no comprimento de onda máximo de absorção teremos somente uma medida experimental por amostra.

A calibração de primeira ordem utiliza um vetor de medidas instrumentais (um espectro ou cromatograma, Figura 12b) para cada uma das amostras analisadas. Diferentemente da calibração de ordem zero, essas análises possibilitam a análise na presença de interferentes presentes no conjunto de amostra. Entretanto, esses

interferentes devem estar presentes no conjunto de calibração. Os métodos quimiométricos de regressão por mínimos quadrados parciais (PLS) e regressão por componentes principais (PCR) são os mais utilizados para construção de modelos de calibração multivariada em dados químicos.

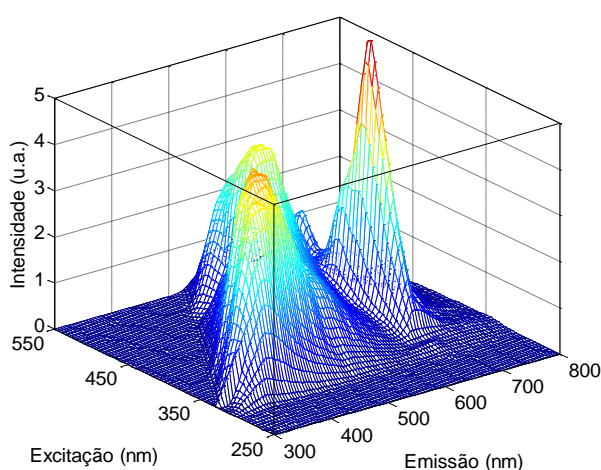
E, por fim, a calibração de segunda ordem é obtida a partir de um conjunto de dados em arranjos tridimensionais (superfícies de excitação-emissão de fluorescência ou dados de espectros obtidos a diferentes em tempos, Figura 12c). A grande vantagem desses dados é a possibilidade de quantificar a espécie de interesse na presença de interferentes, mesmo quando eles não são conhecidos no conjunto de calibração. Essa vantagem é conhecida como vantagem de segunda ordem. Para o tratamento desses dados podem ser utilizados os métodos PARAFAC, MCR – ALS⁴⁹, PARALIND, dentre outros⁷⁰.

Concentração de quinino (mg/L)	Intensidade de Fluorescência
0,02	105,19
0,05	190,03
0,08	305,37
0,12	432,87
0,18	609,64
0,20	665,89



(a)

(b)



(c)

Figura 12. Representação esquemática dos dados que são utilizados em cada uma das calibrações: (a) ordem zero; (b) primeira ordem; (c) segunda ordem.

4.1.2.1. MCR-ALS com restrição de correlação

Essa restrição no MCR-ALS tem como objetivo a criação de modelos de calibração baseados em padrões com concentrações conhecidas. Em cada etapa iterativa os escores calculados pelo MCR são regredidos contra os valores de concentração indicados previamente. A Figura 13 representa esquematicamente o funcionamento dessa restrição tendo como base uma matriz D que é formada pelos conjuntos de calibração e validação. O *toolbox* MCR – ALS GUI 2.0⁴⁸ apresentou recentemente essa implementação da restrição de correlação.

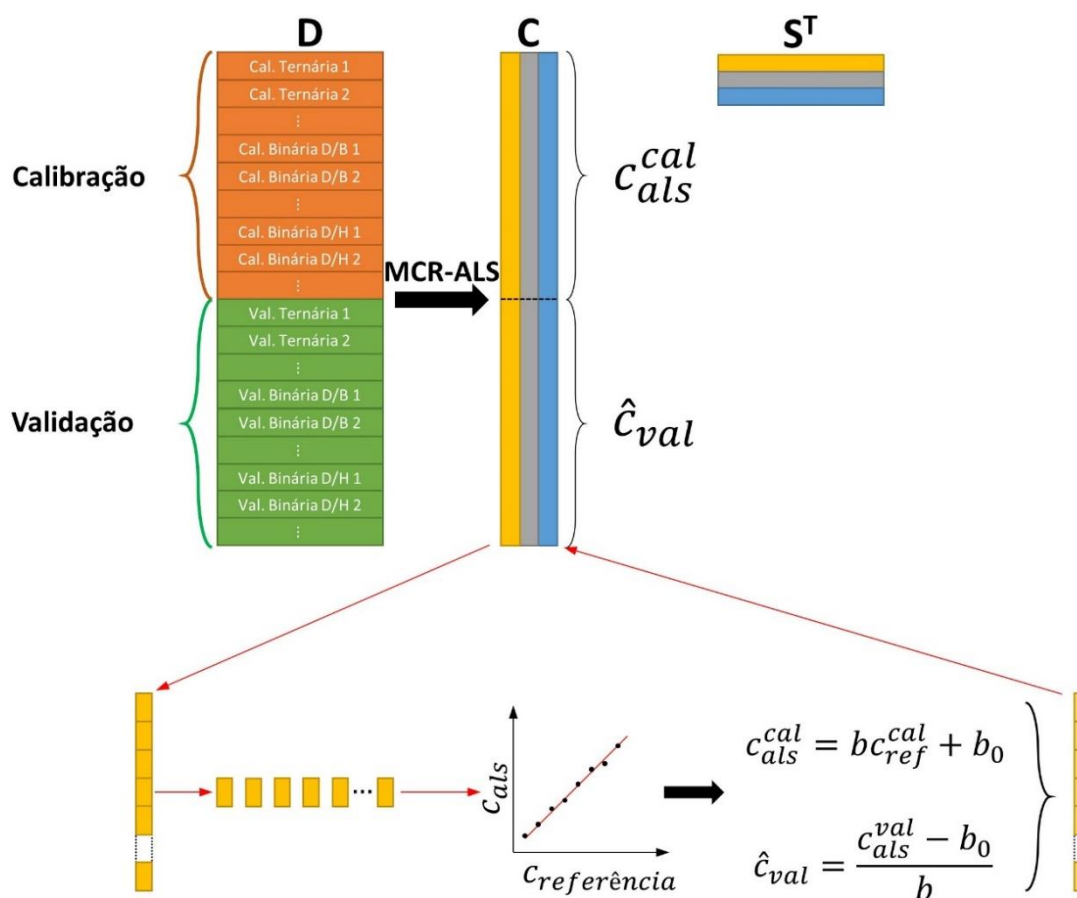


Figura 13. Esquema do funcionamento da restrição de correlação que é feito pelo *toolbox* MCR – ALS GUI 2.0.

O MCR – ALS pode ser utilizado em dados de calibração de primeira ordem^{71,72}, sendo que nesses casos pode ser observada a vantagem de segunda ordem de calibração. Essa vantagem, como dito anteriormente, possibilita a análise de dados de previsão que estejam contaminados com uma substância desconhecida ao modelo de calibração. Entretanto, Goicoechea *et al*⁷¹ propuseram que nem sempre

será possível a obtenção de resultados satisfatórios, como no caso em que se deseja calibrar diversos componentes ao mesmo tempo.

4.1.2.2. Avaliação dos modelos de calibração

A eficiência dos modelos de calibração multivariada pode ser avaliada pelo cálculo dos valores da raiz quadrada do erro quadrático médio (RMSE – *root mean square error*). Os valores de RMSECV (raiz quadrada do erro quadrático médio da calibração cruzada) são utilizados para definição do número de fatores a ser utilizados no modelo.

Tais valores expressam a proximidade entre o valor calculado pelo modelo (y_{prev}) e o valor verdadeiro ou obtido por um método de referência (y_{real}) com relação ao número de amostras (n), ou seja, eles expressão a exatidão do modelo (Equação 16).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_{prev} - y_{real})^2}{n}} \quad \text{Equação 16}$$

Esses erros podem ser calculados a partir da auto predição das amostras de calibração (raiz quadrada do erro quadrático médio de calibração - RMSEC) e da predição de amostras externas (raiz quadrada do erro quadrático médio de previsão - RMSEP).

4.2. Objetivos

- Utilizar a espectroscopia Raman para analisar misturas binárias e ternárias de diesel, biodiesel e HEFA;
- Desenvolver modelos de calibração para os três combustíveis utilizando a restrição de correlação do MCR – ALS;
- Avaliar a vantagem de segunda ordem de calibração do MCR – ALS frente à presença de adulterantes desconhecidos às amostras de calibração.

4.3. Procedimento Experimental

4.3.1. Preparo de Amostra

Foram preparadas 24 soluções (cada uma com 20,00 mL) para a construção dos modelos de calibração, sendo, 10 misturas ternárias de diesel/biodiesel/HEFA, cinco misturas binárias de diesel/biodiesel, cinco misturas

binárias de diesel/HEFA mais os três combustíveis puros. Já o conjunto de validação foi composto de 14 misturas, sendo quatro ternárias e 10 binárias (cinco de diesel/biodiesel e cinco de diesel/HEFA). A Tabela 1 apresenta o volume de cada combustível nessas amostras. Cada modelo está relacionado com a calibração de diesel (Diesel S10A fornecido pela Petrobrás S.A.), biodiesel (Biodiesel de Soja B-100 fornecido pela Petrobrás S.A.) e HEFA (NExBTL diesel fornecido pela Neste Oil - Finlândia) nas amostras. Com base nesse conjunto de amostra é possível realizar a calibração.

Tabela 1. Porcentagem dos volumes de diesel, biodiesel e HEFA adicionados em cada uma das amostras, sendo que o volume final de 20,00 mL.

Amostra	Calibração			Amostra	Validação		
	Diesel	Biodiesel	HEFA		Diesel	Biodiesel	HEFA
	% (v/v)	% (v/v)	% (v/v)		% (v/v)	% (v/v)	% (v/v)
1	73,50	25,50	1,00	1	71,50	24,00	4,50
2	73,00	25,00	2,00	2	70,50	23,00	6,50
3	72,50	24,50	3,00	3	69,50	22,00	8,50
4	72,00	24,00	4,00	4	68,50	21,00	10,50
5	71,50	23,50	5,00	5	95,00	5,00	-
6	71,00	23,00	6,00	6	85,00	15,00	-
7	70,50	22,50	7,00	7	75,00	25,00	-
8	70,00	22,00	8,00	8	65,00	35,00	-
9	69,50	21,50	9,00	9	55,00	45,00	-
10	69,00	21,00	10,00	10	95,00	-	5,00
11	100,00	-	-	11	85,00	-	15,00
12	90,00	10,00	-	12	75,00	-	25,00
13	80,00	20,00	-	13	65,00	-	35,00
14	70,00	30,00	-	14	55,00	-	45,00
15	60,00	40,00	-				
16	50,00	50,00	-				
17	-	100,00	-				
18	100,00	-	-				
19	90,00	-	10,00				
20	80,00	-	20,00				
21	70,00	-	30,00				
22	60,00	-	40,00				
23	50,00	-	50,00				
24	-	-	100,00				

A segunda parte desse trabalho envolveu a análise de 15 amostras adulteradas de diesel B-10 (10 amostras binárias, sendo cinco de diesel/biodiesel e cinco de diesel/HEFA). Foram adicionados a essas misturas seis adulterantes (etanol, cicloexano, gasolina, mistura de gasolina/etanol na proporção 75/25, hexano e querosene) que podem ser encontrados em amostras de diesel. Para a calibração foram utilizadas as amostras de calibração descritas anteriormente.

A adulteração foi realizada com sucessivas adições de 0,50 mL de adulterante na amostra, após cada adição foi medido o espectro da amostra. O volume inicial era de 20,00 e após as quatro adições o volume final foi de 22,00 mL. A partir do volume de amostra foi calculada a porcentagem volume/volume de cada uma das substâncias. A Tabela 2 apresenta a porcentagem dos volumes de cada composto nas amostras. Todas as amostras foram preparadas e guardadas em frascos de vidro transparente.

Tabela 2. Porcentagem dos volumes de diesel, biodiesel, HEFA e adulterante adicionados em cada uma das amostras, sendo que o volume final de 22,00 mL.

Amostra	Diesel	Biodiesel	HEFA	Adulterante
	% (v/v)	% (v/v)	% (v/v)	% (v/v)
1	90,00	10,00	-	0,00
2	87,80	9,76	-	2,44
3	85,71	9,52	-	4,76
4	83,72	9,30	-	6,98
5	81,82	9,09	-	9,09
6	90,00	-	10,00	0,00
7	87,80	-	9,76	2,44
8	85,71	-	9,52	4,76
9	83,72	-	9,30	6,98
10	81,82	-	9,09	9,09
11	90,00	5,00	5,00	0,00
12	87,80	4,88	4,88	2,44
13	85,71	4,76	4,76	4,76
14	83,72	4,65	4,65	6,98
15	81,82	4,55	4,55	9,09

4.3.2. Equipamento Utilizado

Os espectros Raman foram obtidos utilizando o espectrômetro dispersivo Raman Station 400F (PerkinElmer, Waltham, MA, EUA). Foi utilizado um laser de diodo que emite luz em 785 nm, potência máxima de 250 mW na fonte. A detecção foi feita com a CCD resfriada a -50°C por resfriamento termoelétrico (sistema Peltier). As amostras foram expostas ao laser por cinco segundos e foram realizadas 32 varreduras. Os espectros foram obtidos na faixa de 1800 a 200 cm^{-1} , com resolução de 1 cm^{-1} . As medidas foram realizadas diretamente no frasco em que as amostras foram guardadas.

4.3.3. Tratamento de Dados

O primeiro pré-processamento dos dados foi a remoção de sinais referentes à radiação espúria (*spikes*). Posteriormente, foi necessária a correção da linha de base dos espectros, para essa correção foi utilizada a primeira derivada e normalização por vetor unitário. Por último foi realizado o alisamento dos dados, em virtude da alta resolução espectral escolhida. Todos os pré-processamentos foram realizados com PLS Toolbox 7.3 (Eigenvector Research Inc., Wenatchee, WA, USA) em ambiente Matlab 2016a (MathWorks, Natick, MA, USA).

Após a etapa de pré-processamento foi possível o uso do MCR – ALS GUI 2.0⁴⁸ para a construção dos modelos de calibração. Antes da construção dos modelos foi preciso determinar o número de componentes químicos presentes nas amostras. Para as amostras de calibração e validação a SVD resultou na presença de três componentes. No caso das amostras adulteradas a SVD resultou na presença de quatro componentes. É importante ressaltar que as amostras adulteradas com a mistura de gasolina/etanol também resultaram na presença de quatro componentes químicos. Para esse conjunto de dados esperava-se a presença de cinco componentes, mas a SVD não foi capaz de determinar diferenças significativas para a presença de cinco componentes. Uma possível explicação para essa discrepância está no conjunto de dados utilizados, pois na matriz de dados não foram adicionados os espectros puros referentes à gasolina e etanol.

A etapa seguinte à determinação do número de componentes foi a estimativa da matriz S , pois na maioria dos casos esses espectros são conhecidos.

No caso das amostras de calibração e validação foram indicados quais eram as derivadas dos espectros referentes aos combustíveis puros.

Já para as amostras adulteradas foram estimadas quais eram as derivadas espectrais com as variáveis mais puras, ou seja, foram selecionadas as amostras com maior diferença significativa entre suas variáveis. Com base nas amostras selecionadas, foi observado se entre elas estavam as amostras referentes aos combustíveis puros. A necessidade dessa comparação era para identificar quais componentes estavam relacionados com os combustíveis e qual era o componente do adulterante. Caso as amostras selecionadas não contivessem os três combustíveis puros os seus espectros foram inseridos na matriz **S** e somente foi mantido o espectro correspondente à amostra adulterada. Em todos os casos foi observado que a amostra adulterada que havia sido selecionada era a que possuía a maior concentração de adulterante.

Na última etapa para a construção dos modelos de calibração são indicadas as restrições que o algoritmo do MCR – ALS irá utilizar para que seus resultados possuam algum sentido químico. Foram utilizadas as restrições de não-negatividade e correlação para o perfil de concentração das amostras. A restrição de correlação possibilitou a criação dos modelos de calibração, assim como demonstrado na seção 5.1.2.1. Não houve a necessidade de utilizar restrições para o perfil espectral das amostras.

Ao final da etapa iterativa do MCR – ALS são apresentados os valores de RMSEC, LOF e R^2 . Com os resultados dos escores do MCR – ALS referentes às amostras de validação foram calculados os valores de RMSEP. Esses parâmetros foram utilizados para a avaliar os modelos de calibração para o sistema proposto.

Já os espectros recuperados pelo MCR – ALS foram comparados com os espectros de referência. Para cada uma dessas comparações foi calculado o valor da correlação entre os dois espectros para que fosse possível afirmar que o MCR – ALS recuperou corretamente todos os perfis espectrais.

4.4. Resultados e Discussão

4.4.1. Modelos de Calibração e Amostras sem Adulteração

A Figura 14 apresenta os espectros Raman dos três combustíveis utilizados para construção dos modelos de calibração. A partir dos espectros foram observadas

algumas diferenças na faixa espectral analisada. Essas diferenças espectrais auxiliam na criação de modelos de calibração mais exatos para quantificação de cada um dos compostos presentes. A Tabela 3 apresenta os modos vibracionais mais importantes para esses diesel e biodiesel. O HEFA apresentou um espectro similar ao do diesel.

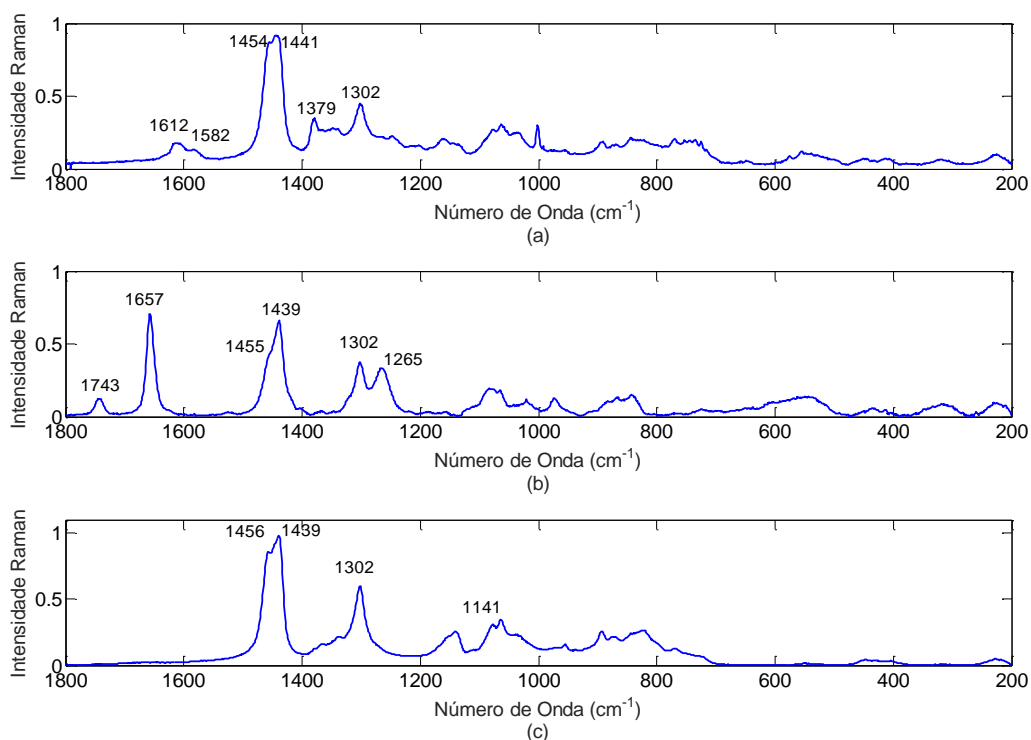


Figura 14. Espectros Raman dos três combustíveis puros: (a) diesel; (b) biodiesel; (c) HEFA.

Tabela 3. Atribuição de bandas do espectro Raman para diesel e biodiesel.

Composto	Número de Onda (cm ⁻¹)	Atribuição ^{27,28}
Diesel	1454	Deformação assimétrica - CH ₂ /CH ₃
	1348	Deformação simétrica dos grupos CH ₂ /CH ₃
	1063	Estiramento do esqueleto C-C do grupo -(CH ₂) _n -
Biodiesel	1743	Estiramento C = O
	1657	Estiramento C = C
	1455	Deformação assimétrica - CH ₂ /CH ₃
	1439	Deformação simétrica - CH ₂ /CH ₃
	1302	Deformação do grupo CH ₂
	1265	Deformação do grupo =C-H
	1200 - 1000	Estiramento do esqueleto C-C do grupo (-CH ₂) _n -
900 - 800	Estiramento C ₁ -C ₂ [*] , CH ₃ <i>rocking</i> , estiramento C-O	

* 1 e 2 em C₁-C₂ representa o primeiro e segundo carbono na estrutura molecular do éster próximo ao grupo C-C(=O).

As diferenças espectrais observadas não foram suficientes para que o MCR – ALS separa-se corretamente os três combustíveis. Os espectros do diesel e do HEFA são mais semelhantes, com isso não foi possível a correta identificação dos mesmos, por isso, foi proposto o uso da primeira derivada como pré-processamento para maximizar a diferença entre esses espectros. A Figura 15 apresenta os espectros referentes às amostras calibração e validação antes deles serem derivados. A Figura 16 apresenta o mesmo conjunto de dados só que com a primeira derivada dos espectros.

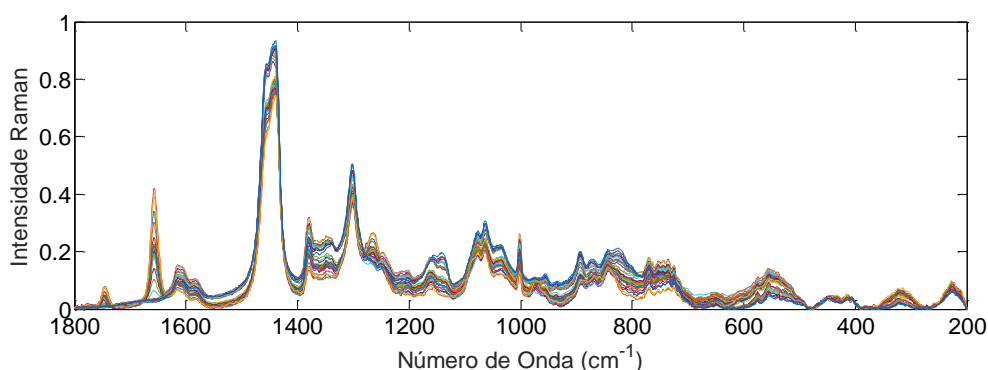


Figura 15. Espectros Raman obtidos experimentalmente para os conjuntos de calibração e validação.

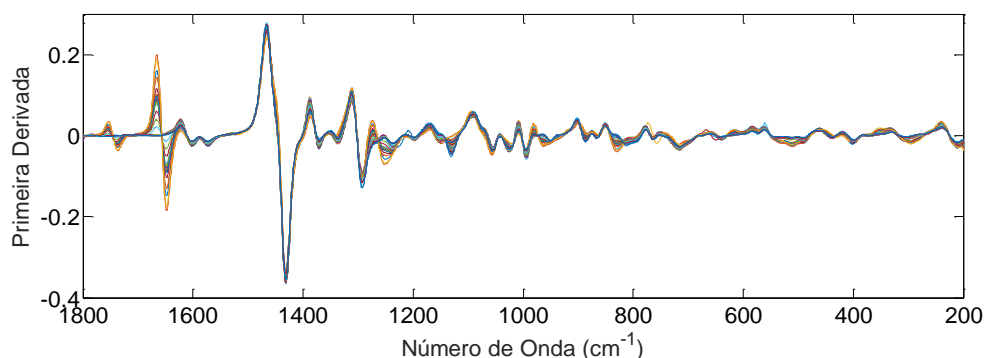


Figura 16. Primeira derivada dos espectros apresentados na Figura 15.

A partir das derivadas espectrais apresentadas na Figura 16 e dos valores de concentração das amostras de calibração foi possível a construção dos modelos de calibração pelo MCR – ALS com restrição de correlação. Os escores do MCR – ALS puderam ser relacionados diretamente com as concentrações das amostras, sendo que a partir de uma regressão linear é possível determinar a equação que os relaciona. Na Figura 17 estão apresentados os três modelos de calibração obtidos pelo MCR – ALS, as amostras de calibração e validação estão representadas em azul

e preto, respectivamente. A reta em vermelho representa o ajuste linear para as amostras de calibração.

Algumas das amostras de validação do diesel (quadrados em preto da Figura 17a) apresentaram valores de escores muito diferentes que os valores de referência. O mesmo comportamento não foi observado nas curvas de calibração de biodiesel e HEFA.

A avaliação dos modelos não é realizada somente com a observação do ajuste linear dos resultados. A Tabela 4 apresenta os resultados de avaliação dos modelos, sendo que ao final do processo iterativo, o MCR – ALS já calcula os valores de RMSEC, R^2 e LOF. A partir dos escores para as amostras de validação foram calculados os valores de RMSEP, Equação 16.

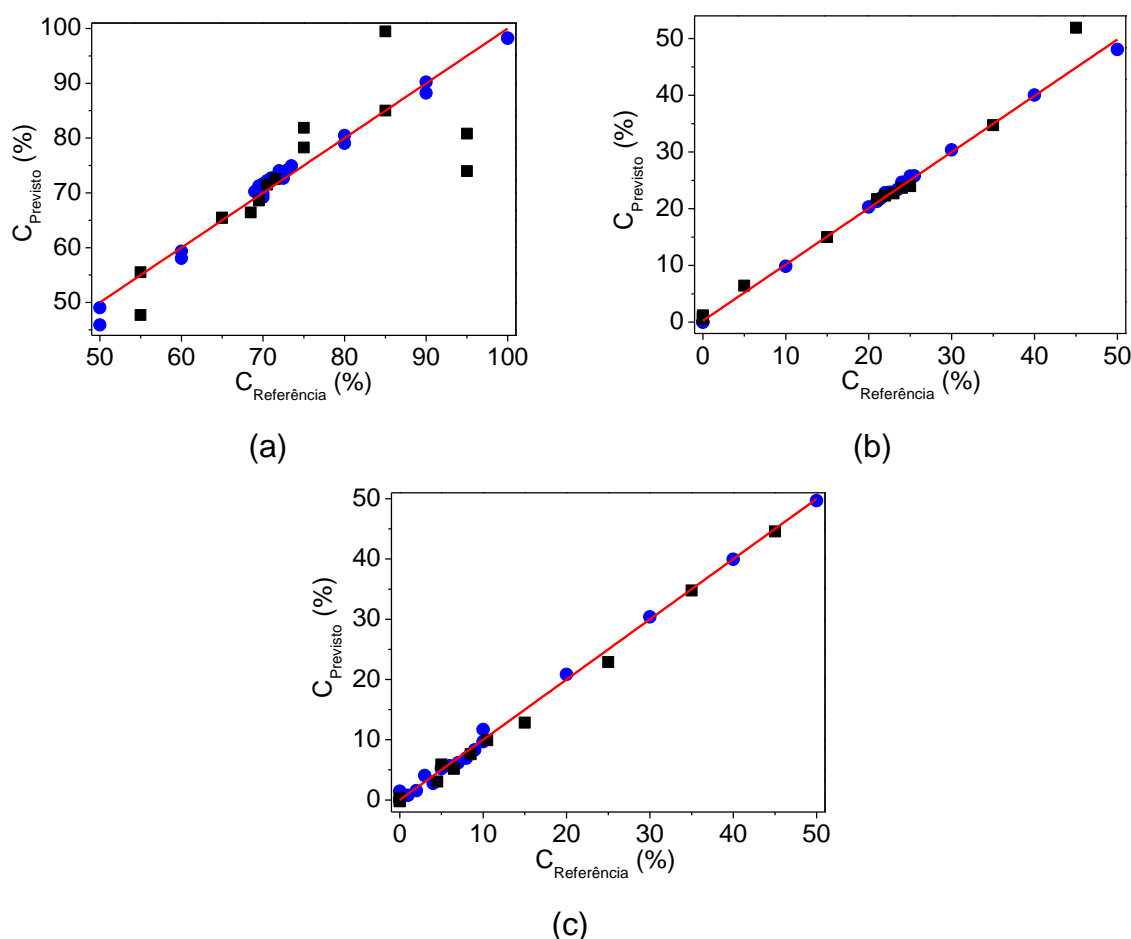


Figura 17. Curvas de calibração para os combustíveis: (a) diesel; (b) biodiesel; (c) HEFA. Amostras de calibração e validação estão representadas em azul e preto, respectivamente. A reta em vermelho é o ajuste linear para as amostras de calibração.

Tabela 4. Valores de R^2 , LOF, RMSEC e RMSEP para os modelos propostos.

	Diesel	Biodiesel	HEFA
R^2	0,9907	0,9973	0,9973
LOF	0,0355	0,0355	0,0355
RMSEC (%)	1,76	0,78	1,06
RMSEP (%)	3,12	0,70	0,70

Como apresentado na Tabela 4, os valores de R^2 foram superiores a 0,99 e o LOF foi igual a 0,0355. Esse resultado confirma a adequação do modelo de calibração do MCR – ALS para o conjunto de dados proposto. Para a obtenção do LOF o MCR – ALS utiliza a Equação 5, esse baixo valor indica uma boa recuperação das matrizes C e S^T .

Ainda com relação aos resultados da Tabela 4, foi possível avaliar os baixos valores de RMSEC e RMSEP. Com exceção do diesel, os resultados de RMSEP para biodiesel e HEFA foram iguais a 0,70%, valor muito próximo do sugerido pelas agências reguladoras que é de 0,50%^{61,73–75}. Essa proximidade indica uma possível aplicação da espectroscopia Raman e do MCR – ALS para análise de combustíveis adulterados.

Além da avaliação da matriz C é necessário observar os resultados da matriz S . Essa matriz traz as informações referentes à recuperação das derivadas dos espectros de cada combustível. Caso haja muita discrepância na comparação entre os resultados obtidos pelo MCR – ALS e as derivadas espectrais de referência os resultados de concentração não poderão ser utilizados para a construção de modelos de calibração. A Figura 18 apresenta uma comparação entre as derivadas dos espectros obtidos pelo MCR – ALS (em vermelho) e dos combustíveis puros (em azul). Essa comparação possibilita o cálculo da correlação entre esses espectros; todos os valores obtidos foram maiores que 0,9930. Com base nesses valores de correlação pode-se afirmar que o MCR – ALS foi capaz de identificar os três combustíveis presentes nas amostras.

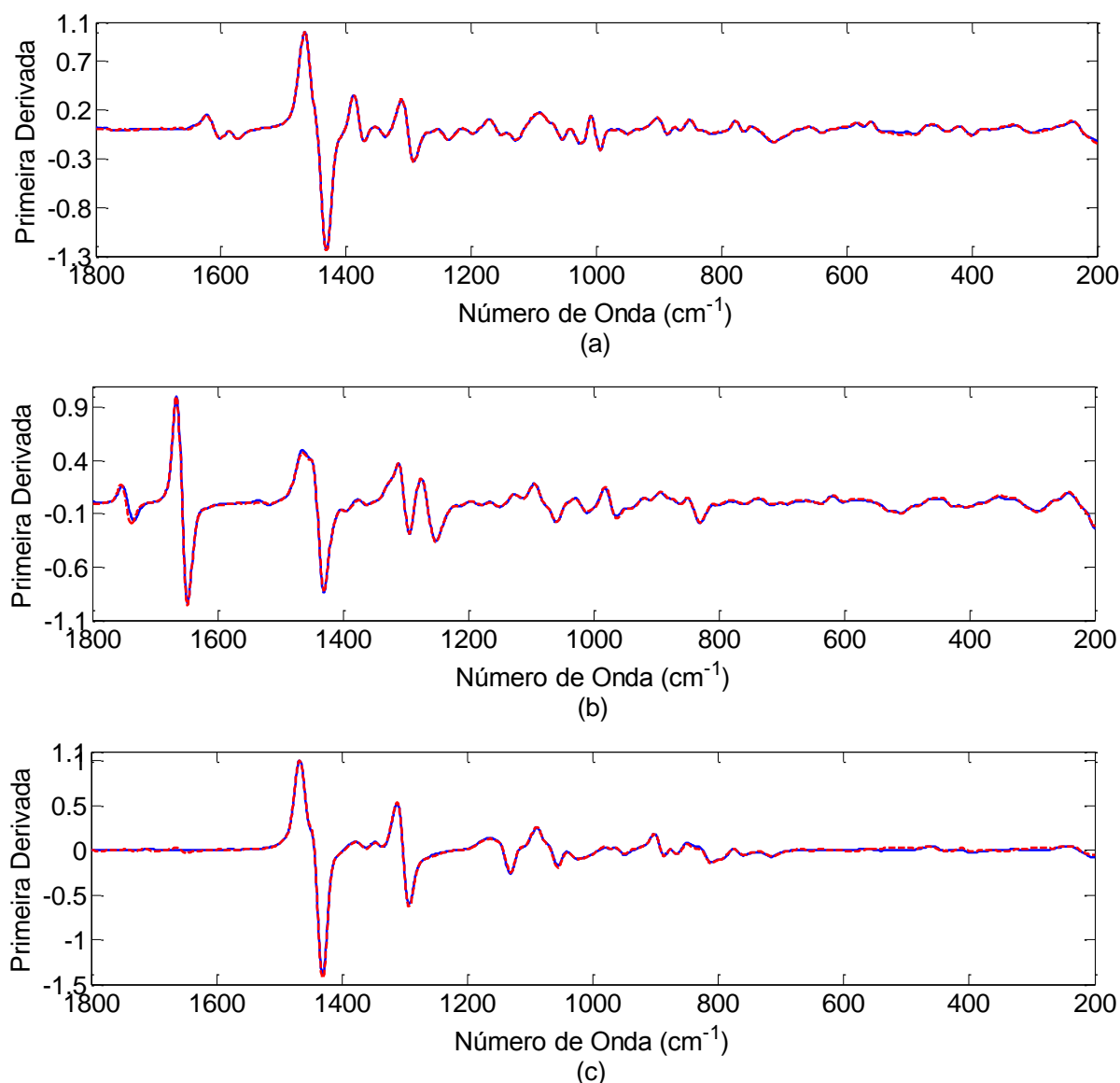


Figura 18. Comparação entre as derivadas dos espectros dos combustíveis puros (em azul) e dos recuperados pelo MCR (em vermelho), (a) diesel – correlação de 0,9993, (b) biodiesel – correlação de 0,9933 e (c) HEFA – correlação de 0,9977.

Portanto, os resultados do MCR – ALS indicaram que foi possível a identificação de todos os combustíveis presentes nas amostras e, além disso, foi possível a construção de seus modelos de calibração. Esses resultados possibilitaram a utilização de amostras adulteradas afim de identificar quais seriam os adulterantes e, se possível, uma estimativa de sua concentração.

4.4.2. Amostras com Adulteração

Foram propostas a construção de dois tipos de matrizes para a otimização do MCR – ALS, a primeira continha as 24 amostras de calibração e as 5 amostras de

diesel B10 adulterado (Figura 19a). O segundo conjunto foi proposto com a utilização das 24 amostras de calibração e as 15 amostras de diesel B10 adulterado (Figura 19b), amostras binárias e ternárias com o mesmo adulterante.

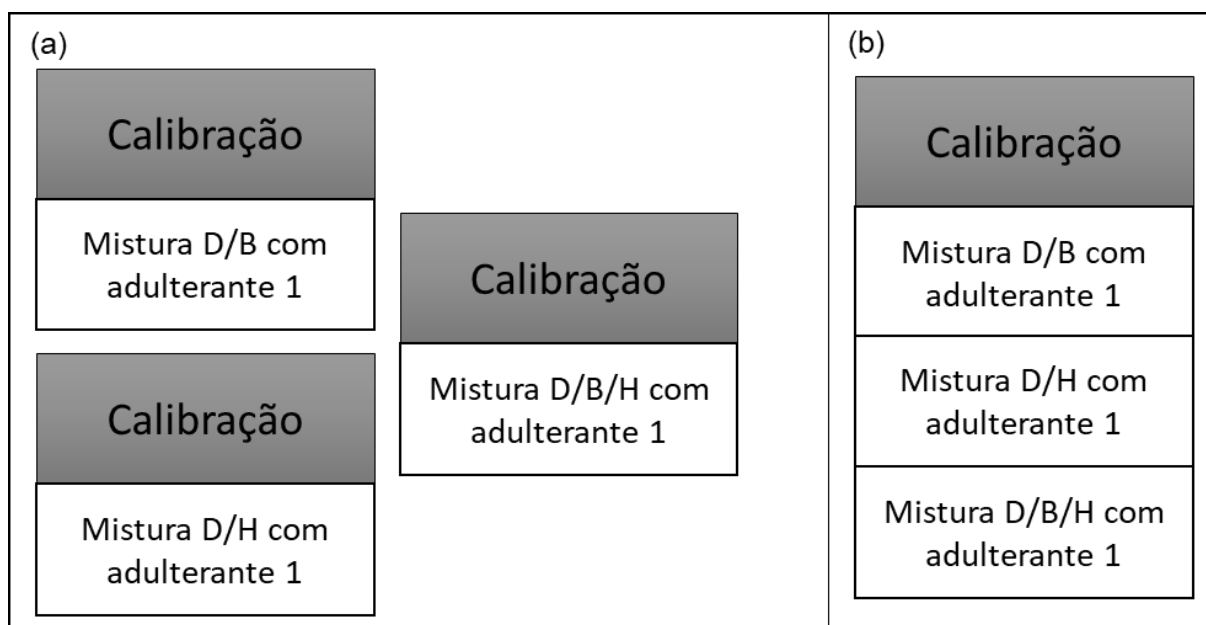


Figura 19. Representação esquemática dos dois conjuntos de dados propostos para a identificação de adulterações em amostras de diesel B10: (a) matrizes com somente um adulterante; (b) matriz aumentada com todas as amostras adulteradas. D = diesel, B = biodiesel e H = HEFA.

Levando em conta o conjunto de dados proposto, é possível que se obtenha a vantagem de segunda ordem pois a quantificação será possível na presença de um adulterante que não estava na curva de calibração. Nesse caso, além da previsão da concentração dos combustíveis, poderá ser possível a identificação do adulterante. A Tabela 5 apresenta os resultados de RMSEP e de correlação entre as derivadas dos espectros recuperados e os de referência.

Tabela 5. Resultados de RMSEP e correlação espectral para as misturas binárias e ternárias das amostras adulteradas. Também estão especificados os valores de correlação espectral entre os espectros puros e recuperados pelo MCR.

Combustíveis e Adulterantes	Mistura de D/B		Mistura de D/H		Mistura de D/B/H		Matriz Aumentada	
	RMSEP (%)	Corr.	RMSEP (%)	Corr.	RMSEP (%)	Corr.	RMSEP (%)	Corr.
Diesel	6,01	0,9996	5,59	0,9996	4,79	0,9996	5,66	0,9994
Biodiesel	2,20	0,9973	0,00	0,9971	2,02	0,9973	1,76	0,9969
HEFA	0,40	0,9985	2,24	0,9985	1,96	0,9985	1,70	0,9980
Cicloexano	8,78	0,9679	8,23	0,9745	8,43	0,9717	8,34	0,9829
Diesel	3,82	0,9996	5,10	0,9996	6,28	0,9996	5,11	0,9993
Biodiesel	1,26	0,9972	0,00	0,9969	1,32	0,9971	1,13	0,9967
HEFA	0,76	0,9984	3,55	0,9985	3,34	0,9984	2,52	0,9978
Etanol	7,68	0,8660	7,67	0,8830	7,79	0,8648	7,80	0,9167
Diesel	2,76	0,9998	2,78	0,9997	2,40	0,9998	2,46	0,9998
Biodiesel	1,45	0,9969	0,00	0,9967	1,56	0,9969	1,25	0,9964
HEFA	0,00	0,9989	1,74	0,9986	2,27	0,9988	1,96	0,9989
Gasolina	5,19	0,8414	3,89	0,8332	4,23	0,8345	4,28	0,8555
Diesel	7,29	0,9997	8,20	0,9997	9,73	0,9997	8,54	0,9995
Biodiesel	0,40	0,9941	0,00	0,9943	0,20	0,9942	0,27	0,9938
HEFA	0,96	0,9986	1,94	0,9985	2,56	0,9986	1,78	0,9983
Hexano	2,26	0,7890	3,07	0,8604	1,64	0,8486	2,13	0,8906
Diesel	4,39	0,9997	5,33	0,9996	3,69	0,9997	4,73	0,9997
Biodiesel	0,87	0,9970	0,00	0,9973	1,04	0,9970	0,73	0,9968
HEFA	0,45	0,9987	0,45	0,9984	0,87	0,9986	0,67	0,9984
Querosene	6,39	0,8781	5,93	0,9109	5,69	0,8925	6,21	0,9108
Diesel	1,69	0,9998	7,28	0,9996	3,35	0,9997	4,30	0,9997
Biodiesel	0,97	0,9969	0,00	0,9966	0,94	0,9970	1,05	0,9965
HEFA	0,22	0,9989	5,20	0,9980	1,97	0,9987	2,75	0,9972
Mistura de G/E	2,82	0,7664	3,32	0,6088	3,14	0,7741	2,80	0,7717

Para as amostras binárias de diesel/biodiesel os valores de RMSEP dos biocombustíveis variaram de 0,40 a 2,20% para biodiesel e de 0,00 a 0,96% para HEFA. Já as amostras binárias de diesel/HEFA esses valores foram de 0,00 para biodiesel e de 0,45 a 5,20% para HEFA. As amostras ternárias apresentaram valores de RMSEP de 0,20 a 2,02% para biodiesel e 0,87 a 3,34% para HEFA. O diesel apresentou valores de RMSEP na faixa de 1,69 a 9,73%, sendo que esses valores foram maiores do que os obtidos na etapa de validação, mas ainda são úteis para aplicações práticas.

Esse aumento nos valores de RMSEP estão relacionados com a presença dos adulterantes, pois eles perturbam a decomposição do MCR – ALS. Caso fossem criados modelos de calibração na presença de adulterantes essa alteração não seria observada; entretanto, não seria possível a utilização da “vantagem de segunda ordem” dos métodos multivariados.

Como representado pela Tabela 5 alguns valores de RMSEP foram iguais a zero. A Equação 16 demonstra como esses valores são obtidos; nela é possível verificar que os escores do MCR – ALS e as concentrações de referência são iguais. Concomitantemente, foram obtidos os valores de correlação espectral para a matriz *S*. As correlações encontradas foram superiores a 0,9940, indicando que o MCR – ALS foi capaz de identificar corretamente os três combustíveis na presença de adulterantes.

Os adulterantes, por sua vez, apresentaram resultados de correlação mais divergentes do que os combustíveis. Não houve grandes mudanças entre os valores de RMSEP calculados entre os conjuntos de dados com mesmo adulterante, mas houve uma grande variação na comparação desses valores para adulterantes diferentes. Essa diferença foi mais pronunciada para o caso das adulterações de cicloexano (8,78%) e hexano (2,26%) nas misturas binárias de diesel/biodiesel. As correlações espectrais também apresentaram resultados diversos, sendo que o pior dos cenários foi o da adulteração com a mistura de gasolina/etanol.

Por causa dessa discrepância foi proposto o uso da matriz aumentada (Figura 19b) utilizando todas as amostras adulteradas com um único adulterante. Essa nova matriz possui mais informações sobre os adulterantes; com isso espera-se que sejam obtidos resultados mais parecidos com os valores de referência. Entretanto, não houve mudanças significativas para os valores de correlação e RMSEP com o uso dessa matriz aumentada, como demonstrado na Tabela 5.

A comparação entre as derivadas dos espectros puros (em azul) e recuperados pelo MCR – ALS (em vermelho) pode ser vista na Figura 20. Os altos valores de correlação apresentados na Tabela 5 demonstraram que foi possível a identificação de todos os componentes químicos das amostras.

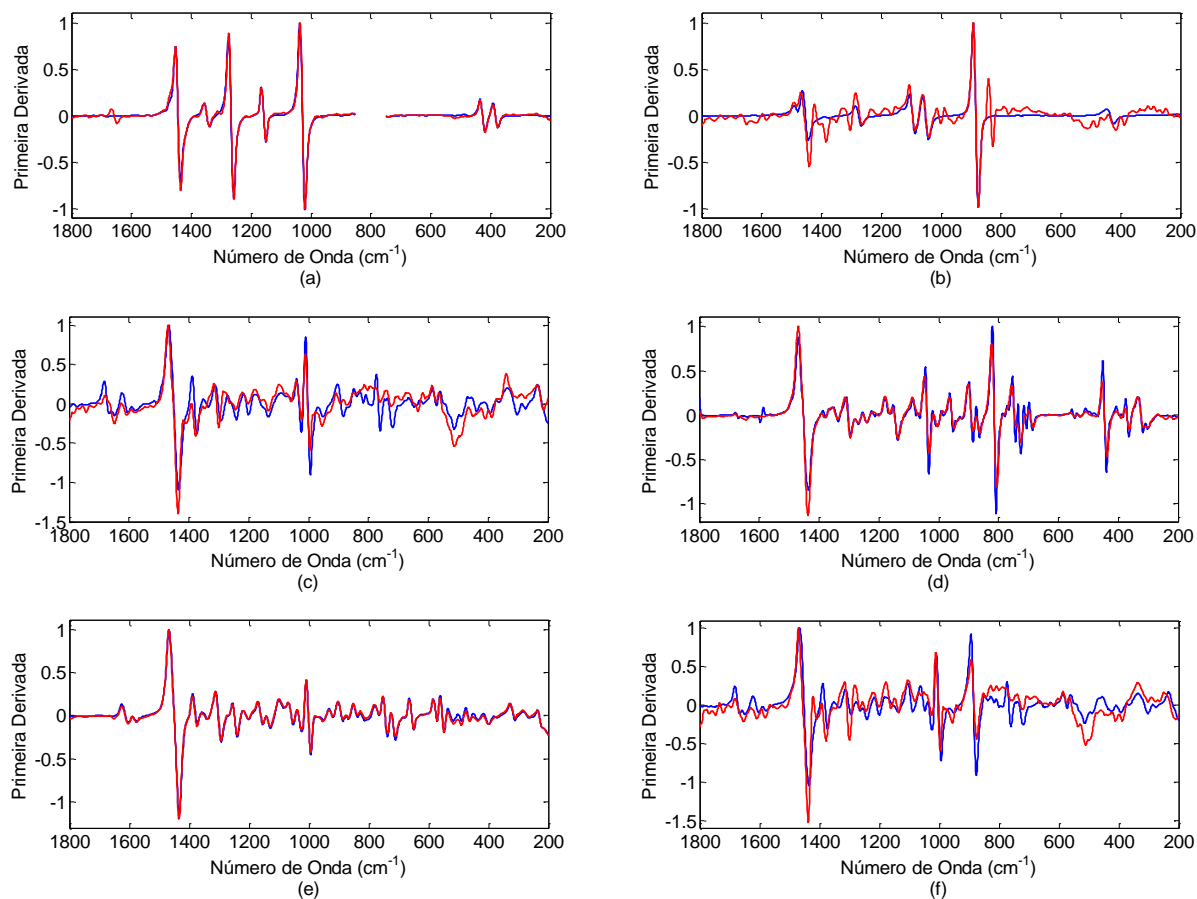


Figura 20. Comparação entre as derivadas dos espectros dos combustíveis puros (em azul) e recuperados pelo MCR (em vermelho): (a) cicloexano; (b) etanol; (c) gasolina; (d) hexano; (e) querosene; (f) mistura de gasolina/etanol.

Cicloexano apresenta uma banda muito intensa na região de 750 a 850 cm^{-1} . Essa banda está relacionada com o estiramento C-C⁷⁶ e a construção de modelos utilizando essa banda apresentaram altos valores de RMSE. A Figura 21 apresenta a primeira derivada do cicloexano com e sem essa região. Com isso foi proposta a remoção dessa banda e um novo modelo foi criado, tendo os seus resultados apresentados na Tabela 5. Já foi reportado na literatura⁷⁷ que a diferença de intensidade (ambiguidade de intensidade) entre bandas pode afetar os resultados das matrizes **C** e **S** do MCR – ALS, por isso optou-se pela remoção dessa região espectral.

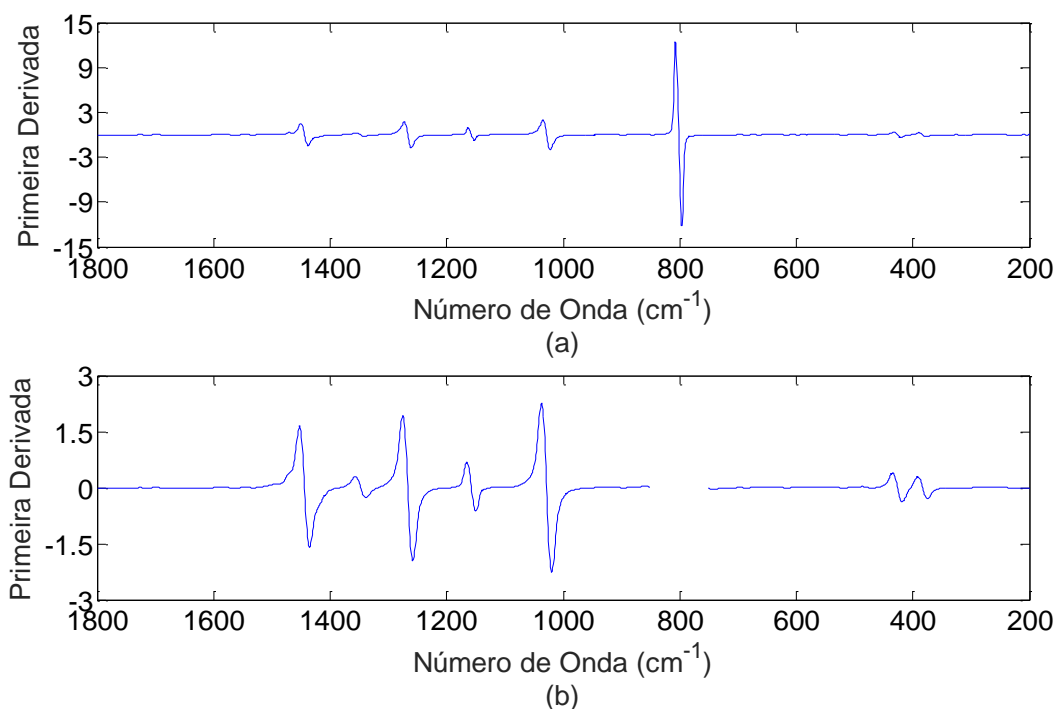


Figura 21. Derivada do espectro do cicloexano (a) antes e (b) depois da remoção da região de 750 a 850 cm⁻¹.

4.5. Conclusões Parciais

Este trabalho demonstrou a possibilidade da combinação da espectroscopia Raman com MCR – ALS com restrição de correlação para o desenvolvimento de uma metodologia analítica de identificação e quantificação de biocombustíveis e adulterantes presentes no diesel. Os resultados para a quantificação de biodiesel e HEFA apresentaram RMSEP de 0,70%, indicando a aplicação prática da metodologia proposta.

Os métodos padrão, como EM 17078 e ASTM D7371-07, indicam o uso da espectroscopia de infravermelho médio com métodos quimiométricos univariados e multivariados (PLS) para a determinação de biodiesel em diesel⁷⁸. Entretanto, para o uso do PLS é necessário que haja mais amostras calibração e validação do que o proposto neste trabalho⁷⁹.

Para amostras com adulterantes, foi possível identificar o composto utilizado para adulteração com alta confiança com base nos espectros recuperados e estimar sua concentração com erros abaixo de 8,34%. Nestas amostras, também foi possível quantificar os biocombustíveis com erros abaixo de 2,52%, além de identificar a ausência de determinado biocombustível, uma vez que o modelo poderia prever sua

concentração nula. Além disso, a vantagem de segunda ordem e o uso do MCR – ALS possibilitaram a correta identificação e quantificação de todos os adulterantes propostos nesse estudo.

O uso da matriz aumentada melhorou os resultados para quantificação de biocombustíveis, mas não para resolução de espectros e quantificação dos adulterantes. Essa melhoria pode estar aliada ao fato de que a matriz aumentada apresenta uma quantidade de informação adicional sobre os combustíveis, facilitando assim sua decomposição pelo MCR – ALS e melhorando os resultados de quantificação.

**CAPÍTULO 5 – CARACTERIZAÇÃO
DE AMOSTRAS DE COCAÍNA E
CRACK**

5.1. Introdução e Motivação

Drogas ilícitas, de modo geral, representam um grande problema social e de saúde pública. Cocaína (em pó ou como pasta base), maconha, LSD (dietilamida do ácido lisérgico), ecstasy (MDMA - metilenedioximetanfetamina), dentre outros⁸⁰⁻⁸³ são algumas das drogas mais difundidas. Algumas delas podem gerar dependência química com uso contínuo, como no caso da cocaína.

Além disso, visando maior lucro se tornou comum a “diluição” dessas drogas. Para a cocaína, que será abordada nesse trabalho, começaram a ser encontrados amido, anestésicos e até mesmo outros medicamentos misturados. Essas alterações podem causar efeitos adversos, como potencialização ou minimização do efeito da droga e baratear seu custo. Dentre os compostos químicos mais comumente utilizados podem ser citados a benzocaína, cafeína, dipirona, levamisol, lidocaína e teofilina⁸⁴.

Dentre esses adulterantes um dos mais controversos é o levamisol^{85,86}, que é um anti-helmíntico restrito ao uso veterinário. Ele deixou de ser utilizado em seres humanos devido a alguns efeitos colaterais no sistema imunológico. O Instituto de Criminalística de Campinas já havia evidenciado que há um aumento do uso desse fármaco em amostras de cocaína⁸⁴. Seu uso crônico causa agranulocitose, manchas pretas na pele, febre e feridas dolorosas na boca e no ânus. Ele é utilizado para potencializar o efeito da cocaína e fazer com que a mesma chegue mais rápido ao sistema nervoso central⁸³.

5.2. Objetivos

- Analisar amostras cocaína e crack apreendidas pela Polícia Militar de Campinas por espectroscopia de imagem Raman;
- Caracterização inicial das amostras utilizando a ICA com o algoritmo JADE;
- Criação de um “Limite de Identificação” dos compostos presentes nas amostras baseado nos valores dos escores do MCR – ALS;
- Realizar uma análise semi-quantitativa dos compostos químicos identificados em cada uma das amostras.

5.3. Procedimento Experimental

5.3.1. Amostras

Foram analisadas 22 amostras de cocaína e crack apreendidas pela Polícia Militar de Campinas. Nenhum tipo de preparo de amostra foi realizado antes das amostras serem analisadas neste trabalho. Todas as amostras foram cedidas pelo Núcleo de Perícias Criminais de Campinas do Estado de São Paulo (NPC-Campinas). Essas amostras foram analisadas (cromatografia em camada delgada) pelos peritos criminais e foi determinada a presença de seis compostos (benzocaína, cafeína, dipirona, levamisol, lidocaína e teofilina) além de cocaína ou crack em suas composições. O Anexo 1 apresenta a autorização do Diretor do NPC-Campinas para a análise das amostras de cocaína/crack utilizando a espectroscopia Raman.

Além das amostras de cocaína e crack, foram obtidos junto ao NPC-Campinas os compostos: benzocaína, cafeína, dipirona, levamisol, lidocaína e teofilina, utilizados como padrão nesse estudo.

5.3.2. Equipamento Utilizado

As amostras foram analisadas utilizando a espectroscopia de imagem Raman com o espectrômetro dispersivo Raman Station 400F (PerkinElmer, Waltham, MA, EUA). Foi utilizado um laser de diodo que emite luz em 785 nm, potência máxima de 250 mW na fonte. A detecção foi feita com uma CCD resfriada a -50°C por resfriamento termoelétrico (sistema Peltier).

Os espectros Raman foram obtidos com um segundo de exposição do laser e 16 varreduras por *pixel*. A área analisada foi de $5,76\text{ mm}^2$ ($2,4\text{ mm} \times 2,4\text{ mm}$) com tamanho do *pixel* $100\text{ }\mu\text{m} \times 100\text{ }\mu\text{m}$. Com isso foi obtida uma imagem de $25 \times 25\text{ pixels} = 625$ pontos da amostra analisados para cada amostra. Os espectros foram obtidos na faixa de 600 a 1800 cm^{-1} com resolução de 2 cm^{-1} . Os compostos puros tiveram uma área analisada de $0,36\text{ mm}^2$ ($0,6\text{ mm} \times 0,6\text{ mm}$), com *pixel* de $100\text{ }\mu\text{m} \times 100\text{ }\mu\text{m}$. A imagem obtida foi de $6 \times 6\text{ pixels} = 36$ pontos, a resolução espectral e número de varreduras foram os mesmos das amostras.

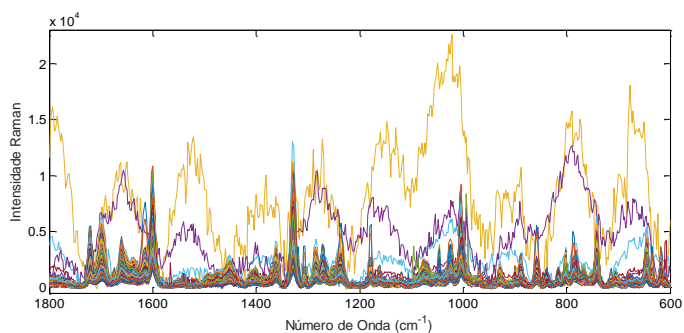
5.3.3. Tratamento de Dados

5.3.3.1. Identificação de Pixels Anômalos e Remoção de Spikes

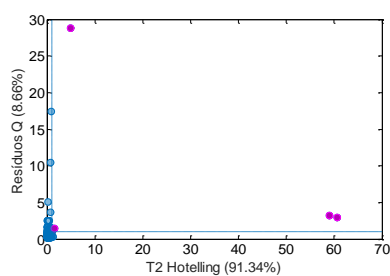
Antes de qualquer tipo de pré-processamento houve necessidade de identificar a presença de espectros (*pixels*) anômalos no conjunto de dados. Essa etapa é importante para que haja remoção de dados que possam vir a prejudicar os resultados dos mapas de concentração das amostras analisadas. A seguir estão descritas as etapas de identificação e substituição dos *pixels* anômalos:

- Realizar uma Análise de Componentes Principais (PCA) dos espectros da amostra (Figura 22a);
- Avaliar o gráfico de T^2 Hotelling contra Resíduos Q (Figura 22b);
- Selecionar as amostras que possuírem altos valores de T^2 Hotelling e Resíduos Q, simultaneamente (amostras em rosa na Figura 22b);
- Substituir os espectros selecionados na etapa anterior pelo espectro mediano dos seus vizinhos (Figura 22c); e
- Obtenção da nova matriz de dados com os novos espectros (Figura 22d).

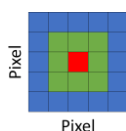
Ao final dessa análise foram identificados os espectros que possuíam *spikes* e os mesmos foram removidos do conjunto de dados. A Figura 22e apresenta o conjunto de dados obtidos após a remoção dos *spikes*. O Anexo 2 apresenta a rotina utilizada para a identificação e substituição desses *pixels* anômalos.



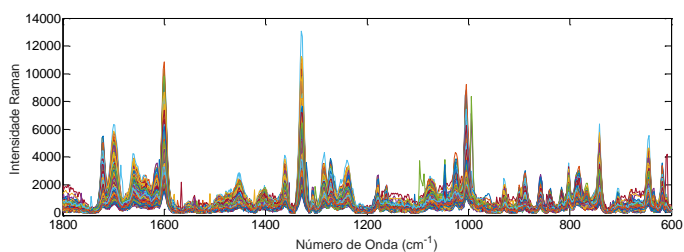
(a)



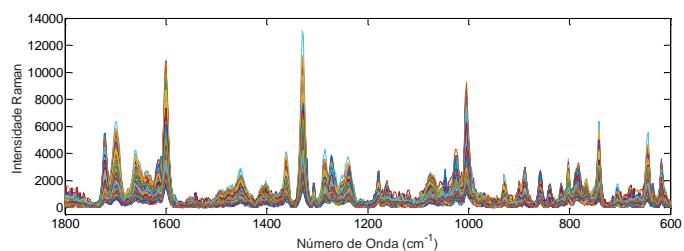
(b)



(c)



(d)



(e)

Figura 22. Identificação dos *pixels* anômalos e remoção de spikes: (a) espectros obtidos experimentalmente; (b) gráfico de T^2 Hotelling contra Resíduos Q; (c) *pixel* em vermelho é um *pixel* anômalo e terá seu espectro substituído pelos espectros dos *pixels* em verde; (d) conjunto de dados após remoção dos *pixels* anômalos; (e) conjunto de dados após remoção dos spikes.

5.3.3.2. *Simulação das Imagens Químicas*

Na impossibilidade de serem obtidas imagens químicas para os possíveis componentes químicos presentes nas amostras há a possibilidade de simular esse conjunto de dados. Essa impossibilidade pode estar ligada com a ausência do componente puro em laboratório ou mesmo algum problema com o equipamento utilizado nas análises anteriores. Para essas eventualidades é possível estimar o conjunto de dados do composto pretendido somente com base em seu espectro oriundo da utilização do processo de resolução de curvas. Neste estudo será apresentada a obtenção do conjunto de dados para a fenacetina (Figura 23a), um composto que não era esperado, mas que apareceu em algumas amostras.

Foi considerado como espectro de referência o sinal obtido pela otimização da ICA, para esse composto. Foi criada uma matriz de ruído aleatório com a mesma intensidade do ruído apresentado pelas amostras dos outros componentes puros (Figura 23b), e esse ruído foi adicionado ao sinal da ICA. No final, foi obtida uma matriz com o mesmo tamanho dos dados referentes aos outros compostos puros (imagem com 36 espectros, Figura 23c).

Além disso, será mostrado adiante que foi criada uma imagem referente ao deslocamento da linha de base dos espectros de cada uma das amostras de cocaína. Como a ICA é capaz de identificar diferenças físicas entre as ICs calculadas, também, foi proposta a criação da imagem dessa diferença de linha de base para que os dados a serem otimizados pelo MCR – ALS pudessem ser melhor interpretados e para que houvesse diminuição da ambiguidade rotacional do MCR – ALS. Essa imagem também foi composta de 36 espectros, sendo que o sinal que a deu origem foi um dos ICs resultantes da utilização do ICA.

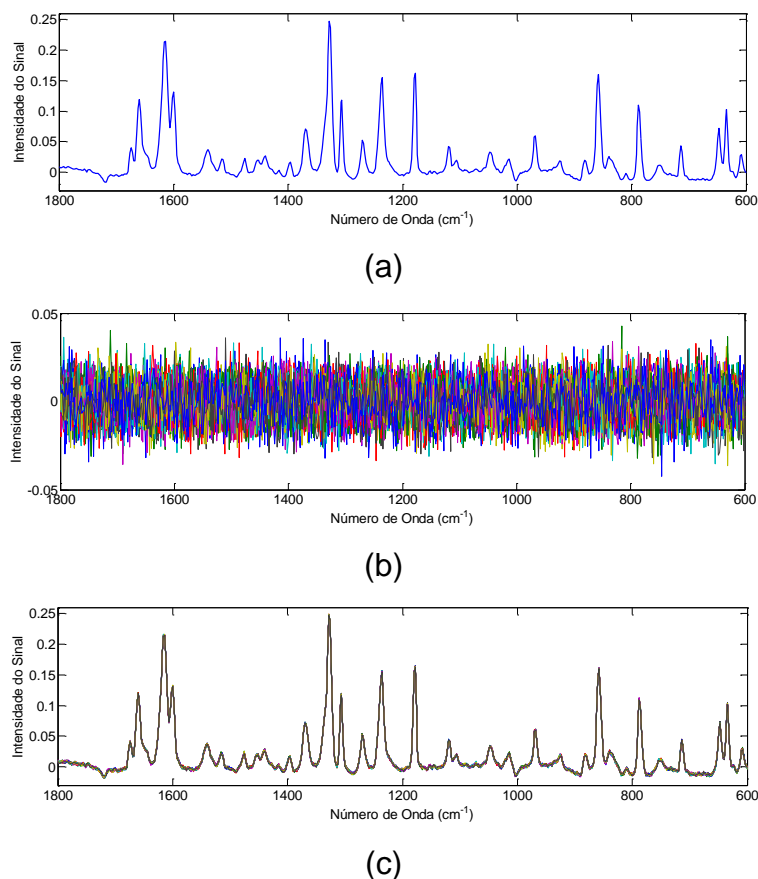


Figura 23. (a) Sinal da ICA referente à fenacetina; (b) conjunto de dados de ruído aleatório gerado para ser adicionado ao sinal da fenacetina; (c) espectros referentes à imagem Raman simulada da fenacetina.

5.3.3.3. Construção da Matriz de Dados

A Figura 24 representa de forma esquemática a construção da matriz de dados que foi analisada pelos métodos de resolução de curvas. O cubo em vermelho representa o conjunto de dados da amostra e os cubos em verde representam os dados dos compostos puros. Cada um desses cubos foi desdobrado e essas matrizes foram concatenadas como representado na Figura 24.

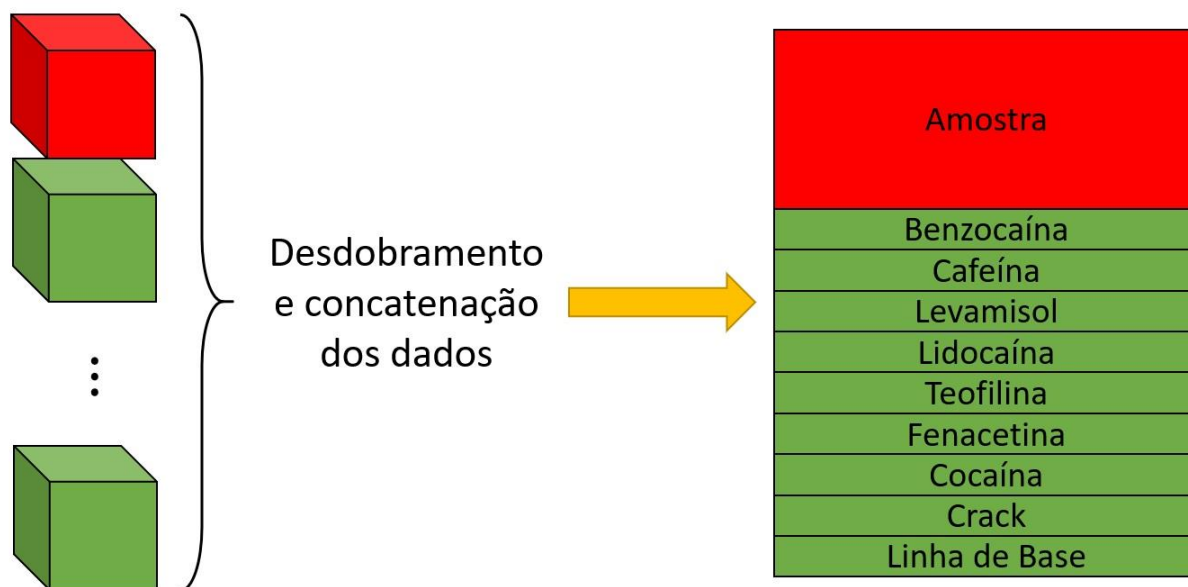


Figura 24. Desdobramento e concatenação dos dados das imagens hiperespectrais e construção da matriz de dados a partir da concatenação das matrizes da amostra e dos compostos puros.

5.3.3.4. Pré-processamento das Imagens Hiperespectrais

A remoção da radiação espúria (*spikes*) foi o primeiro pré-processamento utilizado, como apresentado pela Figura 22e na secção 6.3.3.1. Além disso, todos os espectros foram normalizados por vetor unitário para corrigir a variação de intensidade dos espectros. Por fim, foi realizado um alisamento com o algoritmo Savitzky-Golay com uma janela de 15 pontos e com um polinômio de segunda ordem. Todos os pré-processamentos foram realizados com PLS Toolbox 7.3 (Eigenvector Research, Inc., Wenatchee, WA, USA) em ambiente Matlab 2016a (MathWorks, Natick, MA, USA).

5.3.3.5. Determinação do Número de Componentes

A determinação do número de componentes é a etapa mais importante de ser realizada antes da otimização pelos métodos de resolução de curvas. Como dito previamente, essa escolha irá moldar os resultados obtidos. Se essa escolha for realizada de forma aleatória, os resultados obtidos não irão representar corretamente o conjunto de dados.

A ICA – JADE foi o primeiro método a ser utilizado para uma caracterização inicial das amostras; com isso foram utilizadas a ICA por Blocos (Figuras 25 e 26) e Critério Durbin-Watson (Figura 27) para determinar quantas fontes de sinais independentes (componentes) estariam presentes nas amostras.

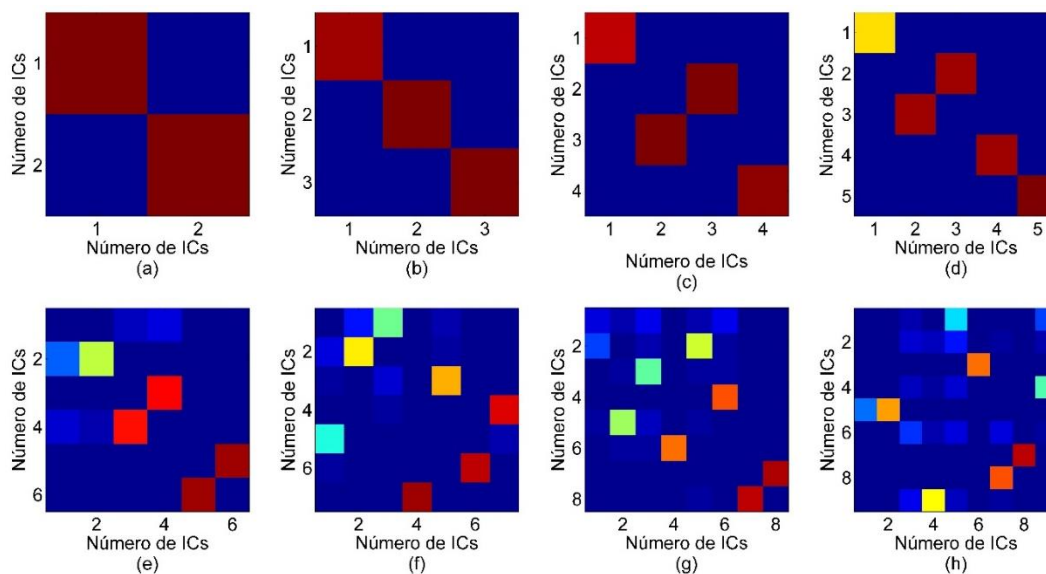


Figura 25. Resultado do ICA por Blocos para uma amostra (Amostra 1). As imagens de (a) até (h) simbolizam os resultados para a extração de 2 a 9 sinais, respectivamente. Em vermelho estão apresentadas as maiores correlações obtidas para as componentes extraídas de cada um dos dois blocos.

Na Figura 25, não foi representado o primeiro componente independente, pois ele está relacionado com a média dos dados presentes na matriz. Os resultados da Figura 25 demonstram que acima de cinco ICs há uma mistura de sinais. Essa mistura de sinais é representada pela presença de dois ou mais *pixels* com cores diferentes de azul escuro e vermelho na mesma linha ou coluna (como observado para as Figuras 25 e-h). Os dados foram divididos em dois blocos sendo que cada um dos eixos representa os ICs extraídos para um desses blocos. Portanto, com base nos resultados apresentados estima-se que a Amostra 1 possua no máximo cinco ICs a serem extraídos.

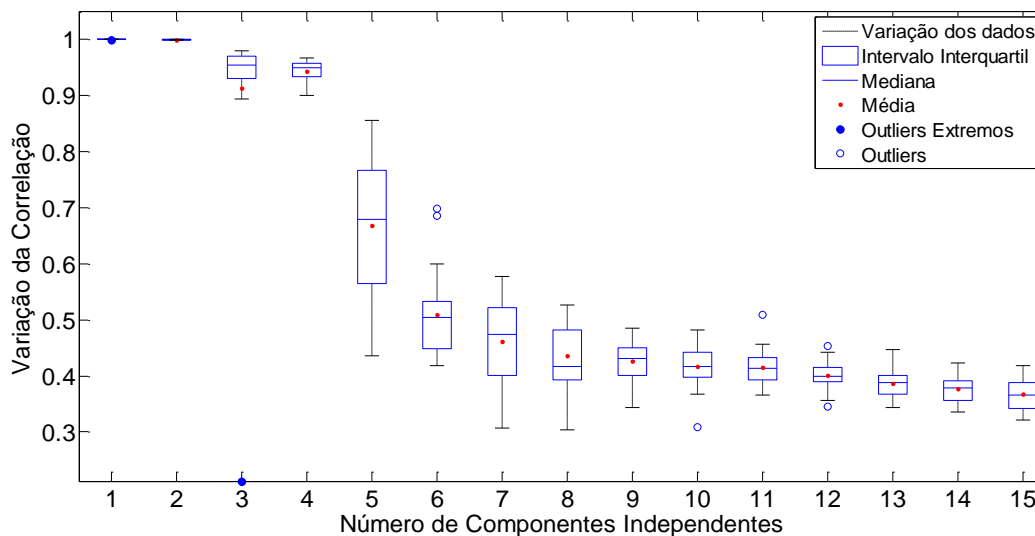


Figura 26. Resultado dos 20 cálculos do ICA por blocos obtidos para a Amostra 1.

Na Figura 26 é possível observar que há uma grande variação entre os resultados de correlação e o número de ICs extraídas. A etapa de criação dos blocos é a mais importante, pois observa-se uma variação de correlação de 0,44 até 0,86 para cinco ICs. Esses resultados devem ser interpretados com base na menor variação existente no intervalo interquartil e ao mesmo tempo altos valores de correlação para os ICs extraídos. Com isso, estima-se que a Amostra 1 possua no máximo quatro ICs a serem extraídos.

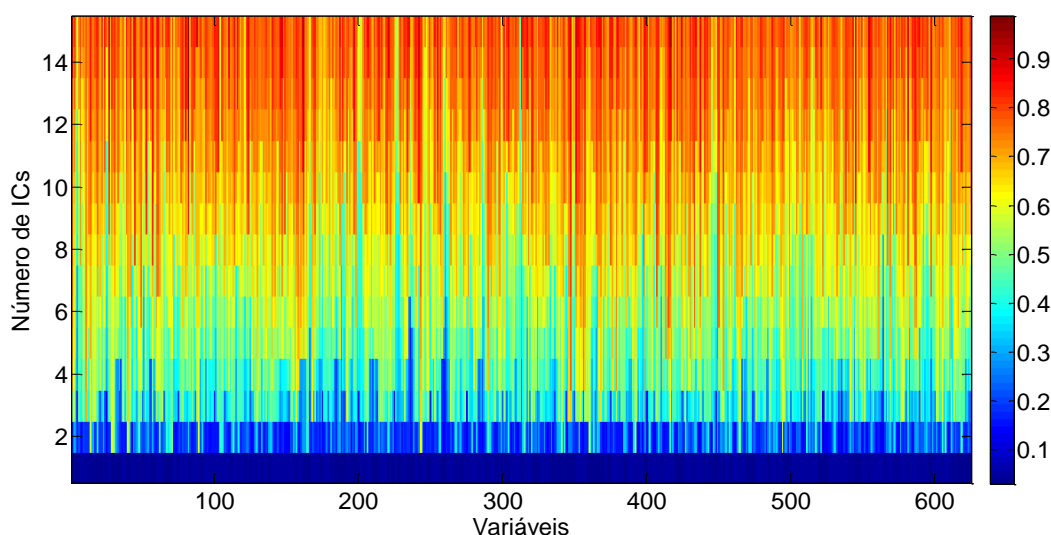


Figura 27. Resultado da determinação do número de ICs com base no critério de Durbin-Watson. A coloração azul indica alta correlação entre o número de componentes e as variáveis. Já a coloração vermelha indica a observação de ruído nas ICs extraídas.

Finalmente, na Figura 27 identifica-se que, no máximo, quatro ICs devem ser utilizadas para explicar a Amostra 1. Como apresentado pela Equação 11 os ICs extraídos não podem possuir ruído (cor vermelha na Figura 27), com isso observa-se em quais ICs ainda há a presença de variáveis com a coloração azul e ausência do vermelho. Os três métodos indicaram que para a Amostra 1 poderiam ser extraídos até quatro ICs. O mesmo procedimento foi realizado para cada uma das 22 amostras analisadas.

Para o MCR – ALS foi utilizada a SVD (Figura 28) para determinar o número de componentes presentes na matriz aumentada. Diferentemente da ICA, o número de componentes presentes na matriz aumentada poderia ser estimado com base nos resultados prévios de cada amostra mais o número de imagens químicas adicionadas.

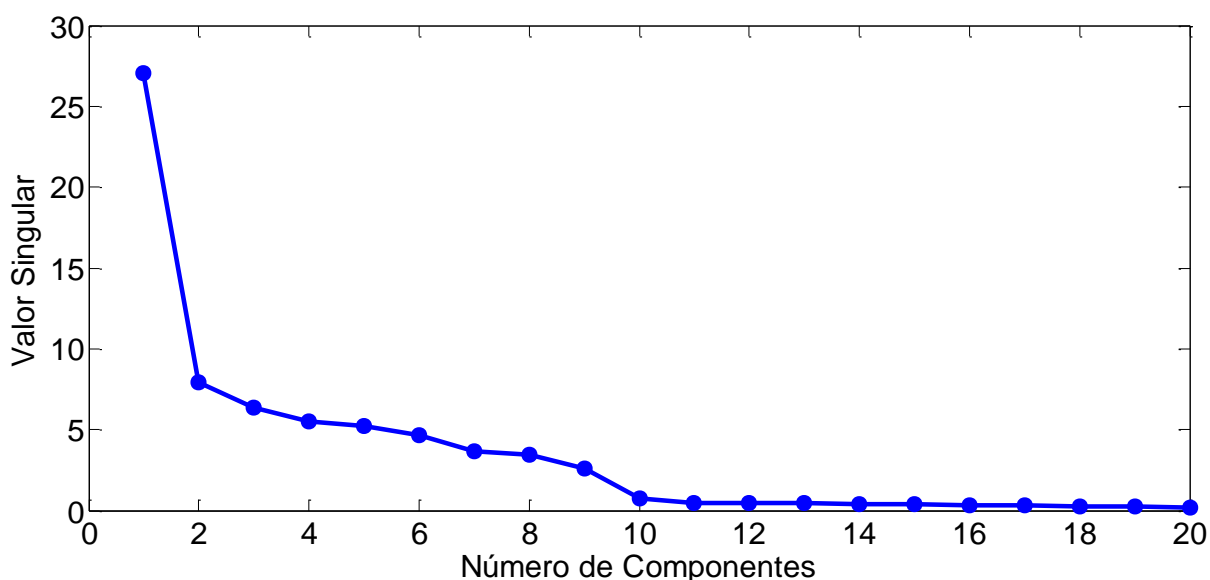


Figura 28. Análise dos valores singulares da matriz aumentada da Amostra 1.

A análise de valores singular (Figura 28) demonstra que poderiam ser utilizados até 10 componentes para a Amostra 1. Com base na matriz aumentada era de se esperar no mínimo nove componentes, pois os mesmos estão relacionados com as imagens puras (7 imagens) e simuladas (2 imagens). A mesma análise foi realizada para todas as outras 22 amostras.

5.3.3.6. Métodos de Resolução de Curvas

Após o pré-processamento dos dados eles foram analisados com os métodos de resolução de curvas MCR – ALS e ICA – JADE. O MCR – ALS permite o uso de restrições para que os resultados possuam sentidos físico e químico, com isso

foi utilizada a restrição de não-negatividade para os perfis de concentração e espectral. A Figura 29 apresenta de forma esquemática como foi montada a matriz de dados **D** e as duas matrizes resposta dos métodos de resolução de curvas.

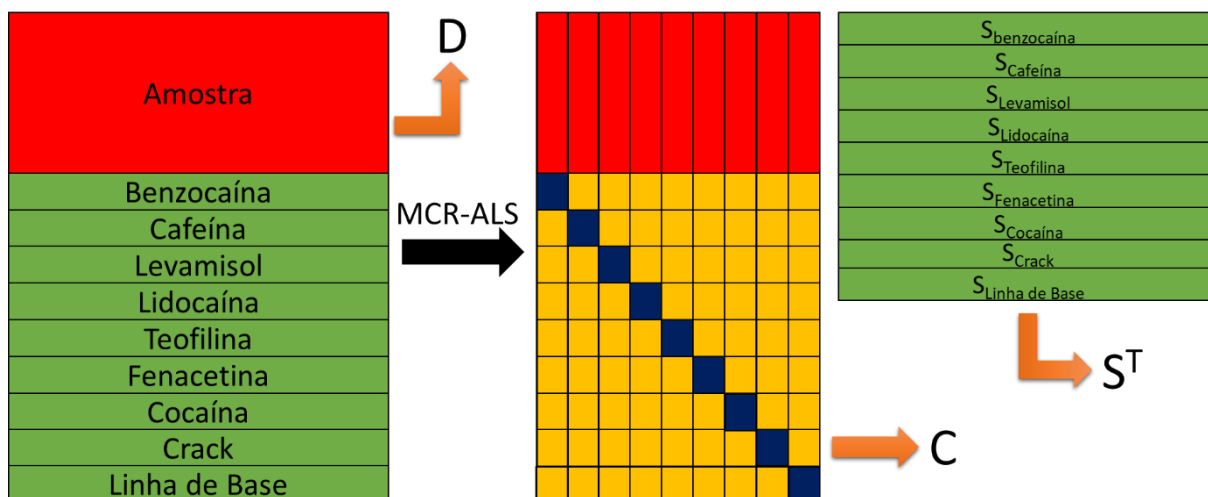


Figura 29. Arranjo da matriz aumentada utilizada para a decomposição do MCR – ALS. A matriz **D** apresenta retângulos vermelhos e verdes que estão relacionados com os espectros da amostra e dos fármacos/drogas, respectivamente. A matriz **C** apresenta os resultados dos escores do MCR para a amostra (em vermelho) e dos compostos puros (em amarelo e azul). Em amarelo e azul estão representados os escores referentes a ausência e presença dos compostos puros, respectivamente.

5.3.3.7. Limite de Identificação

A partir dos escores obtidos na etapa anterior foram calculados os valores mínimos de escores para se afirmar quais compostos estão presentes nas amostras analisadas. A Figura 30a representa a matriz de escores, onde os retângulos em vermelho são os escores referentes aos *pixels* da amostra de cocaína/crack. Esses *pixels* foram utilizados para a construção das imagens químicas. Já os *pixels* representados pelos quadrados azuis e amarelos estão relacionados diretamente com as imagens químicas dos compostos puros. Por causa disso foram calculados os valores mínimos de escores para se identificar todos os compostos que estariam presentes nas amostras.

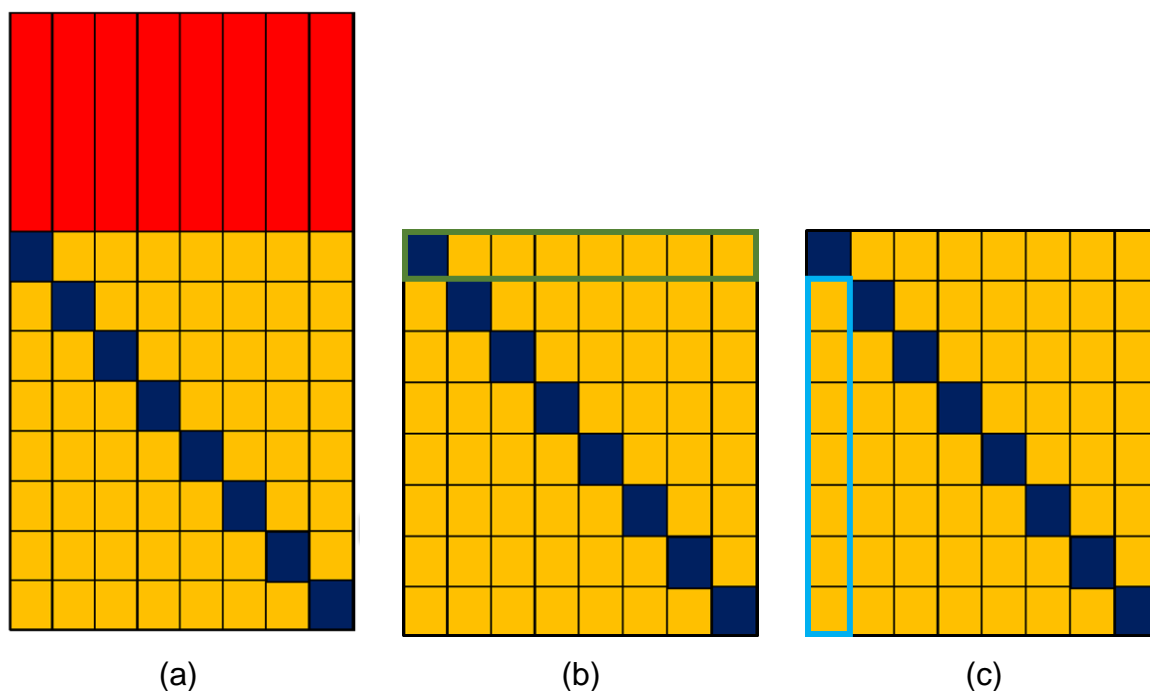


Figura 30. Matriz de escores obtida após a otimização da matriz **D** com os métodos de resolução de curvas: (a) matriz **C** completa; (b) escores a serem normalizados, cada linha representa a imagem de um componente puro (retângulo verde); (c) escores a serem utilizados para a média de cada composto na amostra. Os valores de 0 e 1 representam os escores esperados para ausência e presença de um composto químico na matriz **C**.

A seguir estão listadas as etapas para determinação do “limite de identificação”.

- Separar os escores em amarelo e azul (Figura 30b);
 1. Normalizar cada uma das imagens dos N compostos puros (escores dentro do retângulo verde na Figura 30b), para que os escores em amarelo e azul sejam próximos de 0 e 1, respectivamente;
 2. Calcular a média por coluna dos escores normalizados (escores dentro do retângulo azul claro na Figura 30c), desconsiderando os escores do quadrado azul;
- Após realizar os passos 1 e 2 para todas as amostras, calcular a média dos escores do Composto N com o resultado do passo 2;
- O “limite de identificação” do Composto N é o valor da média calculada anteriormente acrescido de três desvios-padrão.

Com esses valores de escores limites é possível identificar a presença de todos os componentes nas amostras analisadas. A última etapa é comparar os escores das amostras com os limites calculados. Se o escore for maior do que o limite pode-se afirmar que aquele composto está presente na amostra.

5.4. Resultados e Discussão

Inicialmente o NPC – Campinas não informou quais eram os aditivos presentes nas amostras. Entretanto, na literatura^{17,82,83,87-90} encontra-se a massiva presença dos aditivos: benzocaína; cafeína; levamisol; lidocaína; teofilina. Todas as amostras analisadas eram contra-prova; por isso não foi autorizado o uso de análises analíticas destrutivas. Dessa forma, foi utilizada a análise de componentes independentes para a caracterização inicial dessas amostras. Em seguida, foi proposta a criação de um limite de identificação que, em tese, seria capaz de identificar quais eram os possíveis compostos presentes nas amostras.

5.4.1. Caracterização das amostras com JADE – ICA

Diferentemente do MCR – ALS, a ICA não necessita de estimativas iniciais das matrizes **A** e **S**. Por causa disso, diz-se que ela realiza uma análise às cegas do conjunto de dados. Essa diferença permitiu seu uso para a caracterização das amostras de cocaína e crack, sendo que nesse caso cada amostra foi tratada de forma independente, onde um modelo foi construído para cada imagem.

Além disso, o algoritmo JADE não utiliza restrições, como a não-negatividade nos perfis de proporções e sinais, na obtenção de seus resultados. Assim como em todos os métodos quimiométricos, é crucial a etapa de determinação de número de componentes, pois um erro nessa etapa causaria obtenção de resultados equivocados.

A seção 6.3.3.5 apresentou todos os critérios utilizados para a determinação do número de componentes independentes para cada uma das amostras. Alguns desses componentes puderam ser relacionados diretamente com os aditivos indicados pelos peritos criminais. Além disso, a ICA é capaz de identificar sinais relacionados com as diferenças físicas presentes nas amostras (no caso de imagens hiperespectrais devido a focos distintos na mesma amostra). Os componentes independentes, também, podem ser relacionados com o ajuste da linha

de base dos sinais extraídos, como o sinal extraído pela ICA para a Amostra 3 representado pela Figura 31.

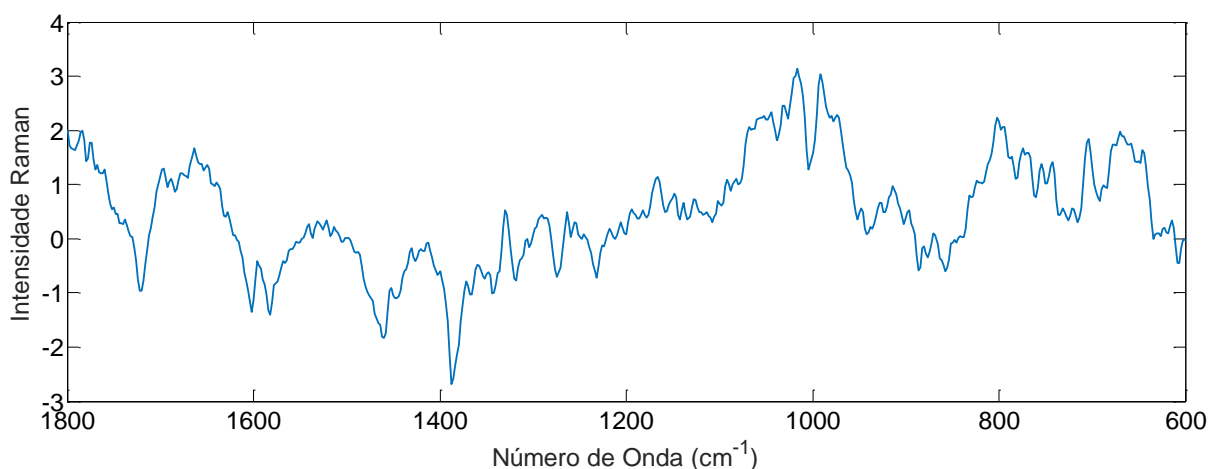


Figura 31. Sinal obtido com o JADE para a Amostra 3, onde é possível observar que esse sinal não apresenta nenhuma banda que poderia ser relacionada com algum tipo de informação química. Esse sinal está relacionado com a variação de linha de base dos espectros da Amostra 3.

A mesma amostra que apresentou o sinal da Figura 31 teve como resultado os espectros apresentados na Figura 32. A partir dos espectros de referência foi possível a identificação dos sinais extraídos. Na Figura 33 estão representadas as imagens químicas correspondentes aos espectros da Figura 32. Os métodos de determinação de componentes indicaram que para a Amostra 3 era necessário o uso de quatro componentes independentes os quais foram apresentados nas Figuras 31 e 32.

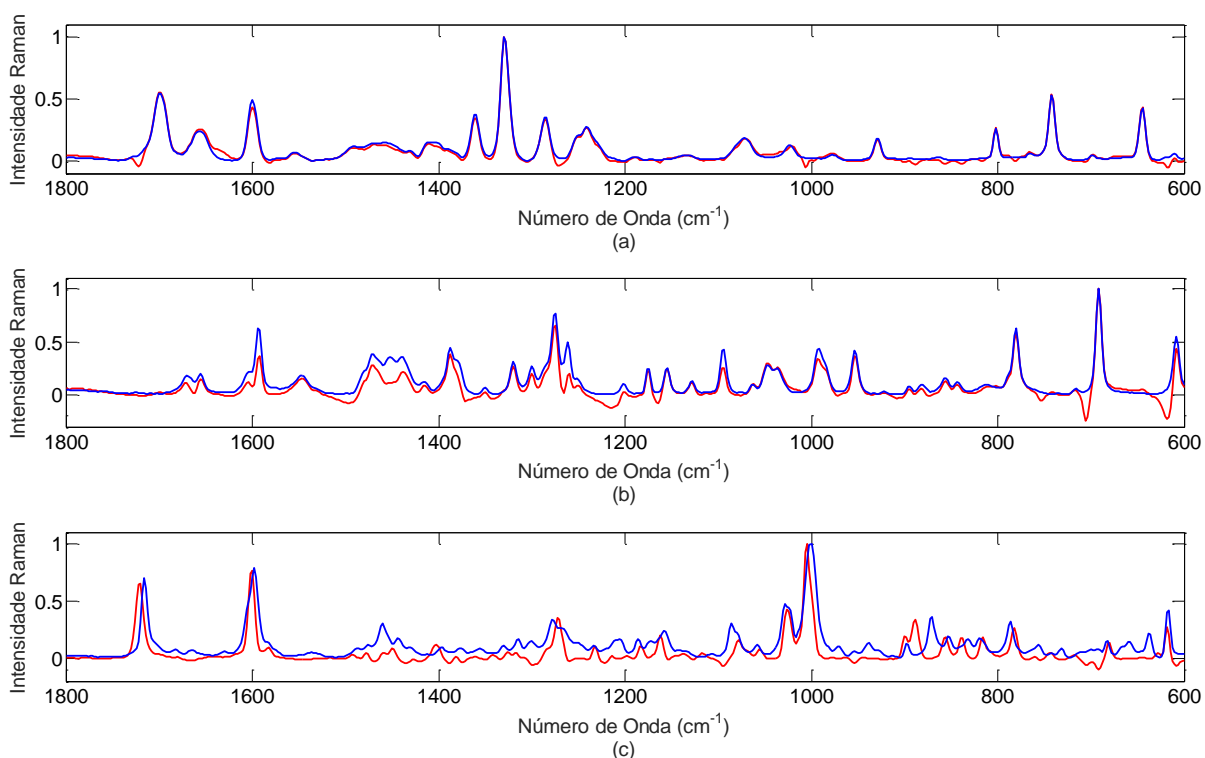


Figura 32. Comparação entre os espectros de referência (em azul) e os sinais (em vermelho) obtidos pela ICA – JADE: (a) cafeína; (b) lidocaína; (c) cocaína.

Os sinais obtidos pela ICA – JADE apresentam algumas regiões negativas, como apresentado nas Figuras 32b e 32c. Como o JADE não possui restrições, como a não-negatividade, seus sinais, bem como as proporções químicas, podem apresentar resultados negativos. Por causa disso, as imagens químicas (Figura 33) possuem o seu mínimo de proporções negativo. Esse fato pode ser comprovado pela escala de cores apresentada na mesma Figura. A escala de cores do azul para o vermelho indicam um gradiente da ausência até a presença do componente químico analisado.

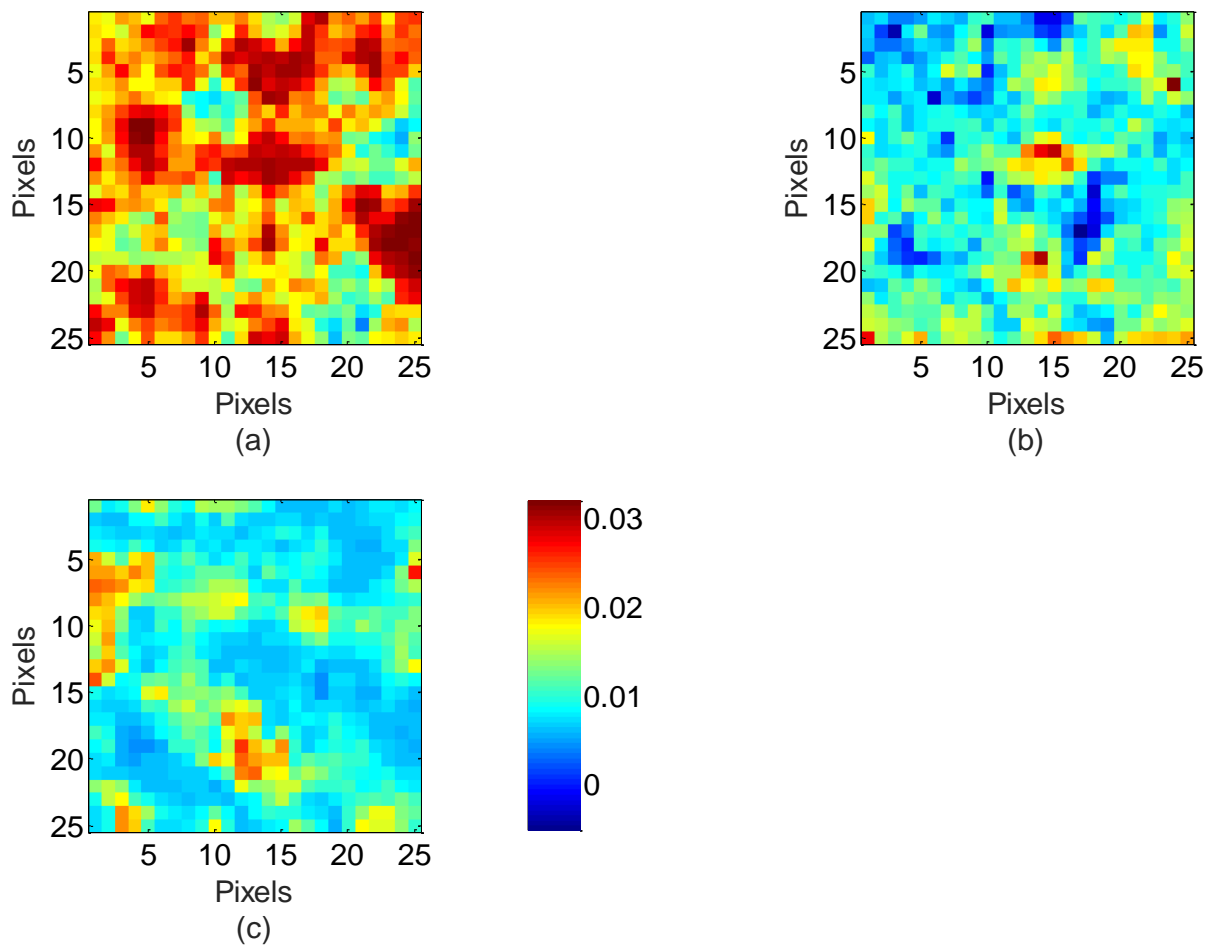


Figura 33. Imagens químicas da Amostra 3: (a) cafeína; (b) lidocaína; (c) cocaína. As cores azul e vermelho simbolizam um gradiente de proporções químicas desses compostos da sua ausência até a presença nos *pixels* analisados.

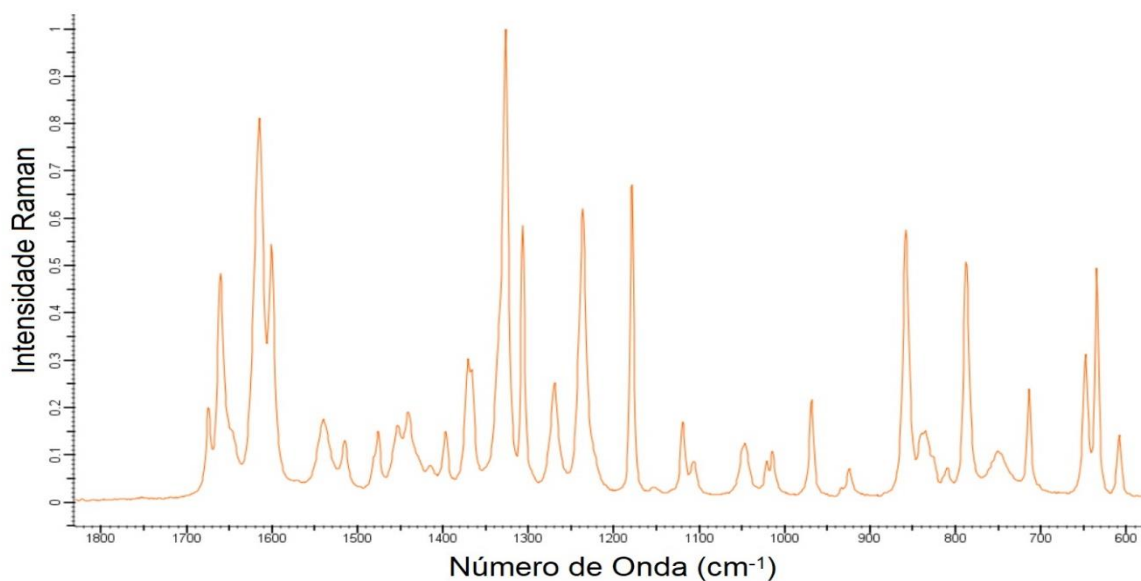
Ao final de todos os modelos da ICA foi possível a caracterização das amostras cedidas pelo NPC – Campinas. A Tabela 6 contém o resultado da composição das 22 amostras analisadas. O crack foi encontrado em 11 amostras e a cocaína em 6 amostras. Os aditivos encontrados em maior parte das amostras foram cafeína (11 amostras) e lidocaína (8 amostras). O levamisol foi encontrado em duas amostras, benzocaína e teofilina foram encontrados em somente uma amostra e a dipirona não foi encontrada em nenhuma amostra.

Tabela 6. Identificação dos compostos presentes nas amostras de cocaína/crack com base nos resultados da ICA – JADE.

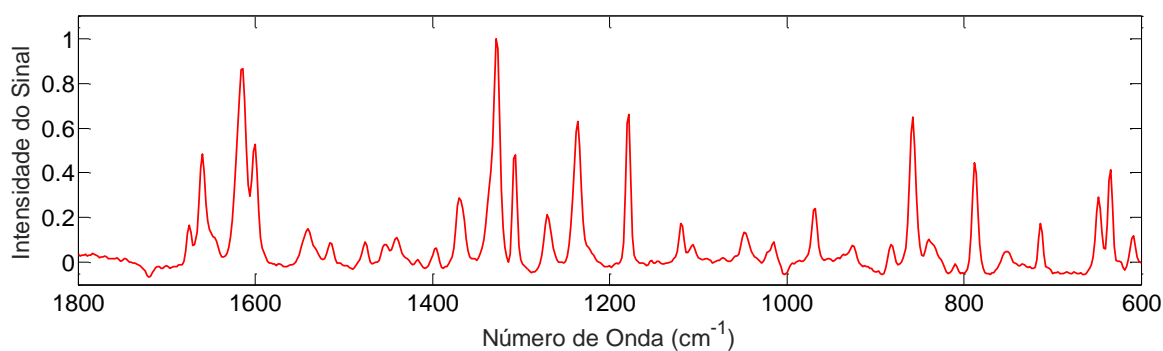
Amostra	Composição
1	Cafeína e Lidocaína
2	Cafeína e Cocaína
3	Cafeína, Lidocaína e Cocaína
4	Cafeína, Lidocaína e Cocaína
5	Cafeína e Lidocaína
6	Benzocaína, Cafeína e Teofilina
7	Fenacetina e Crack
8	Fenacetina e Crack
9	Fenacetina e Cocaína
10	Fenacetina e Crack
11	Lidocaína e Crack
12	Fenacetina e Crack
13	Cafeína, Fenacetina e Crack
14	Cafeína, Lidocaína e Crack
15	Cafeína, Lidocaína e Crack
16	Fenacetina e Crack
17	Cafeína, Lidocaína e Crack
18	Cafeína, Fenacetina e Lidocaína
19	Crack
20	Levamisol e Cocaína
21	Levamisol e Cocaína
22	Cocaína

Em algumas amostras foi determinada a presença da fenacetina, a qual, inicialmente, não havia sido considerada como aditivo para esse conjunto de amostras. O NPC – Campinas não forneceu uma amostra desse composto puro, uma vez que não possuíam esse composto no laboratório. A fenacetina foi encontrada em 8 amostras, sendo que, normalmente, ela é utilizada como aditivo da pasta base de cocaína. A identificação do espectro foi realizada com base em espectros encontrados

na literatura^{91,92} (Figura 34). O Anexo 3 apresenta a atribuição de bandas para a fenacetina e para os demais aditivos.



(a)



(b)

Figura 34. (a) Espectro da fenacetina encontrado na literatura⁹²; (b) sinal extraído pelo JADE referente à fenacetina.

Os resultados da Tabela 6 ainda demonstram que duas amostras (Amostras 19 e 22) apresentaram somente um composto químico. Nessas duas amostras somente foram identificados os espectros referentes ao cloridrato de cocaína e à pasta base de cocaína. Apesar do número de componentes independentes ser maior do que um, não foram extraídos outros sinais que pudessem ser relacionados com qualquer outro aditivo. A Figura 35 apresenta o espectro médio dessas duas amostras.

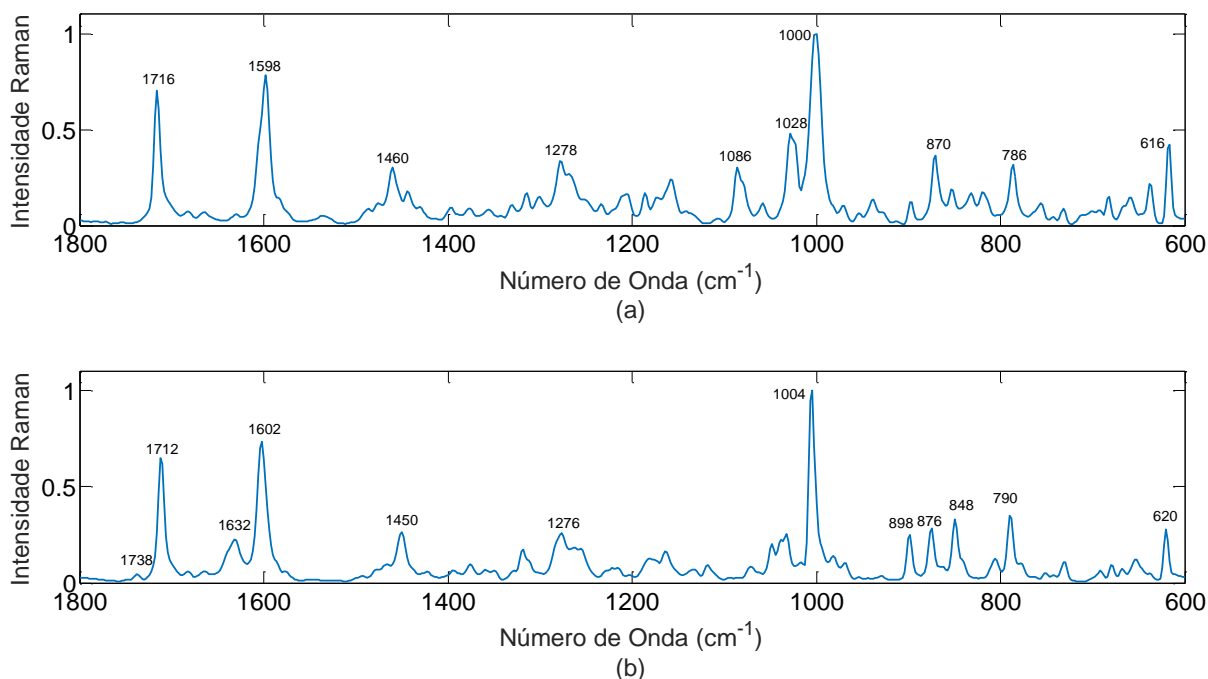


Figura 35. Espectro médio das amostras que apresentaram somente um componente químico após a ICA: (a) Amostra 22 – cocaína; (b) Amostra 19 – crack.

Há algumas semelhanças entre os espectros apresentados na Figura 35, como demonstrado na Tabela 7 que contém a atribuição das bandas mais importantes. Entretanto, a principal diferença está relacionada com a banda em 1632 cm^{-1} que representa o estiramento do ácido benzóico. Em teoria, apesar da similaridade entre os espectros, isso não irá impedir a utilização dos métodos de resolução de curvas.

Tabela 7. Atribuição das bandas mais características para cocaína e crack.

Composto	Número de onda (cm^{-1})	Atribuição ^{17,82}
Cocaína e Crack	848, 874 e 898	Estiramento C-C – tropano
	1000 – 1004	Estiramento simétrico – respiração do anel aromático
	1453	Deformação assimétrica do grupo CH_3
	1605	Estiramento C=C do anel aromático
	1712	Estiramento simétrico C=O
	1735	Estiramento assimétrico C=O
Crack	1632	Estiramento do ácido benzóico

5.4.2. Caracterização das amostras com MCR – ALS e aplicação do limite de identificação

Os resultados obtidos com o JADE proporcionaram maior conhecimento sobre as amostras. Entretanto, para determinar quais eram os aditivos presentes sempre será necessária a comparação entre todos os sinais obtidos com os espectros de referência. Além disso, pode ocorrer de alguns sinais não possuam nenhum significado químico, como foi o caso dos sinais relacionados com o ajuste de linha de base.

Uma forma de tentar contornar esse problema seria utilizando uma matriz aumentada como a da Figura 29. O número de componentes estimados foi condizente com o conjunto de dados, mas alguns dos sinais extraídos não puderam ser correlacionados com os espectros de referência. Pode-se dizer que o tamanho da matriz aumentada foi o fator mais importante para a não obtenção dos sinais esperados. Essa matriz possui grande quantidade de informação sobre todas as imagens obtidas experimentalmente e essa variabilidade acaba fazendo com que a ICA misture os sinais ao invés de extraí-los. Conseqüentemente não foi possível criar um limite de identificação com base nas proporções químicas da ICA (resultados da matriz A).

Por conta desses resultados, foi proposto o uso do MCR – ALS para a criação do limite de identificação. Antes de montar a matriz aumentada, onde para cada amostra são acrescentadas as imagens dos compostos puros, todas as amostras foram analisadas pelo MCR – ALS do mesmo modo que foi realizado para a ICA. Ao final dessas otimizações verificou-se que o MCR não foi capaz de recuperar o espectro da fenacetina como ocorreu na ICA. O perfil recuperado era uma mistura do espectro da fenacetina com o espectro da pasta base de cocaína. Entretanto, todos os outros aditivos foram recuperados pelo MCR.

Por fim, foi montada a matriz aumentada com as imagens da amostra e dos sete componentes puros (benzocaína, cafeína, levamisol, lidocaína, teofilina, cocaína e crack) além das imagens simuladas de fenacetina e do ajuste de linha de base. Cada uma dessas matrizes foi decomposta pelo MCR – ALS com o número de componentes adequados e com as estimativas iniciais de cada um dos compostos adicionados.

Antes da decomposição verificou-se a matriz de estimativa inicial de perfil espectral para cada uma das matrizes construídas. A escolha desses espectros não é aleatória, ela é baseada na procura das variáveis mais puras que estão presentes no conjunto de dados. Em outras palavras, são escolhidos os espectros mais distintos entre si no conjunto de dados, sendo que serão escolhidos quantos componentes forem determinados pela decomposição em valores singulares. Para todas as amostras ficou evidente a escolha de um espectro de cada uma das imagens dos componentes puros.

Após a verificação da estimativa inicial todas as matrizes foram decompostas linearmente pelo MCR – ALS. Ao final dessa decomposição foram obtidos os espectros recuperados para os componentes puros (Figura 36). A comparação visual e os altos valores de correlação indicaram que o MCR – ALS foi capaz de reconhecer todos os componentes presentes na matriz de dados. Além dessa comparação visual foi realizada a atribuição das bandas mais características para esses compostos (Anexo 3).

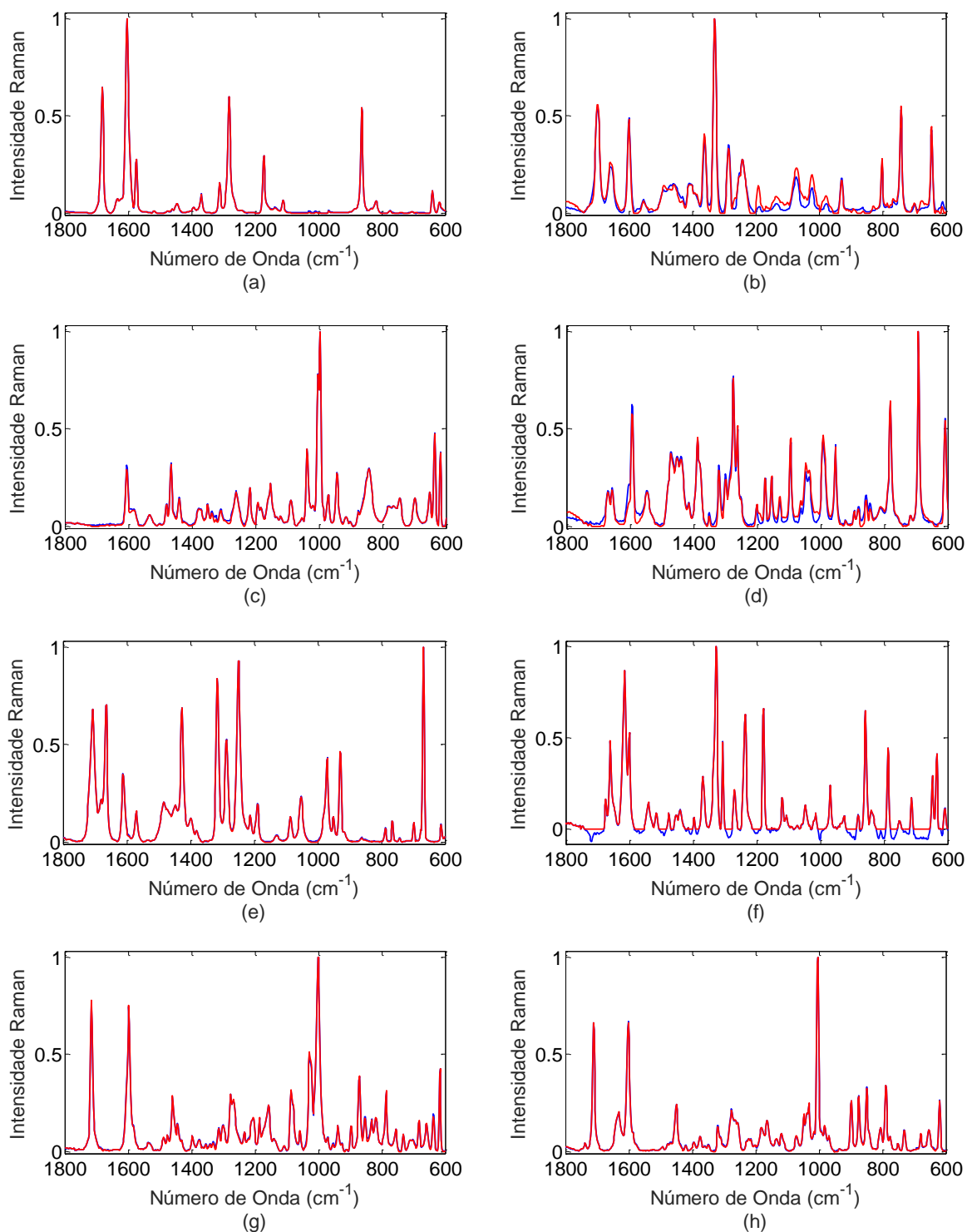


Figura 36. Comparação entre os espectros Raman de referência (em azul) e recuperados pelo MCR – ALS (em vermelho) para a Amostra 1: (a) benzocaína – correlação de 0,9999; (b) cafeína – correlação de 0,9800; (c) levamisol – correlação de 0,9991; (d) lidocaína – correlação de 0,9870; (e) teofilina – correlação de 0,9999; (f) fenacetina – correlação de 0,9954; (g) cocaína – correlação de 0,9992; (h) crack – correlação de 0,9996.

Após essa comparação foram obtidas as imagens químicas desses componentes (Figura 37). A escala de cores de azul para vermelho indica o acréscimo dos valores de escores do MCR. Essas imagens indicam quais compostos estão presentes nas amostras analisadas.

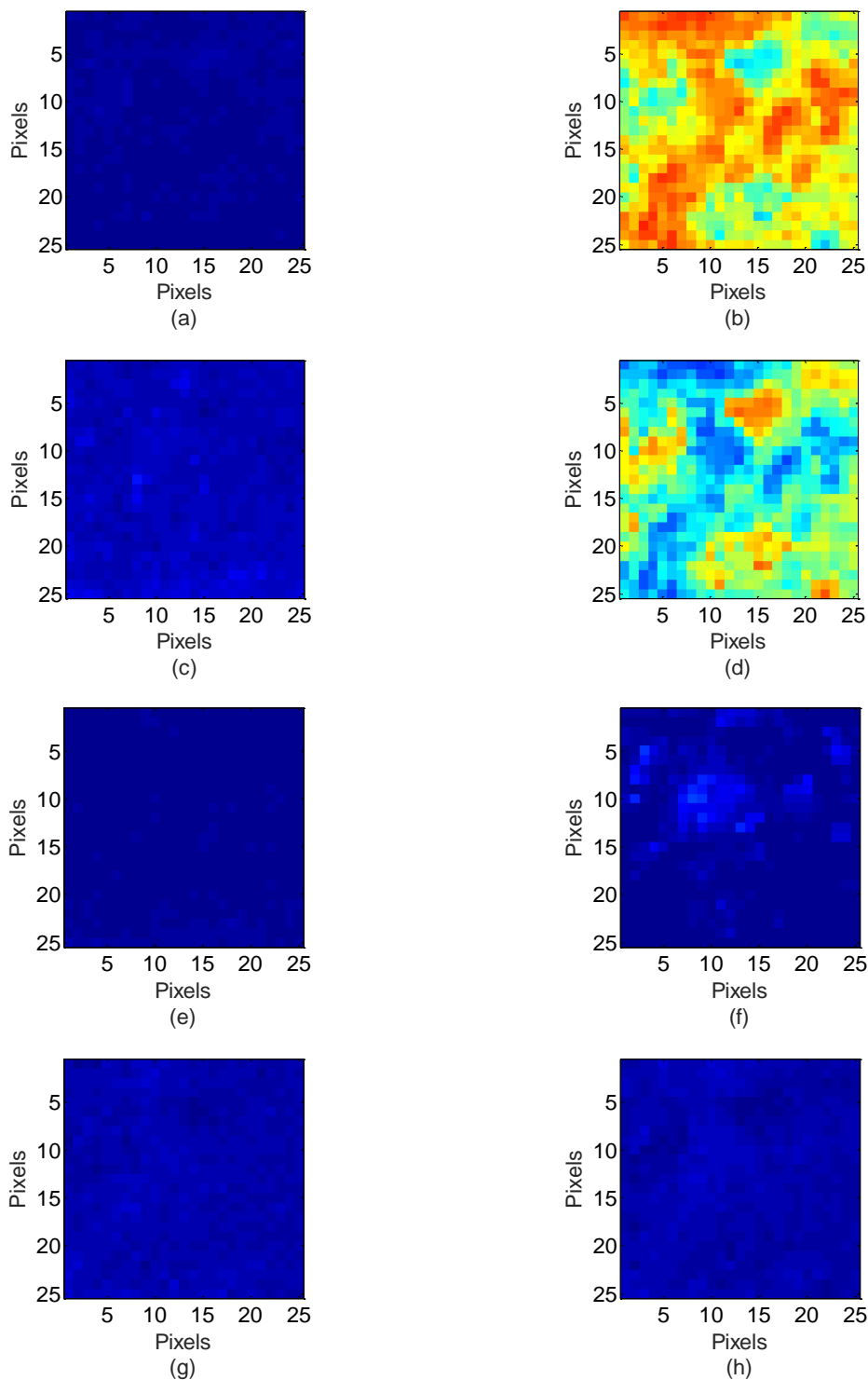


Figura 37. Imagens químicas da Amostra 1: (a) benzocaína; (b) cafeína; (c) levamisol; (d) lidocaína; (e) teofilina; (f) fenacetina; (g) cocaína; (g) crack.

Esses resultados indicaram a presença de cafeína, lidocaína e talvez fenacetina na Amostra 1. Mesmo com esses indícios visuais é possível que algum outro composto esteja presente só que em menor escala, visto que essas imagens foram criadas com a mesma escala de cores. Essa dúvida está baseada no fato de que todos os resultados da matriz C possuem unicamente um sentido matemático. Sentido esse que surge quando são utilizadas restrições para a construção dos modelos, critérios de convergência e matrizes aumentadas. O sentido químico deve ser estabelecido pelo usuário dos métodos, pois ele deve avaliar concomitantemente as matrizes C e S^T . Separadamente essas matrizes podem induzir a conclusões tendenciosas sobre o conjunto de dados.

A literatura é rica em trabalhos que utilizam matrizes aumentadas para analisar as imagens hiperespectrais. Todavia, esses trabalhos muitas vezes não utilizam informações de compostos que não estão presentes no conjunto de dados da amostra. Desse modo, não há métodos estatísticos que irão ajudar na determinação de quais compostos estão presentes na amostra. A criação desse “limite de identificação” visa usar os escores do MCR para apontar quais seriam os aditivos de cada uma das amostras.

Na seção 6.3.3.7 foi apresentada a metodologia para o cálculo do “limite de identificação”; os mesmos estão apresentados na Figura 38. Foram calculados os limites para todos os componentes puros, inclusive para a fenacetina que teve a sua imagem química simulada.

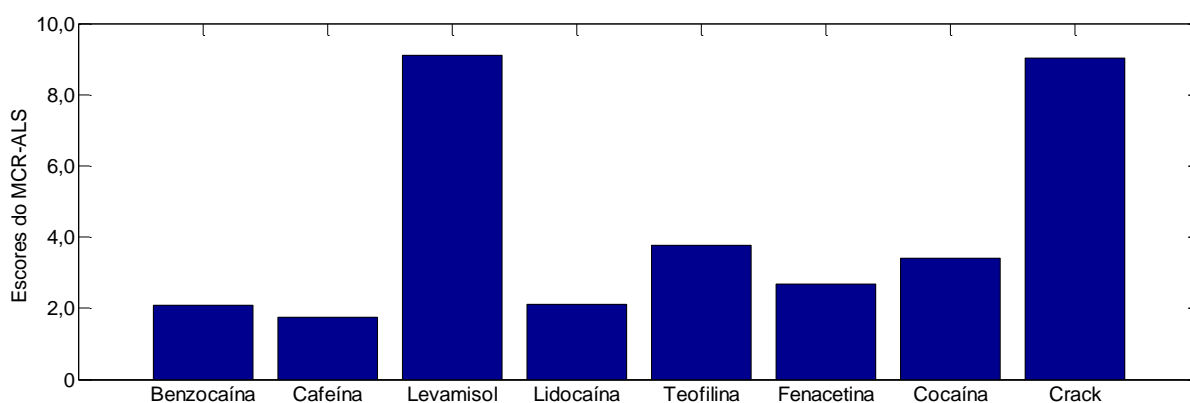


Figura 38. Valores dos “limites de identificação” em termos de escores para os compostos analisados.

Com exceção do levamisol e do crack, todos os compostos apresentaram limites inferiores a 1,50%. Sendo que os resultados mais expressivos foram para a

cocaína (0,21%) e cafeína (0,36%) por apresentarem os menores limites calculados. Esses dois resultados podem estar diretamente ligados com a presença desses compostos nas amostras analisadas. Em outras palavras, as amostras analisadas possuíam mais informações sobre esses dois compostos, o que pode fazer com que eles acabem contaminando todos os *pixels* analisados na aplicação dos métodos de resolução de curvas.

A partir desse ponto pode-se inferir que, possivelmente, o levamisol não será encontrado em muitas amostras, pois ele possuiu o maior limite calculado (3,40%). Para esse composto seria importante haver maior quantidade de amostras e avaliar a alteração do limite em função de um conjunto de dados maior.

O uso das imagens simuladas (fenacetina e ajuste da linha de base) influenciaram positivamente na criação desses limites, como verificado na Tabela 8. Nessa Tabela estão comparados os valores dos limites de identificação para os conjuntos de dados com e sem as imagens simuladas.

Tabela 8. Comparação entre os “limites de identificação” com e sem as imagens simuladas. Esses limites estão apresentados em porcentagem de escores do MCR.

Composto	Limites de identificação	
	Com imagens simuladas	Sem imagens simuladas
Benzocaína	2,08	11,32
Cafeína	1,76	1,76
Levamisol	9,10	20,62
Lidocaína	2,10	7,18
Teofilina	3,76	5,57
Cocaína	3,39	24,06
Crack	9,02	6,05

Os resultados da Tabela 8 demonstram que as imagens simuladas possibilitaram obter menores valores para os limites propostos. Com exceção da cafeína, todos os outros compostos apresentaram melhorias expressivas para o uso das imagens químicas. Entretanto, pode-se afirmar que essas melhoras podem estar mais relacionadas com o uso da imagem química da linha de base dos espectros das amostras. Principalmente para o caso da cocaína que teve o seu limite diminuído de 7,02% para 0,21%.

Colocando em um gráfico de barras os escores médios de cada composto nas amostras e traçando-se o “limite de identificação” correspondente foi possível

identificar quais seriam os constituintes presentes nas amostras (Figura 39). É importante ressaltar que essa metodologia pode ser mais confiável do que a simples observação das imagens químicas.

Com os resultados da Figura 39 as amostras foram separadas em quatro grupos de acordo com a presença de cocaína e crack. Quatro amostras continham cocaína e crack, dez amostras continham apenas cocaína, três amostras apenas crack e duas amostras não continham cocaína ou crack. As Figuras 40 – 42 apresentam um exemplo de cada grupo. Comparando as imagens da Figura 37 com as da Figura 40 foi possível verificar que o “limite de identificação” proposto pode ser utilizado para indicar quais seriam os constituintes de uma amostra. Outro indicativo foi a complementaridade das imagens químicas das Figuras. Em alguns casos houve certa dificuldade encontrar essa complementaridade, mas no final o padrão foi observado para todas as amostras.

Ainda com relação à Figura 39, verificou-se que cafeína (14 amostras) e lidocaína (12 amostras) foram os compostos mais comumente adicionados nessas drogas de rua. Já a benzocaína apareceu em sete, levamisol e teofilina em uma amostra. Esses resultados estão sumarizados na Tabela 9.

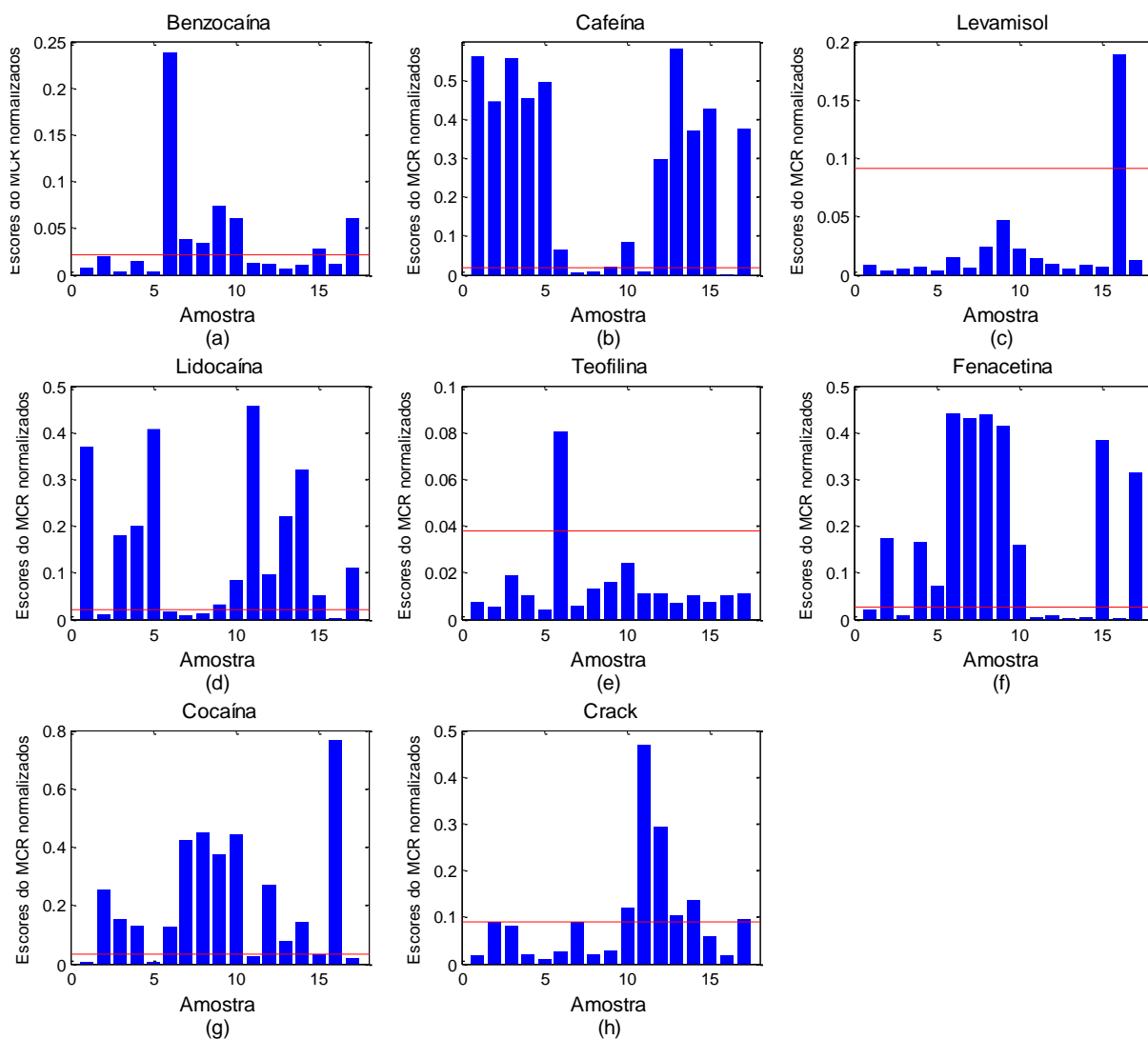


Figura 39. Cálculo dos escores dos compostos presentes nas amostras para: (a) benzocaína; (b) cafeína; (c) levamisol; (d) lidocaína; (e) teofilina; (f) fenacetina; (g) cocaína; (h) crack. A reta em vermelho representa a média acrescida de dois desvios padrão para cada um desses compostos.

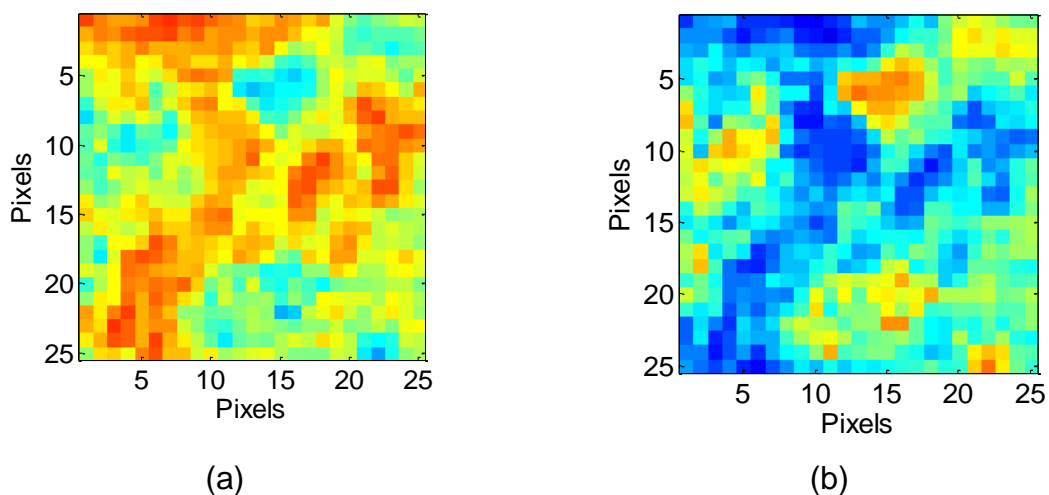


Figura 40. Imagens químicas para os compostos identificados para a Amostra 1: (a) cafeína; (b) lidocaína.

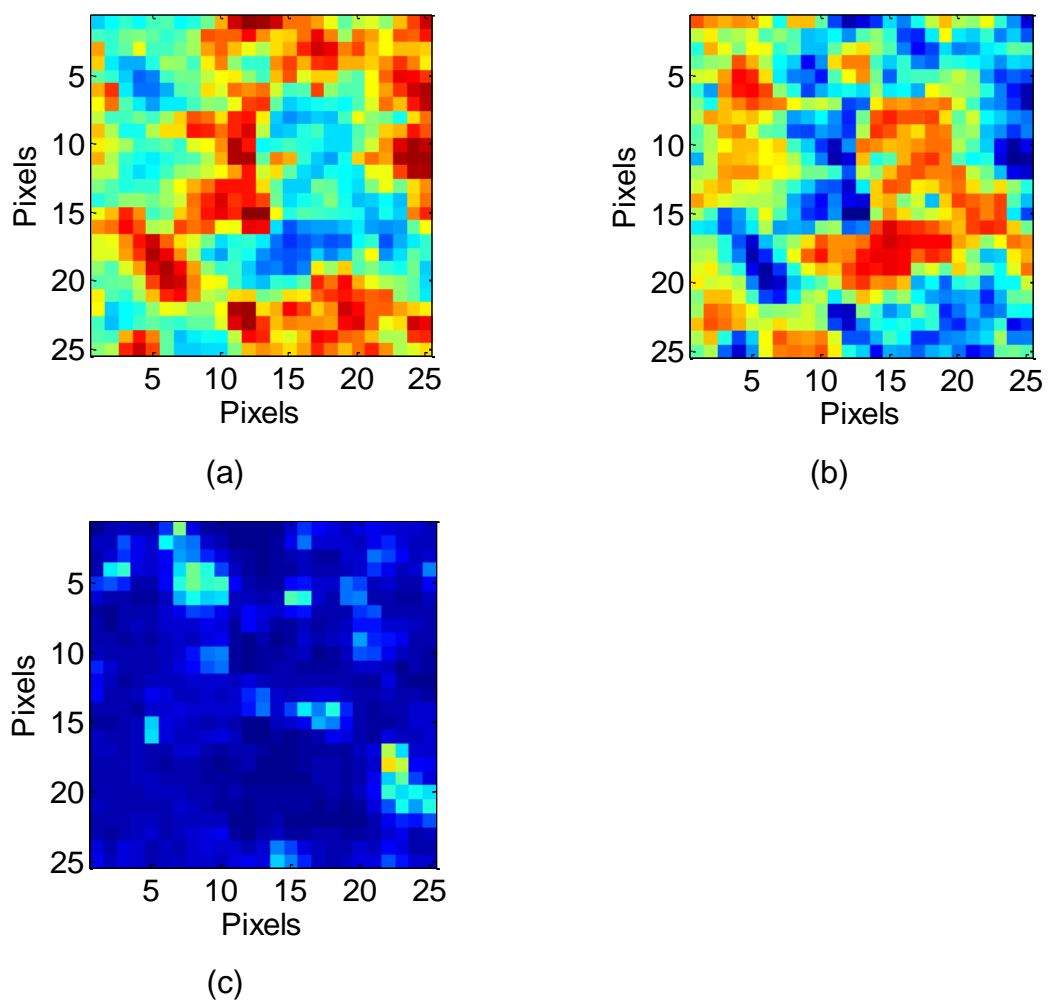


Figura 41. Imagens químicas para os compostos identificados para a Amostra 5: (a) cafeína; (b) lidocaína; (c) fenacetina.

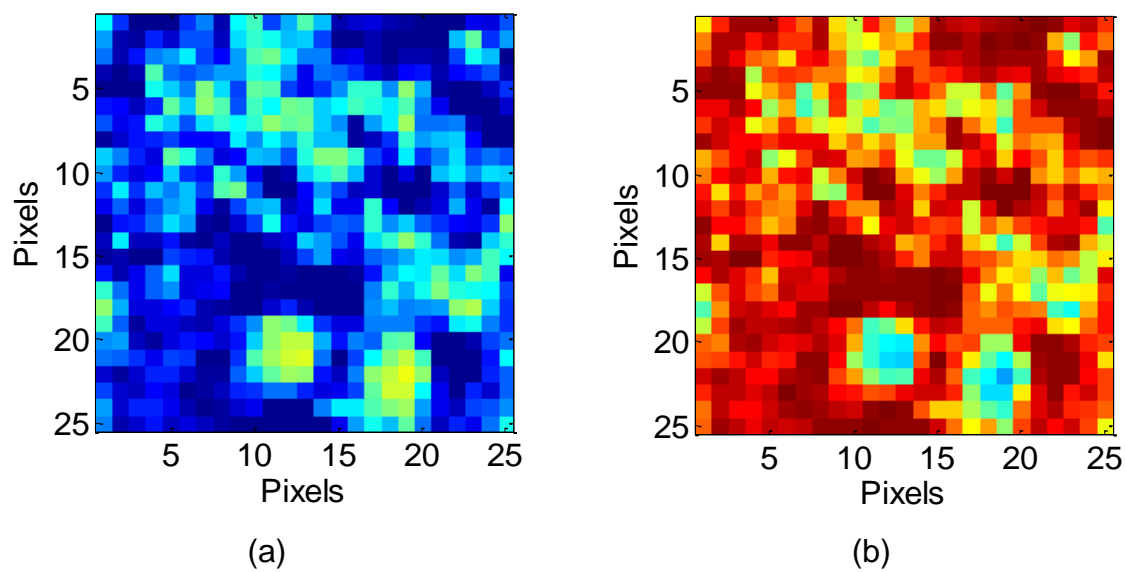


Figura 42. Imagens químicas para os compostos identificados para a Amostra 16: (a) benzocáína; (b) levamisol; (c) cocaína; (d) crack.

Tabela 9. Identificação dos compostos presentes nas amostras de cocaína/crack após a análise do “limite de identificação”.

Amostra	Composição
1	Cafeína e Lidocaína
2	Cafeína, Fenacetina e Cocaína
3	Cafeína, Lidocaína e Cocaína
4	Cafeína, Lidocaína, Fenacetina e Cocaína
5	Cafeína, Lidocaína e Fenacetina
6	Benzocaína, Cafeína, Teofilina, Fenacetina e Cocaína
7	Benzocaína, Fenacetina e Cocaína
8	Benzocaína, Fenacetina e Cocaína
9	Benzocaína, Cafeína, Lidocaína, Fenacetina e Cocaína
10	Benzocaína, Cafeína, Lidocaína, Fenacetina, Cocaína e Crack
11	Cafeína, Lidocaína e Crack
12	Cafeína, Lidocaína, Cocaína e Crack
13	Cafeína, Lidocaína, Cocaína e Crack
14	Cafeína, Lidocaína, Cocaína e Crack
15	Benzocaína, Cafeína, Lidocaína, Fenacetina e Cocaína
16	Levamisol e Cocaína
17	Benzocaína, Cafeína, Lidocaína, Fenacetina e Crack
18	Cocaína
19	Crack

5.5. Conclusões Parciais

Para as amostras de cocaína e crack foram obtidos resultados que puderam ser confirmados pelo Instituto de Criminalística, principalmente para os resultados obtidos com a ICA. Portanto, as técnicas espectroscópicas em conjunto com o ICA – JADE podem ser utilizadas para identificação de drogas apreendidas sem as etapas de preparo de amostra. Como diferencial foram utilizadas técnicas de imagens hiperespectrais para essas análises, sendo que normalmente são utilizados métodos cromatográficos. As imagens químicas trazem maior conhecimento sobre as amostras, mas ao mesmo tempo é obtida uma grande quantidade de dados a serem analisados, por isso os métodos quimiométricos precisam ser empregados.

Além disso, os métodos de resolução de curvas resultam em sinais que podem ser diretamente relacionados com os espectros dos componentes presentes nas amostras. Essa característica pode ser evidenciada com os altos valores de

correlação espectral que foram apresentados na Figura 36. E mais importante, foi possível a identificação de um aditivo (fenacetina) que não era esperado pelos peritos criminais.

A espectroscopia de imagem Raman não necessita de etapas laboriosas de preparo de amostra e mantém a integridade física das mesmas. Essas vantagens são importantes quando há quantidade muito pequena de amostra a ser analisada e ela deve ser guardada para análises futuras. Verificou-se que mesmo com regiões muito pequenas analisadas foi possível a identificação de diversos adulterantes nas amostras de cocaína ou crack.

Por fim, os resultados dos “limites de identificação” obtidos com base nos escores do MCR – ALS foram inferiores a 3,40% (levamisol). Para serem obtidos melhores resultados seria importante o uso de uma quantidade maior de amostras e, provavelmente, uma maior área analisada. Como esta é uma primeira aplicação desse tipo de análise, foi plenamente possível verificar que os resultados são bastante promissores e a espectroscopia de imagem Raman pode ser uma poderosa aliada na identificação de compostos adicionados nas drogas apreendidas pela polícia.

**CAPÍTULO 6 –
ADULTERAÇÃO DE
ÓLEO DE GERGELIM**

6.1. Introdução teórica e Motivação

A adulteração de óleos vegetais nobres a muito tempo é uma fonte de estudos devido ao alto valor agregado dos diferentes tipos de óleos. Com a finalidade de diminuir o custo desses óleos tornou-se prática a venda de misturas comerciais desses óleos comestíveis. Entretanto, essas misturas acabam tendo qualidades nutricionais diferentes dos óleos puros⁹³. Outro fato relevante é que os óleos de gergelim, azeites de oliva e óleos essenciais são comumente adulterados^{94,95}.

Em especial o óleo de gergelim vem ganhando mercado por apresentar em sua composição ligninas, tocoferóis e ácidos graxos insaturados⁹⁵. A adulteração de óleo de gergelim não é apenas uma fraude comercial, mas pode apresentar ameaça potencial à saúde humana⁹⁶. As adulterações mais frequentes são realizadas com óleos de milho, soja e canola, pois apresentam baixo valor agregado.

Por consequência, torna-se indispensável o uso de técnicas analíticas para a identificação e quantificação dos possíveis adulterantes de óleos vegetais. A presença de moléculas fluorescentes na composição desses óleos possibilita o uso da espectroscopia de fluorescência molecular para sua análise. Além disso, há possibilidade de utilização das superfícies de excitação-emissão de fluorescência (EEM – *Excitation-Emission Matrix*) para caracterização de diversos fluoróforos presentes no óleo de gergelim^{93,97}. A identificação desses compostos auxilia na correta identificação dos diferentes tipos de óleos que podem estar presentes nas amostras.

6.2. Objetivos

- Obtenção das superfícies de excitação-emissão para amostras de óleo de gergelim adulteradas;
- Desenvolver e validar modelos de calibração utilizando a restrição de trilinearidade do MCR – ALS e PARAFAC;
- Comparação entre os resultados obtidos para os dois métodos quimiométricos.

6.3. Procedimento Experimental

6.3.1. Preparo de Amostra

Foram obtidas duas marcas comerciais de óleo de gergelim comestível (marcas A e B), a marca A foi selecionada para ser a amostra a ser adulterada. Com

base na literatura⁹⁴ foi verificado que os óleos de canola, girassol, milho, soja e óleo de gergelim não-comestível são os mais utilizados em adulterações. Para cada adulterante foram preparadas dez amostras, sendo seis amostras para a construção da curva de calibração e quatro amostras para validação. A Tabela 10 apresenta a porcentagem massa/massa de cada uma das soluções preparadas. Para o preparo das amostras foram consideradas as densidades de cada óleo. Todas as amostras foram preparadas e acondicionadas em frascos transparentes de 10,00 mL.

Tabela 10. Porcentagem massa/massa de cada um dos óleos utilizados para construção das curvas de calibração e validação.

Curva	Amostra	% (m/m)	
		Óleo de Gergelim	Óleo Adulterante
Calibração	1	95	5
	2	85	15
	3	75	25
	4	65	35
	5	55	45
	6	50	50
Validação	7	90	10
	8	80	20
	9	70	30
	10	60	40

6.3.2. Equipamento Utilizado

Para a obtenção das superfícies de EEM foi utilizado o espectrofluorímetro PC1 (ISS, Champaign, Illinois, IL, EUA). Os espectros de excitação foram obtidos entre 250 a 550 nm com resolução espectral de 5 nm e os espectros de emissão foram obtidos entre 300 a 800 nm com resolução de 5 nm. Foram utilizadas fendas de 2,0 mm e foram coletados três espectros em cada ponto de medida. Durante as leituras 2,00 mL de amostra foram colocados em cubetas de quartzo.

6.3.3. Tratamento de Dados

6.3.3.1. Construção da Matriz de Dados

Os dados trilineares, como as superfícies EEM, necessitam de uma modificação em sua estrutura para que os métodos de resolução de curvas possam ser utilizados. Para o MCR – ALS utilizou-se uma matriz aumentada em direção a

matriz C (Figura 43a). E no caso do PARAFAC foi construído um tensor trilinear \underline{X} como representado na Figura 43b.

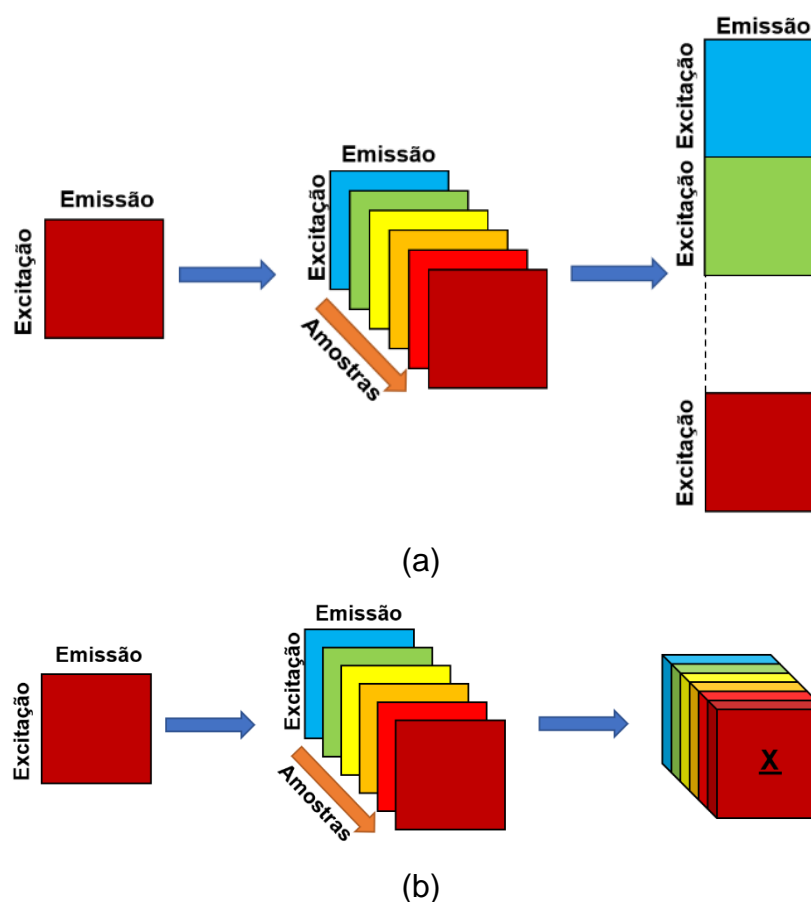


Figura 43. Construção da matriz de dados a partir da concatenação das superfícies de EEM das amostras: (a) MCR – ALS; (b) PARAFAC.

6.3.3.2. Pré-Processamentos

Normalmente os dados trilineares precisam ser normalizados para que os modelos obtidos com PARAFAC ou MCR – ALS possam ser melhor interpretados. A normalização dessas superfícies deve ser feita sempre que forem construídos modelos de calibração para que não haja diferença de intensidade entre as matrizes de dados. São propostos três tipos de pré-processamentos⁵⁸: (1) correção de erros sistemáticos no conjunto de dados, (2) remoção de sinais não relacionados com dados de fluorescência e (3) normalização dos dados.

- *Eliminação de dados não-trilineares*

Os dados trilineares como as superfícies EEM podem apresentar sinais diagonais que estão relacionados com dados não-trilineares (Figura 44). Os espalhamentos Rayleigh e Tyndall acontecem no comprimento de onda de excitação

e possuem maiores magnitudes que os sinais de fluorescência. Já o espalhamento Raman pode ocorrer comprimentos de onda ligeiramente maiores. Os espalhamentos Rayleigh e Raman de segunda ordem podem ocorrer no dobro do comprimento de onda dos espalhamentos primários. Os espalhamentos de maior ordem são bem menos intensos que os espalhamentos primários, e nem sempre são observados. A Figura 44 apresenta a superfície EEM antes e depois de serem removidos todos os dados não-trilineares. Para a remoção dos espalhamentos Rayleigh e Raman desses dados foi utilizada a rotina *eemscat*⁹⁸. Ela remove o espalhamento, assim como (opcionalmente) interpola as áreas removidas de tal forma que não há valores ausentes na nova superfície.

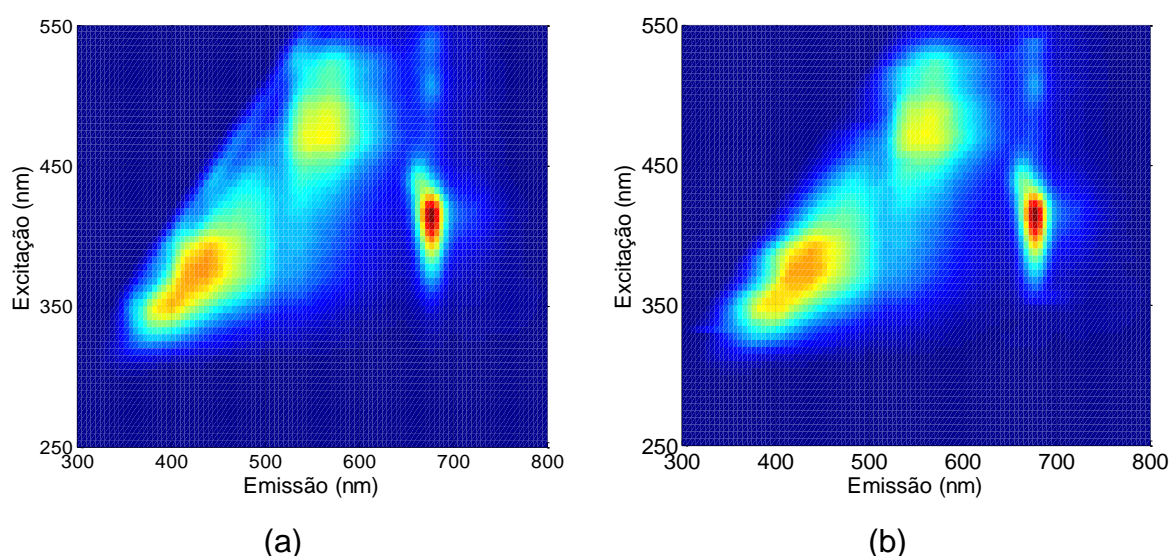


Figura 44. Representação dos espalhamentos Rayleigh e Raman em superfícies EEM.

- *Normalização*

A normalização dos sinais é necessária para todo o conjunto de dados em que houver gradiente de concentração, como no caso de diluições da amostra⁵⁸. Essa normalização é realizada pela divisão do conjunto de dados pela soma dos valores quadráticos de todas as variáveis da amostra. O *toolbox* drEEM⁵⁸ realiza essa etapa de normalização antes de serem criados os modelos.

6.3.3.3. *Determinação do Número de Fatores*

A última etapa de processamento dos dados, antes da construção dos modelos, é a determinação do número de componentes. Há diversas metodologias

que podem ser utilizadas para identificar o número de fatores em um conjunto de dados trilineares. O *toolbox* MCR – ALS GUI 2.0⁴⁸ utiliza a análise de valores singulares, já para o PARAFAC o mais utilizado é a metodologia chamada análise *split-half*, que foi incorporada por Murphy *et al*⁶⁸ no *toolbox* drEEM para análises de dados trilineares. Na Figura 45 está representado esquematicamente o funcionamento da análise *split-half*.

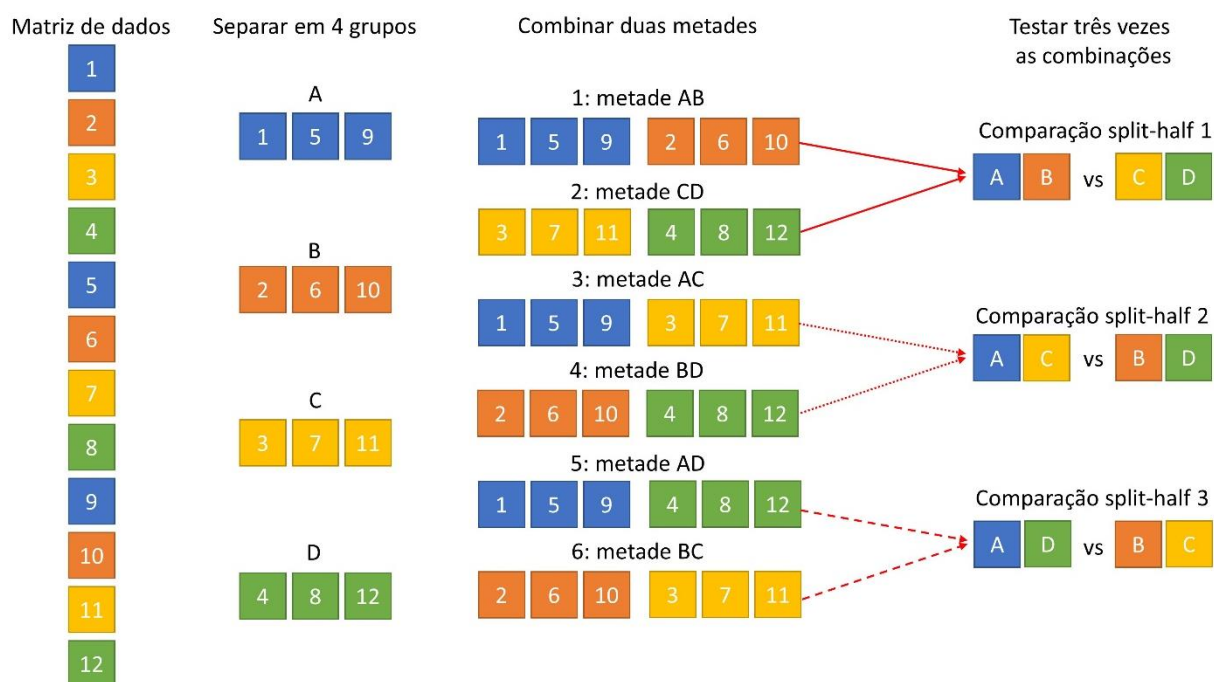


Figura 45. Representação esquemática da análise *split-half*, sendo que a nomenclatura para essa análise é 'S₄C₆T₃'. S simboliza o número de *splits* (pedaços) em que a matriz de dados foi separada, C é o número de combinações criadas e T o número de testes para as combinações propostas.

O algoritmo *split-half* divide o arranjo de dados em blocos (aleatoriamente) e após essa separação os blocos são combinados dois a dois. Os novos blocos formados são submetidos a sucessivos cálculos de PARAFAC e ao final os *loadings* (resultados referentes aos perfis espectrais obtidos) são comparados e o resultado ideal é obtido quando não há diferenças significativas entre os *loadings* obtidos. Na Figura 46 estão representados os *loadings* obtidos para o conjunto de amostras adulterados com óleo de milho, utilizando quatro e cinco fatores.

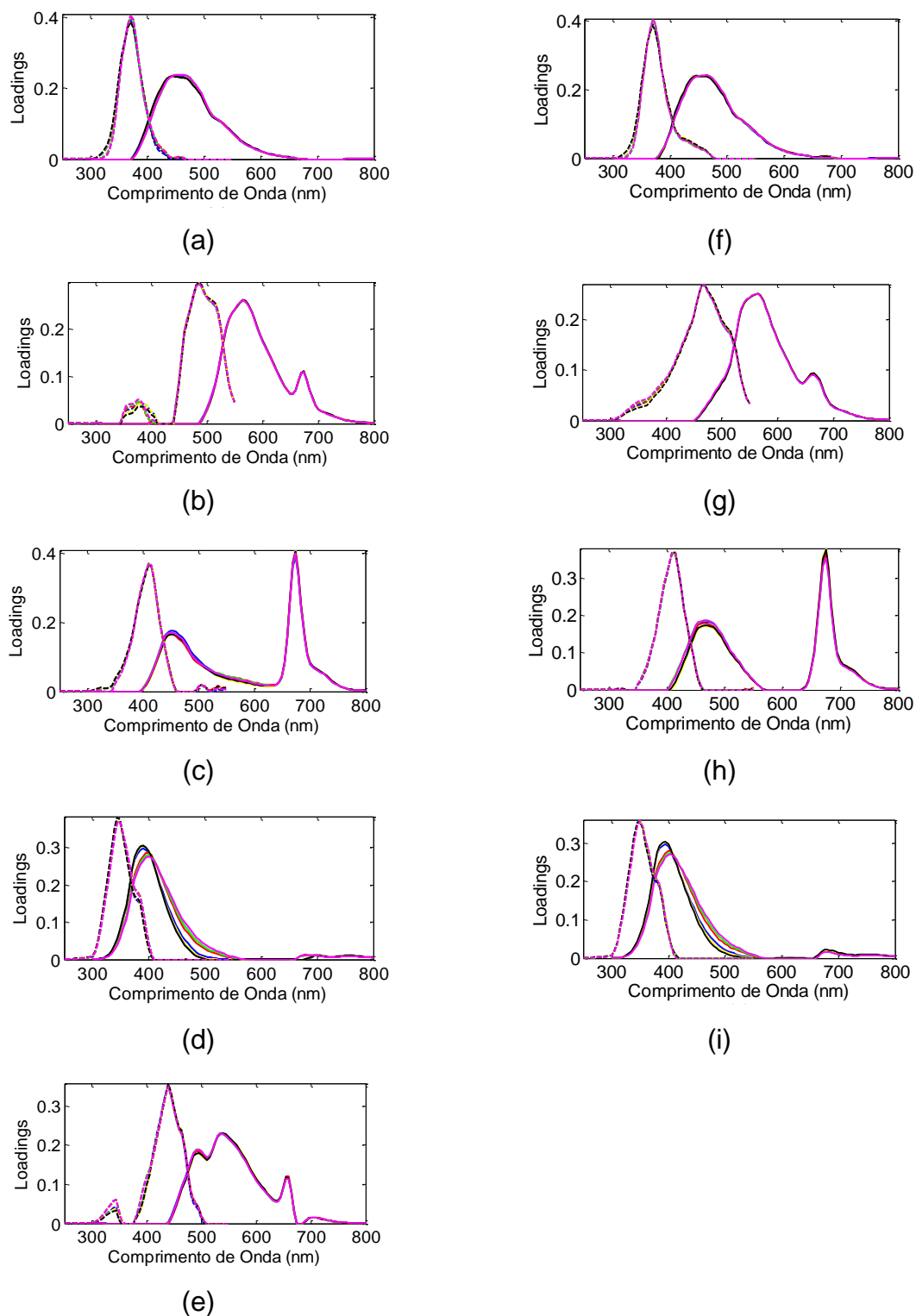


Figura 46. Determinação do número de fatores utilizando a análise *split-half*. Em (a) até (e) para cinco fatores e de (f) até i para quatro fatores. as curvas tracejadas e contínuas estão relacionadas com os espectros de excitação e emissão, respectivamente.

A análise *split-half* foi proposta por Harshman⁹⁹ e há duas formas em que os dados podem ser separados e analisados. Essas variações são $S_4C_4T_2$ e $S_4C_6T_3$, onde S representa o número de *splits*, C o número de combinações e T a quantidade de testes realizados.

6.3.3.4. Métodos de Resolução de Curvas

Com o número de fatores determinado pela análise *split-half* foram obtidos os modelos de calibração para cada um dos sistemas adulterados. As superfícies EEM são tidas como trilineares e para o tratamento correto dessas informações há necessidade da utilização de métodos quimiométricos mais específicos, como o PARAFAC^{13,56,57} e MCR – ALS com restrição de trilinearidade. Para o caso do MCR – ALS poderia ser utilizada a combinação entre as restrições de trilinearidade e de correlação, mas atualmente essa combinação ainda não é possível. No entanto, em 2016, Neves *et al*¹⁰⁰ publicaram um trabalho que apresenta a combinação das restrições de trilinearidade e de correlação, mas essa rotina ainda não foi disponibilizada.

As Figuras 47 e 48 apresentam a decomposição do conjunto de dados para o PARAFAC e MCR – ALS com restrição de trilinearidade, respectivamente. As curvas de calibração foram obtidas a partir do ajuste linear dos valores de concentração de referência *versus* os valores obtidos pelos escores do MCR – ALS e PARAFAC.

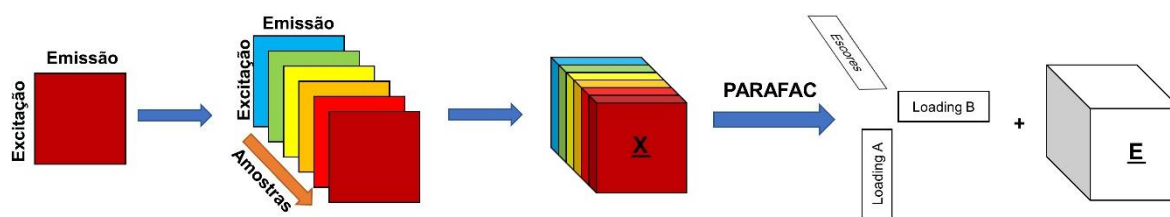


Figura 47. Decomposição da matriz pelo PARAFAC em pesos e escores.

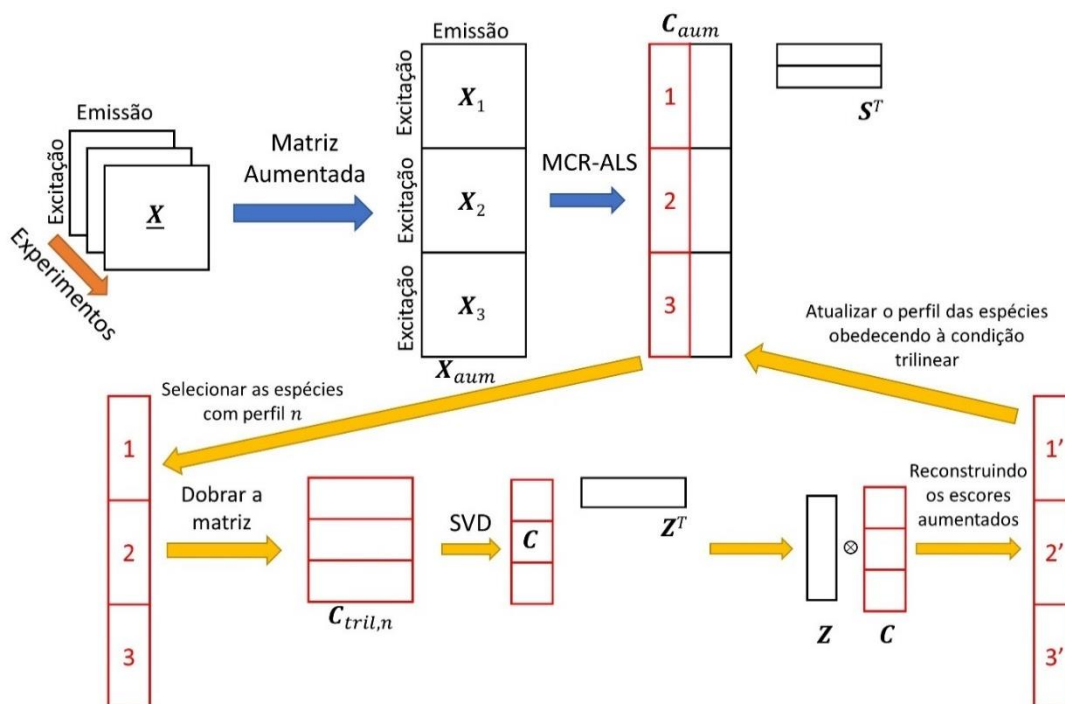
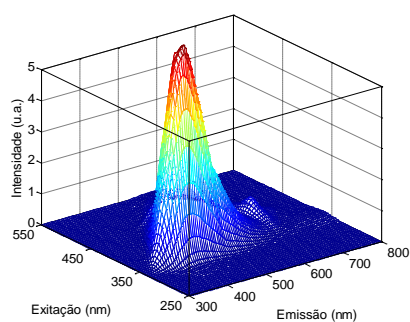


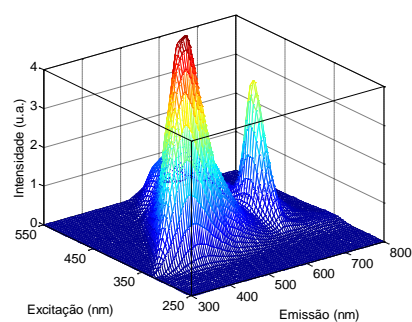
Figura 48. Decomposição da matriz de dados com o MCR – ALS com a restrição de trilinearidade.

6.4. Resultados e Discussão

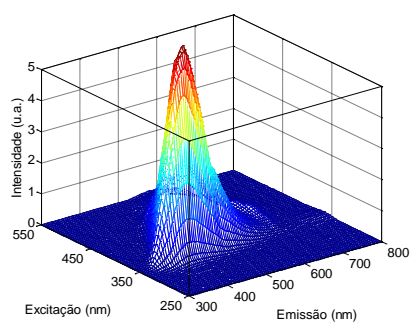
A Figura 49 apresenta as superfícies de EEM dos óleos puros, mostrando que há diferenças espectrais entre os óleos analisados e indicando que é possível a construção dos modelos de calibração pretendidos. Porém, observa-se também que embora alguns desses espectros sejam diferentes há diversas semelhanças entre eles, o que poderá acarretar dificuldades na identificação dos tipos de óleos utilizados como adulterantes. A maior diferença se encontra na região de 700 nm do eixo de emissão, sendo que essa região está ligada com a emissão da clorofila. Pode-se observar que os óleos de gergelim comestíveis e soja possuem uma emissão mais intensa nessa região que os outros óleos estudados.



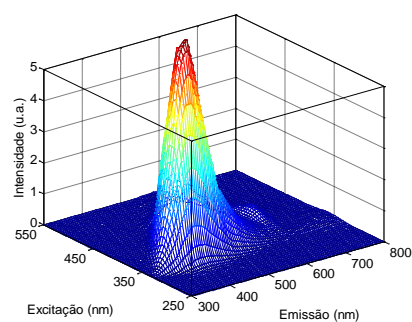
(a)



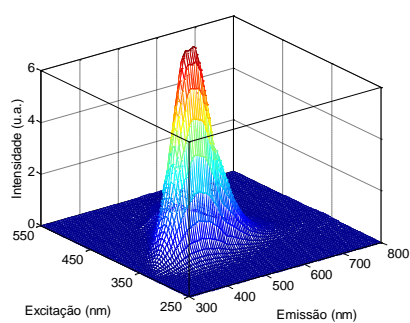
(b)



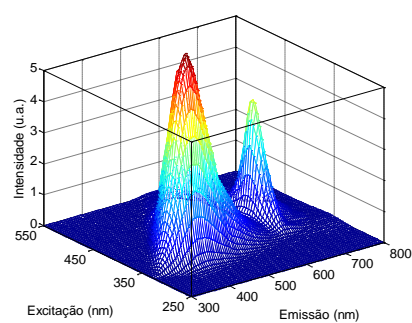
(c)



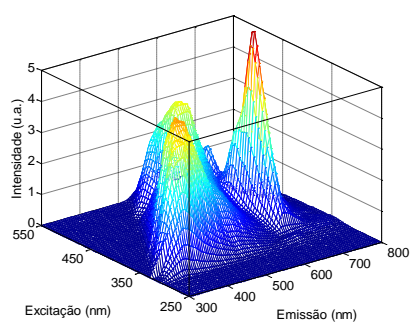
(d)



(e)



(f)



(g)

Figura 49. Superfícies EEM normalizadas dos óleos puros: (a) canola; (b) gergelim não comestível; (c) gergelim comestível marca B; (d) girassol; (e) milho; (f) soja; (g) gergelim comestível marca A.

Os óleos de canola, girassol e soja possuem banda em 700 nm que corresponde à emissão da clorofila e causa superposição de sinais com o óleo de gergelim que também possui essa banda, impedindo que métodos univariados sejam utilizados. Por outro lado, o óleo de soja não possui essa banda e nesse caso calibração univariada poderia ser utilizada.

Serão apresentados a seguir resultados para milho e soja, sendo que para os outros apenas os resultados serão apresentados numa tabela.

A Figura 50 apresenta os resultados de SVD e *split-half* para o óleo de gergelim adulterado com milho. A Figura 50a apresenta o resultado da análise de valores singulares, essa etapa é realizada pelo MCR – ALS. Normalmente o resultado da SVD pode ser utilizado diretamente para se determinar quantos componentes estão presentes no conjunto de dados. No entanto, os óleos vegetais são compostos de diversas estruturas fluorescentes o que acaba dificultando a determinação de quantos componentes químicos são responsáveis pela formação da superfície EEM. Por causa disso, utilizou-se a análise *split-half* para se confirmar qual seria o número ideal de fatores para a decomposição desses dados

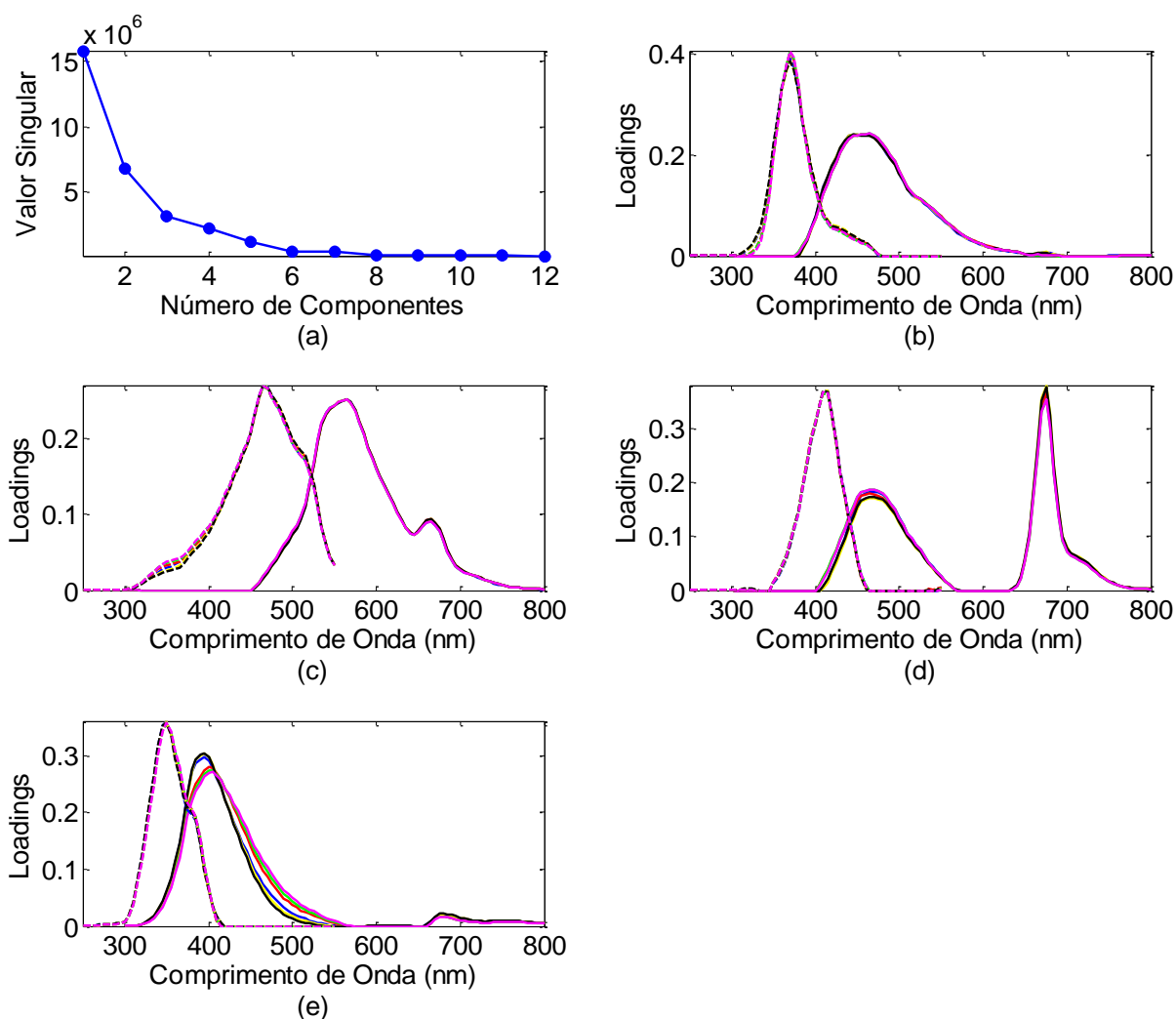


Figura 50. Determinação do número de fatores para o conjunto de amostras adulteradas com óleo de milho utilizando (a) SVD e (b) até (e) análise *split-half* para quatro fatores.

Com base nas Figuras 50b-e, é possível identificar que começa a haver uma alteração nos *loadings* obtidos quando foram utilizados quatro componentes, sendo que o último fator apresentou maiores diferenças entre os *loadings* obtidos e com isso conclui-se que devem ser utilizados 4 fatores nos cálculos. Como apresentado na Figura 46, se fossem utilizados 5 fatores haveria mais de um fator com diferença entre os *loadings* obtidos.

O uso de um maior número de componentes pode acarretar melhora nos resultados de concentração, entretanto, pode não ser possível a correta interpretação dos *loadings* obtidos ao final da aplicação dos métodos. Além disso, com maior

número de componentes os modelos podem ser superajustados levando a resultados pouco confiáveis.

Ao final da etapa iterativa foram obtidos três resultados em ambos os algoritmos: proporções químicas, espectros de excitação e espectros de emissão. A literatura é rica na interpretação das regiões de absorção e emissão de óleos vegetais. Com base nesses trabalhos foi possível a identificação de três dos componentes mais importantes identificados nos resultados dos modelos obtidos.

A Figura 51 apresenta os espectros recuperados pelo MCR – ALS (Figuras 51a, 51c, 51e e 51g) e PARAFAC (Figuras 51b, 51d, 51f e 51h) para as amostras adulteradas com óleo de milho. Já a Figura 52 apresenta os espectros recuperados para as amostras adulteradas com o óleo de soja.

Nas Figuras 51 e 52 os espectros recuperados em azul e vermelho correspondem respectivamente às absorções e emissões dos componentes identificados. Essas três componentes estão relacionadas com tocoferóis e fenóis (Figuras 51a e 52a), compostos oxidativos (Figuras 51b e 52b) e pigmentos de clorofila (Figuras 51c e 52c). Não foi possível a identificação da quarta componente (Figuras 51d e 52d), mas pela região em que se encontram os seus espectros de emissão e excitação leva-se a crer que ela está relacionada com outros tipos de tocoferóis e fenóis presentes nas amostras.

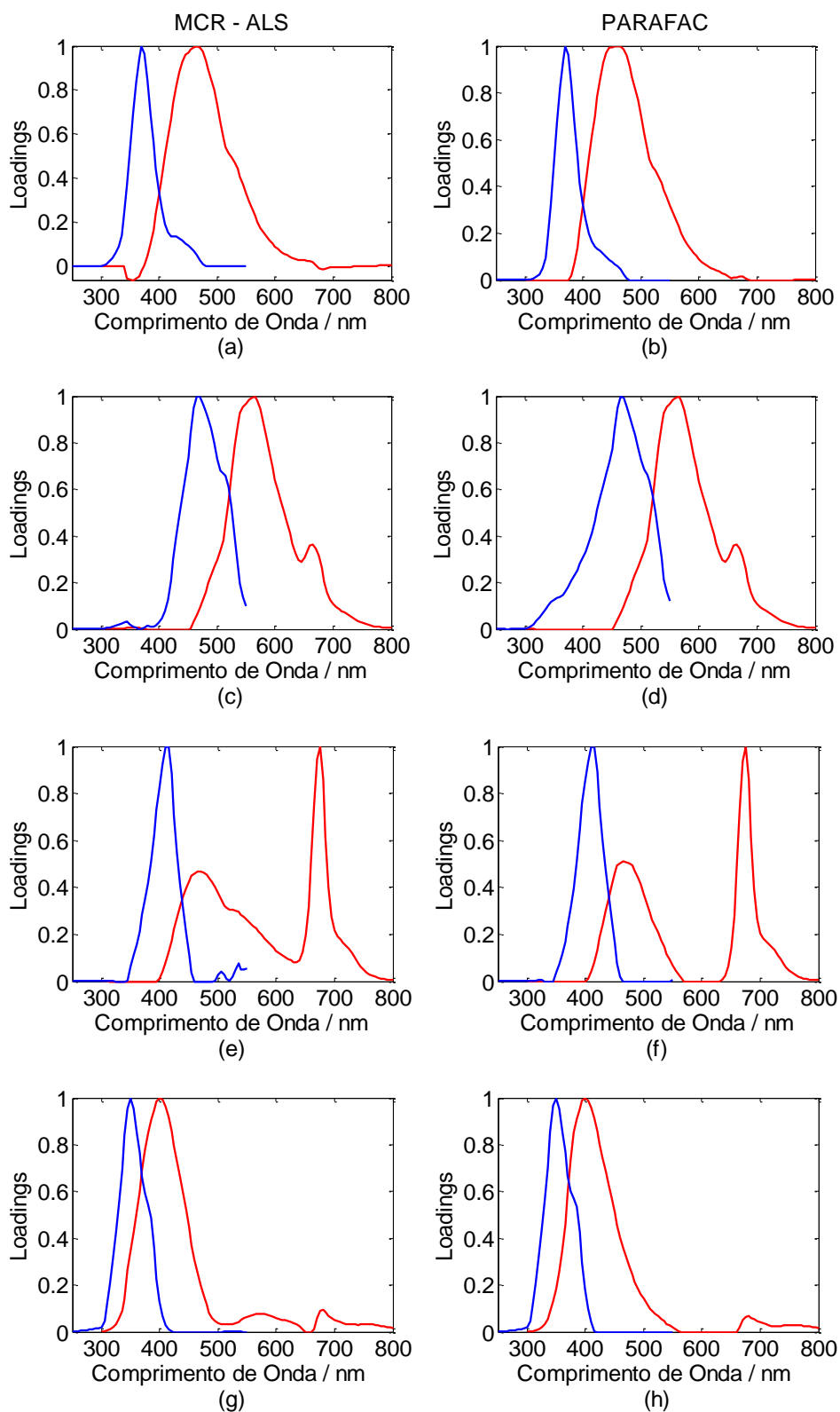


Figura 51. Resultados do MCR – ALS e PARAFAC para as amostras adulteradas com óleo de milho. Espectros de excitação (azul) e emissão (vermelho) para: (a) tocoferóis e fenóis; (b) compostos oxidativos; (c) pigmentos de clorofila; (d) possivelmente relacionado com os diferentes tocoferóis e fenóis presentes.

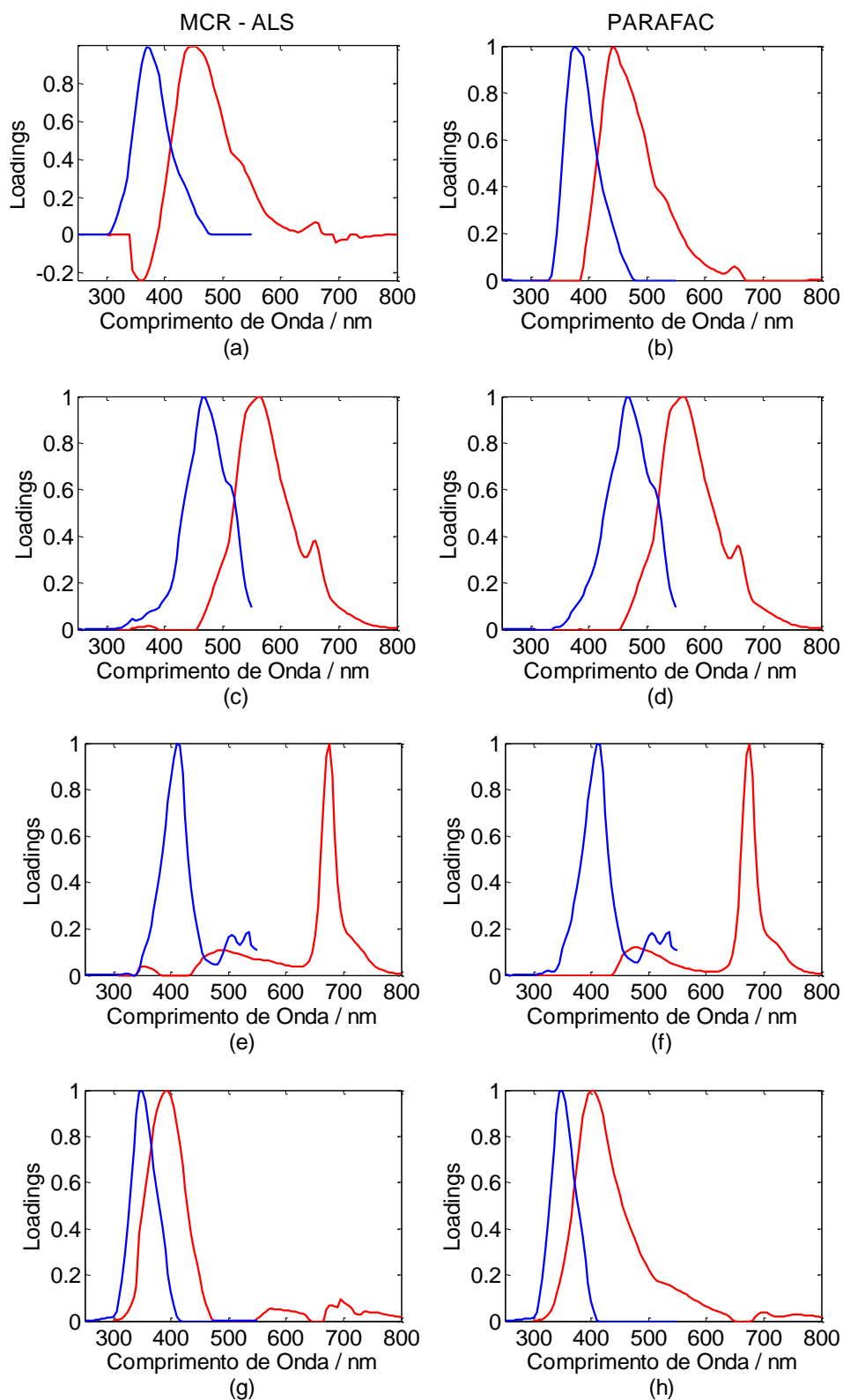


Figura 52. Resultados do MCR – ALS e PARAFAC para as amostras adulteradas com óleo de soja. Espectros de excitação (azul) e emissão (vermelho) para: (a) tocoferóis e fenóis; (b) compostos oxidativos; (c) pigmentos de clorofila; (d) possivelmente relacionado com os diferentes tocoferóis e fenóis presentes.

Observando as Figuras 51a e 52a foi possível verificar que os espectros de emissão recuperados possuem um perfil negativo para comprimentos de onda inferiores a 400 nm. Todas as otimizações com o MCR – ALS foram realizadas com a restrição de não-negatividade, mas mesmo assim foram obtidas regiões negativas. O principal motivo desses resultados está ligado ao uso da restrição de trilinearidade. Uma das etapas dessa restrição é realizar uma análise de valores singulares na matriz **C** durante a etapa iterativa, nesse processo são geradas matrizes que possuem valores negativos. Com isso, mesmo com a restrição de não-negatividade é possível que a matriz final possua valores negativos.

Os resultados apresentados nas Figuras 51 e 52 correspondem apenas a dois dos modos (emissão e excitação) obtidos pelo MCR – ALS e PARAFAC. O terceiro modo está relacionado com o perfil de concentração dos componentes presentes nas amostras. As Figuras 53 e 54 apresentam os resultados das curvas de calibração para as amostras adulteradas com milho e soja, respectivamente.

A construção dessas curvas foi realizada com o uso de apenas um fator para cada óleo. Para o óleo de gergelim foi utilizado o fator apresentado pela Figura 51c e para os demais adulterantes o fator da Figura 51a. Essa seleção foi realizada com base nas superfícies apresentadas na Figura 49, onde é possível observar algumas diferenças entre os espectros dos óleos puros.

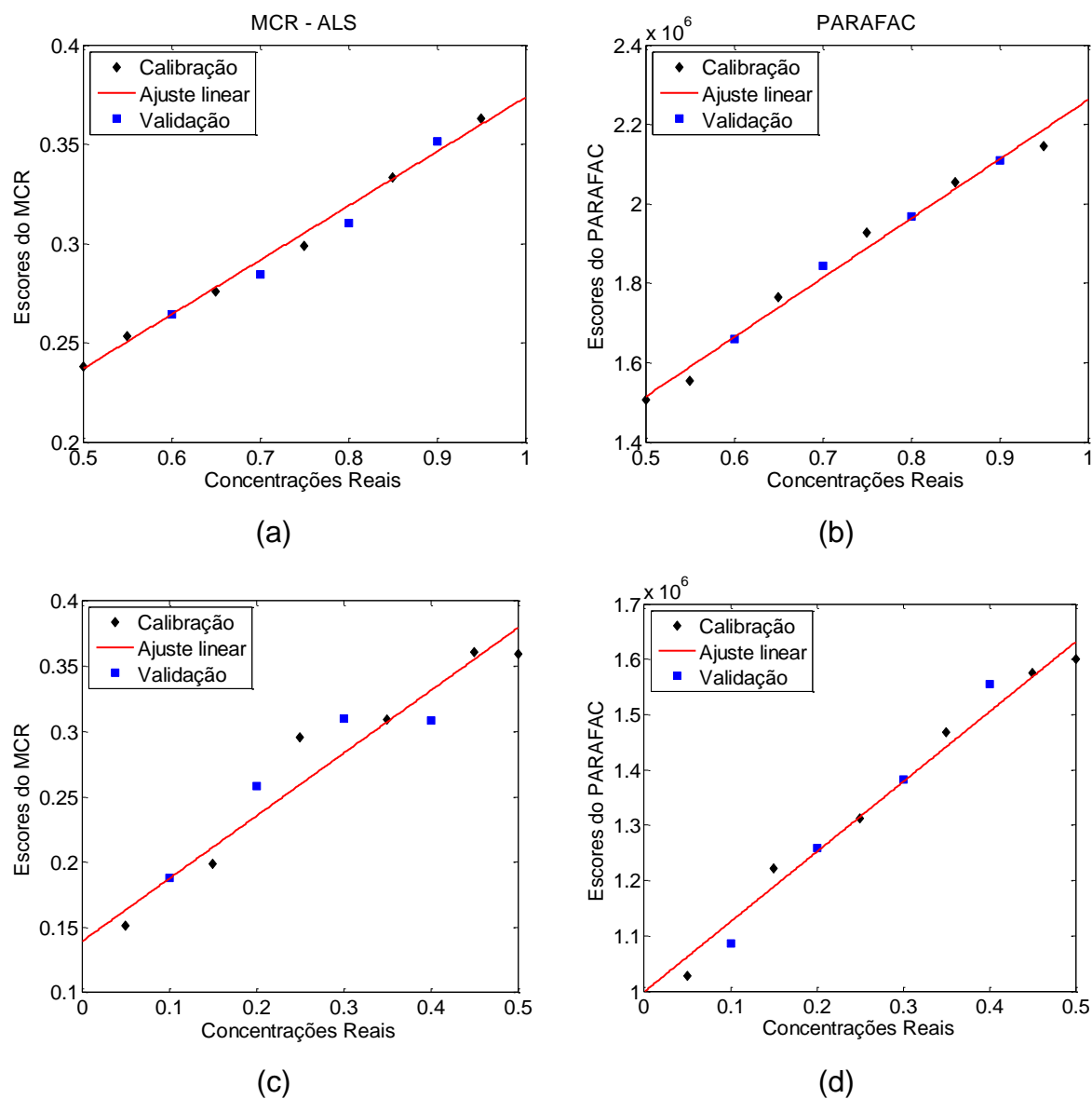


Figura 53. Curvas de calibração para as amostras de óleo de gergelim adulterado com óleo de milho. (a) óleo de gergelim pelo MCR-ALS, (b) óleo gergelim pelo PARAFAC, (c) óleo de milho pelo MCR-ALS, (d) óleo de milho pelo PARAFAC.

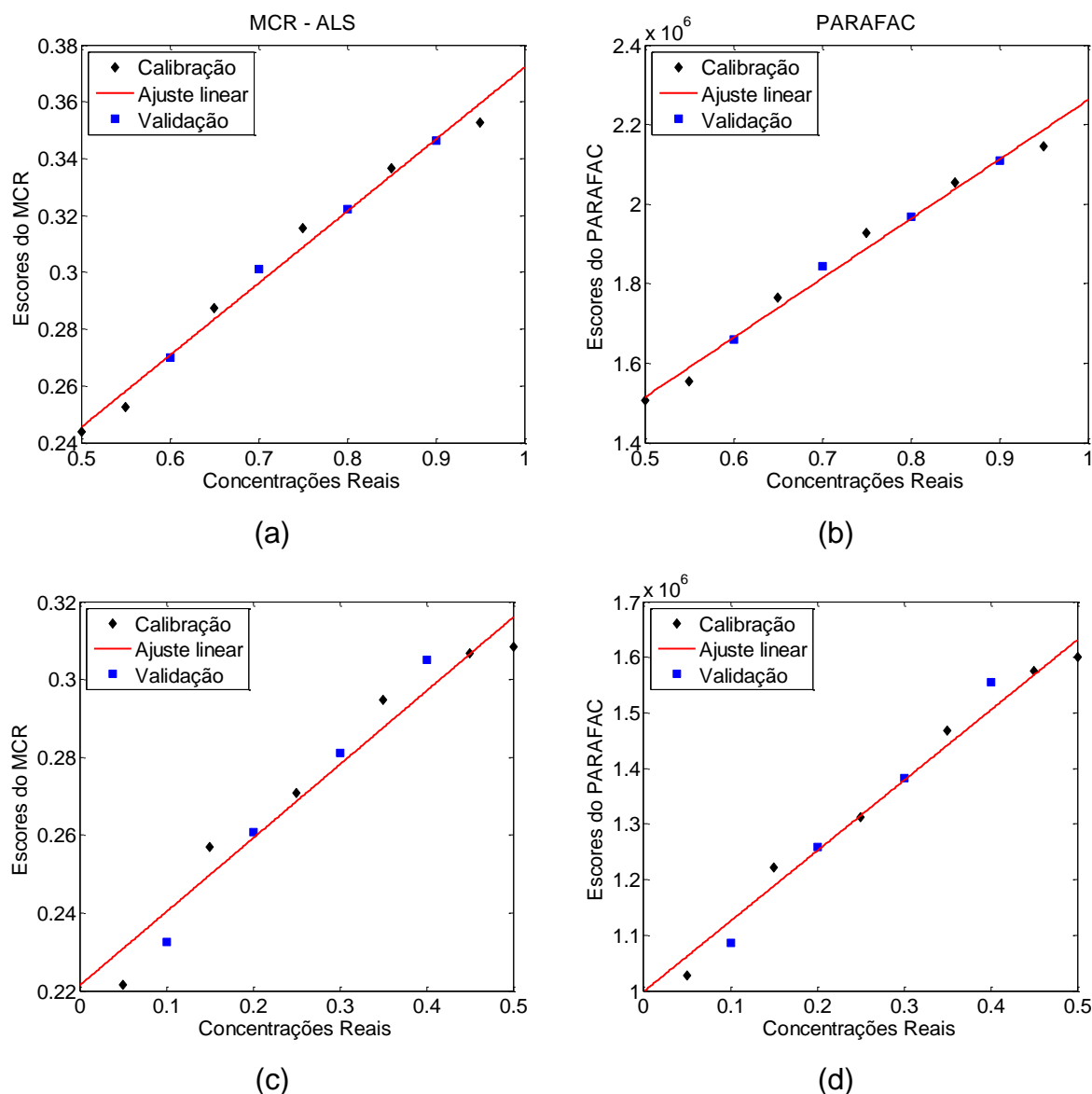


Figura 54. Curvas de calibração para as amostras de óleo de gergelim adulterado com óleo de soja. (a) óleo gergelim pelo MCR-ALS, (b) óleo gergelim pelo PARAFAC, (c) óleo de soja pelo MCR-ALS, (d) óleo de soja pelo PARAFAC.

Com base nas curvas apresentadas nas Figuras 51 e 52 foram calculados o coeficiente de determinação e os erros de calibração (RMSEC) e previsão (RMSEP) para todos os conjuntos de dados, que estão apresentados na Tabela 11. Nessa Tabela estão representados os parâmetros modelos com os resultados do MCR – ALS e PARAFAC para cada conjunto de dados.

Com os resultados obtidos, percebeu-se que não era possível a correta identificação do óleo utilizado como adulterante, devido à grande similaridade entre as superfícies de emissão-excitação dos óleos. Afim de confirmar esse fato foram

construídas matrizes de dados em que o conjunto de calibração era composto por amostras adulteradas com óleo de milho e o conjunto de validação com amostras adulteradas com óleo de soja. Os resultados obtidos foram similares aos apresentados nas Figuras 51 a 54, indicando que o modelo não consegue diferenciar os óleos.

Mesmo os métodos trilineares não sendo capazes de identificarem quais seriam os adulterantes eles indicam que as amostras estavam adulteradas com outro tipo de óleo. Esse fato pode ser confirmado pelos valores de RMSEC e RMSEP, já que em todos os casos eles foram inferiores a 5,50%. Esses baixos valores indicam que as superfícies EEM poderiam ser utilizadas para quantificação de adulterantes em óleo de gergelim.

A Tabela 11 apresenta os resultados gerais para os estudos com todos os adulterantes. Os coeficientes de determinação foram todos próximos ou superiores a 0,9500 e RMSEPs na faixa 1,07 a 5,40%. Esses valores indicam que há alta correlação entre os escores do MCR ou PARAFAC com os valores de referência. Com isso, pode-se afirmar que os modelos construídos são adequados para o conjunto de dados estudado.

No caso das amostras adulteradas com óleo de gergelim não comestível os melhores resultados foram obtidos com MCR – ALS, principalmente para a curva de calibração do óleo de gergelim da marca A, em que foram obtidos valores de RMSEC e RMSEP na ordem de 2,00%. O PARAFAC, por sua vez, resultou em resultados superiores a 4,30% para esse mesmo conjunto de dados. Para as amostras adulteradas com óleo de gergelim comestível (marca B), não foi possível a identificação de nenhum fator que poderia ser correlacionado com a variação de concentração das espécies presentes nas amostras. Isso pode estar relacionado com a grande similaridade existente entre os dois óleos comestíveis utilizados.

Tabela 11. Resultados para as curvas de calibração de todos os conjuntos de amostras para os dois algoritmos.

Algoritmo	Tipo de Óleo	RMSEC (%)	RMSEP (%)	Coefficiente de Determinação	
MCR – ALS	Canola	3,69	3,54	0,9741	
	Gergelim marca A	3,84	5,40	0,9721	
	Gergelim não comestível	5,24	5,30	0,9599	
	Gergelim marca A	2,08	2,00	0,9533	
	Gergelim marca B	-	-	-	
	Gergelim marca A	-	-	-	
	Girassol	1,44	1,97	0,9959	
	Gergelim marca A	1,48	1,45	0,9957	
	Milho	3,86	4,36	0,9719	
	Gergelim marca A	1,20	2,24	0,9972	
	Soja	3,43	3,08	0,9776	
	Gergelim marca A	1,94	1,03	0,9927	
	PARAFAC	Canola	3,13	2,76	0,9813
		Gergelim marca A	3,77	5,40	0,9731
Gergelim não comestível		4,65	5,01	0,9498	
Gergelim marca A		5,05	4,35	0,9916	
Gergelim marca B		-	-	-	
Gergelim marca A		-	-	-	
Girassol		1,31	1,37	0,9966	
Gergelim marca A		1,50	1,52	0,9959	
Milho		3,75	4,21	0,9733	
Gergelim marca A		1,29	2,31	0,9967	
Soja		2,04	2,53	0,9919	
Gergelim marca A		2,02	1,07	0,9920	

6.5. Conclusões Parciais

Foram obtidos resultados satisfatórios para as curvas de calibração por ambos os métodos quimiométricos. Os valores de RMSEC, RMSEP e coeficiente de determinação indicam que os modelos construídos possuem altas correlações com os valores de concentrações de referência. A única desvantagem foi a não identificação de quais seriam os adulterantes se as análises fossem realizadas sem o conhecimento dos compostos presentes nas amostras.

Já o uso de dados trilineares trouxe informações que não poderiam ser obtidas de forma univariada. A banda em 700 nm (Figura 49) corresponde

exclusivamente a clorofila e se fossem realizados experimentos univariados seria possível encontrar uma curva de calibração para o óleo de gergelim, mas o mesmo não pode ser dito para os outros óleos, visto que o fator que representa o aumento de concentração em cada uma das amostras não seria observado.

Por fim, os modelos propostos pelo PARAFAC e MCR-ALS podem ser considerados equivalentes. Entretanto, do ponto de vista espectroscópico os melhores resultados foram obtidos pelo PARAFAC, sendo que nenhum dos espectros obtidos apresentaram valores negativos. O mesmo não pode ser observado para o MCR (Figuras 51 e 52) sendo que há uma região abaixo de 400 nm com *loadings* negativos.

**CAPÍTULO 7 – DETERMINAÇÃO DAS
CONSTANTES DE DESPROTONAÇÃO
DO 2-(2'-HIDROXIFENIL)BENZOTIAZOL**

7.1. Introdução teórica e Motivação

Os equilíbrios químicos baseados em conceitos ácido-base podem ser descritos pela teoria de Brønsted-Lowry. Essas teorias ajudam a explicar as reações de transferência de próton. Essas reações possuem aplicações em catálise¹⁰¹, bioquímica, sensores ópticos, eletrônica orgânica e armazenamento de energia¹⁰². Sendo que os processos de transferência de prótons no estado excitado são mais interessantes devido às suas aplicações em fotofísica e fotoquímica, que são úteis para compreender reações como a fotossíntese e produção de ATP.

Fotofísica é o estudo dos processos de absorção e emissão de radiação, observando as diferentes escalas de tempo e diferentes condições experimentais, com a finalidade de entender o comportamento dos materiais. Já a fotoquímica é o estudo de processos químicos que ocorrem devido a absorção de radiação. O estudo de sistemas fotoquímicos que usam a luz solar para conduzir reações químicas importantes ou para gerar eletricidade é de grande importância para o desenvolvimento de fontes sustentáveis de energia¹¹.

A transferência de próton no estado eletrônico excitado (ESPT – *Excited State Proton Transfer*) é impulsionada por diversas forças. Muitas vezes ela está relacionada com reações acopladas, como isomerização, a formação de compostos cíclicos, processos redox e foto-tautomerização¹⁰³. Independentemente da reação subsequente, a fotoacidez é uma característica comum, geralmente definida como o realce da acidez no estado excitado devido ao rearranjo eletrônico da densidade após a excitação. Essa acidez pode ser estimada pelo ciclo termodinâmico proposto por Förster (Equações 17 e 18).

$$\Delta pK_a^* = pK_a - pK_a^* \quad \text{Equação 17}$$

$$\Delta pK_a^* = \frac{1}{2RT \ln 10} [(h\nu_{AHabs} + h\nu_{AHem}) - (h\nu_{Aabs} + h\nu_{Aem})] \quad \text{Equação 18}$$

Os termos pK_a e pK_a^* são chamadas de constantes de desprotonação nos estados eletrônicos fundamental e excitado, respectivamente. Elas são calculadas a partir de $h\nu_{AHabs}$ e $h\nu_{Aabs}$ que representam as energias da transição da absorção do fotoácido AH e de sua base conjugada A, sendo que suas as energias da transição de fluorescência $h\nu_{AHem}$ e $h\nu_{Aem}$ são as energias de transição das espécies AH e A, respectivamente. R é a constante de gás ideal e T é a temperatura.

Considerando o acoplado ao processo de fototautomerização da transferência de próton intramolecular no estado excitado (ESIPT – *Excited State Intramolecular Proton Transfer*, Figura 55), os derivados da benzotiazol, como por exemplo, o 2-(2'-hidroxifenil)benzotiazol (HBT) são uma classe de moléculas orgânicas emissoras de grande importância. Essas moléculas são chamadas de espécies enol-CIS, E_I^* , pois no estado eletrônico excitado sofrem transferência de próton ultrarrápida para produzir um tautômero, K^* . Essa tautomerização gera grande deslocamento Stokes (geralmente de 6.000 cm^{-1} a 12.000 cm^{-1}). Essa transição é impossível ocorrer no estado fundamental, visto que não é energeticamente favorável.

O enol-CIS, E_I^* , possui a conformação mais estável em solventes apolares e apróticos, e, também é a única configuração reativa da molécula. Entretanto, espera-se que diferentes conformações coexistam em equilíbrio, incluindo o enol-trans, E_{II}^* (para $X = S$ ou O) e enol-trans aberto, E_{III}^* (para $X = NH$). Além disso, em meios polares ou próticos, o enol-cis aberto, E_{IV}^* , pode ser observado. Todas as conformações contribuem para absorção e emissão dos enóis¹⁰⁴.

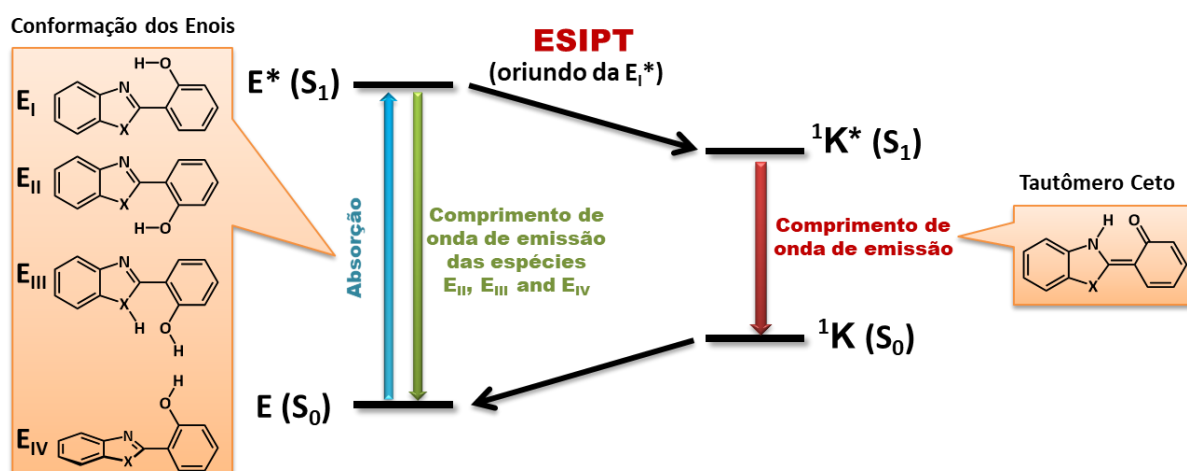


Figura 55. Ciclo fotoquímico do processo de ESIPT a 2-(2'-hidroxifenil)benzotiazol e seus enóis derivados.

7.2. Objetivos

- Obtenção das superfícies de excitação-emissão para HBT em soluções superácidas até soluções alcalinas;
- Desenvolver um modelo no MCR – ALS com a restrição de trilinearidade para a obtenção dos espectros de excitação e emissão HBT;

- Determinar os valores das constantes de desprotonação do HBT na faixa analítica selecionada;
- Desenvolver um modelo no MCR – ALS sem a restrição de trilinearidade para a obtenção do espectro de emissão do HBT e seus valores de constantes de desprotonação.

7.3. Procedimento Experimental

7.3.1. Preparo de Amostra

Foram preparadas 26 soluções extremamente ácidas com H_0^{105} na faixa de -3,78 a 1,00 (seguem a acidez de Hammett^{106,107}). Também foram preparadas 21 soluções tampão na faixa de pH de 1,50 a 11,50. O Anexo 4 apresenta o procedimento experimental seguido para alcançar os valores de pH sugeridos. Em cada uma das amostras a concentração total de HBT e derivados foi de 8,50 $\mu\text{mol/L}$.

7.3.2. Equipamento Utilizado

Para a obtenção das superfícies de EEM foi utilizado o espectrofluorímetro PC1 (ISS, Champaign, Illinois, IL, EUA). Os espectros de excitação foram obtidos entre 260 a 400 nm com resolução espectral de 2,5 nm e os espectros de emissão foram obtidos entre 340 a 630 nm com resolução de 2,5 nm. Foram utilizadas fendas de 2,0 mm e foram coletados três espectros em cada ponto de medida. Durante as leituras 2,00 mL de amostra foram colocados em cubetas de quartzo.

7.3.3. Tratamento de Dados

Ao final da aquisição das 47 superfícies EEM os dados tiveram a remoção dos espalhamentos Rayleigh e Raman e posteriormente normalizados. A determinação do número de componentes foi realizada com base na análise *split-half*. O MCR – ALS com restrição de não-negatividade foi utilizado para a decomposição do conjunto de dados. Todos esses procedimentos foram semelhantes aos apresentados na seção 6.3.3.

7.4. Resultados e Discussão

A Figura 56 apresenta três exemplos das superfícies que foram obtidas experimentalmente. Cada uma dessas superfícies foi obtida em diferentes soluções ácidas: H_0 -3,78 (Figura 56a); pH 5,00 (Figura 56b); pH 11,50 (Figura 56c).

Observando essas superfícies fica evidente que com a diminuição da acidez das soluções há mudanças nos perfis espectrais de excitação e emissão. Essas mudanças são causadas pelos equilíbrios de desprotonação que estão ocorrendo. A Figura 56d apresenta uma visão de todo o conjunto de dados, apresentando todas as 47 superfícies obtidas experimentalmente.

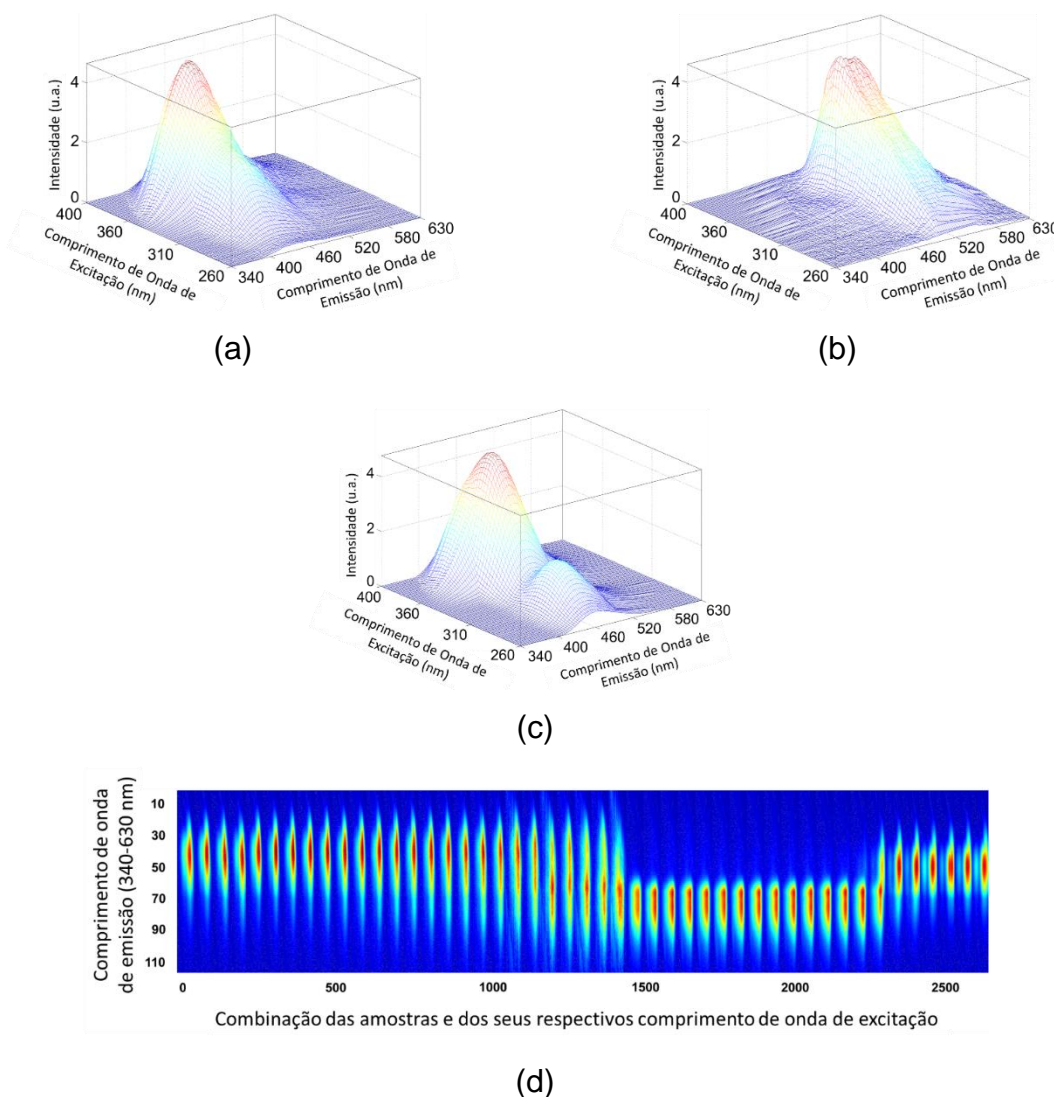


Figura 56. Superfícies de EEM em diferentes concentrações ácidas: (a) H_0 -3,78; (b) pH 5,00; (c) pH 11,50; (d) matriz de dados completa, no eixo y está representado o perfil de emissão e no eixo x é representada a multiplicação entre as 47 amostras e todos os comprimentos de onda de excitação.

As superfícies EEM podem ser tratadas como dados trilineares e, também, como uma única matriz. Nessa matriz utiliza-se apenas um espectro de emissão de cada superfície, mantendo-se fixo o comprimento de onda de excitação. Para o conjunto de dados trilineares serão obtidos como resultados os espectros de excitação

e emissão e as proporções químicas de todos os componentes presentes nas amostras.

Para o caso de transformar os dados em matriz, como descrito anteriormente, não serão obtidos os espectros de excitação dos componentes químicos. Entretanto, em tese, os espectros de emissão e as proporções químicas serão equivalentes às obtidas com os dados trilineares.

Independentemente de qual tratamento de dados for realizado foram retiradas todas as informações referentes aos espalhamentos Raman e Rayleigh, como descrito na seção 6.3.3.2.

Neste estudo, foi determinada *por split-half* a presença de três componentes presentes no conjunto de dados. Na literatura^{108,109} são encontradas informações que indicam que para esse sistema seriam observadas mais que três espécies químicas. No entanto, para a observação dessas outras espécies seria necessária a realização de experimentos com equipamentos capazes de medir o tempo de vida dessas espécies no estado excitado.

A Figura 57 apresenta as estruturas e equilíbrios das espécies que não estão no estado excitado que já foram reportados na literatura para esse sistema.

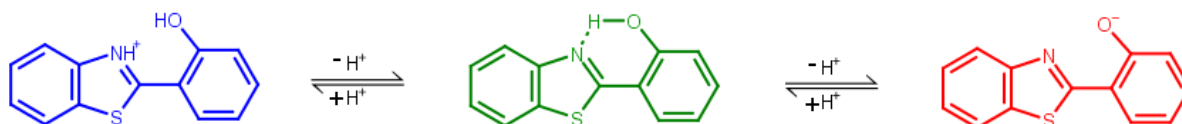


Figura 57. Equilíbrios químicos reportados na literatura para o HBT e seus derivados no estado fundamental.

Após a otimização do MCR – ALS com as restrições de trilinearidade e não-negatividade (para **C** e **S**), foram obtidos os espectros de excitação e emissão (Figura 58) para as três estruturas da Figura 58. Os espectros de excitação foram obtidos a partir da decomposição da matriz **C** do MCR – ALS, como explicado na seção 6.3.3.4. Já os espectros de emissão foram obtidos na matriz **S** do MCR – ALS. Não foi possível uma comparação direta desses perfis espectrais com espectros obtidos em laboratório, mas eles foram comparados com os encontrados na literatura. Essa comparação demonstrou que os espectros recuperados pelo MCR – ALS eram muito semelhantes aos espectros da literatura, indicando uma correta decomposição dos dados pelo MCR-ALS.

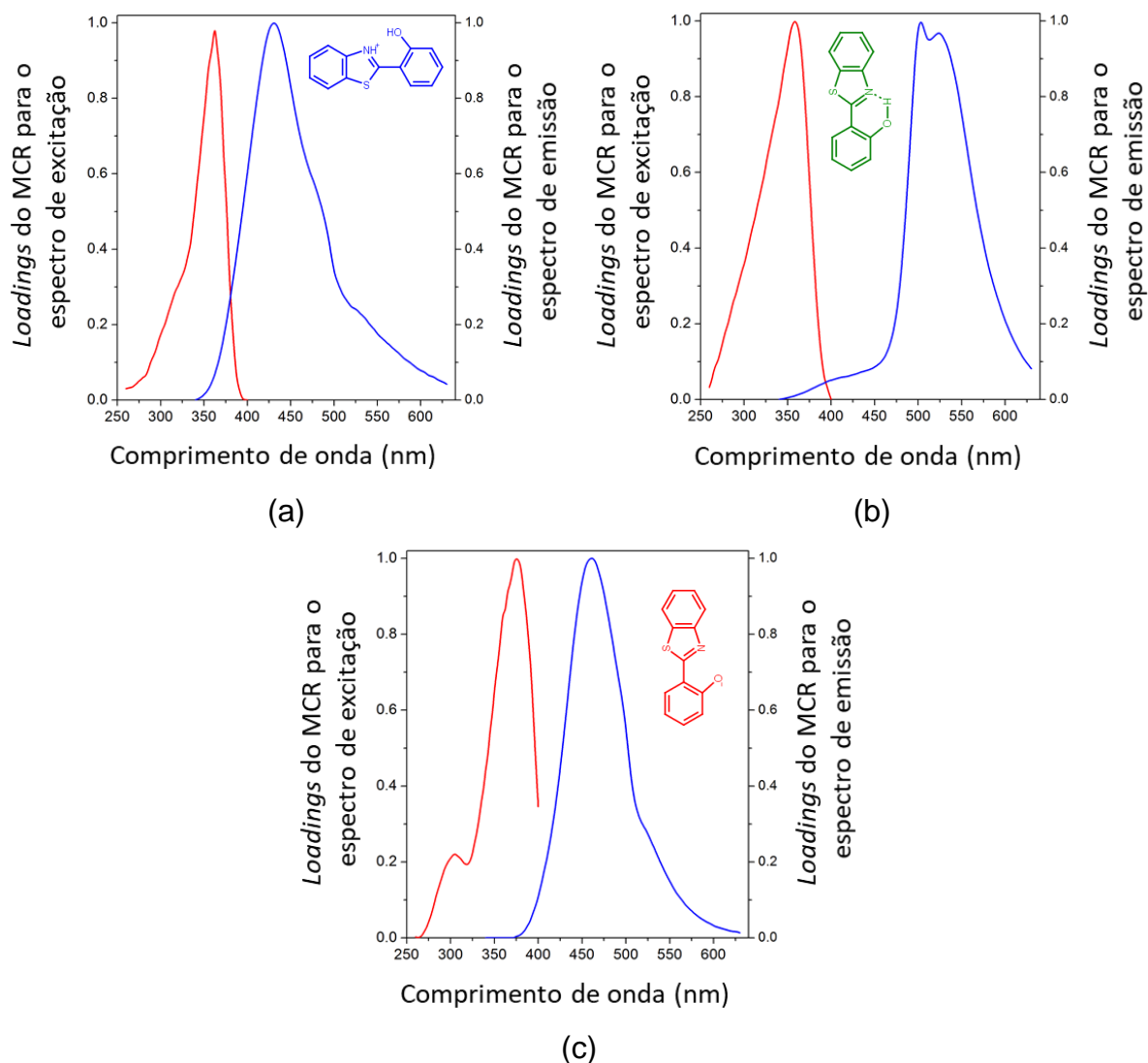


Figura 58. Espectros de excitação (em vermelho) e emissão (em azul) recuperados pelo MCR – ALS para as três estruturas químicas apresentadas na Figura 55: (a) HBT protonado; (b) HBT neutro; (c) HBT desprotonado.

Os equilíbrios são conhecidos há muitos anos e para o HBT e seus derivados esse equilíbrio já foi amplamente estudado. Entretanto, não foram demonstrados quais seriam os espectros de excitação das estruturas apresentadas na Figura 57, e como apresentado na Figura 58, um dos resultados do MCR com a restrição de trilinearidade são esses espectros. Potter *et al*¹⁰⁸ utilizaram soluções alcoólicas para a obtenção desses espectros. Os *loadings* do MCR – ALS foram comparados com os apresentados nesse trabalho e com isso foi possível a identificação das espécies químicas presentes.

Foram observadas diferenças entre os resultados obtidos com os da literatura, às quais foram relacionadas com os diferentes solventes utilizados nos

experimentos da literatura e nesse trabalho. A fluorescência pode ser influenciada pela polaridade do solvente, sendo esse efeito conhecido como solvatocromismo¹⁸. Entretanto, esse efeito não impediu a identificação das espécies presentes no conjunto de dados.

As proporções químicas, resultante da decomposição da matriz **C** do MCR – ALS, como explicado na seção 6.3.3.4, apresentam um perfil de equilíbrio químico das espécies identificadas (Figura 59). Os escores possibilitaram a estimativa dos valores de pK_a para o estado fundamental das diferentes estruturas do HBT. Usando um ajuste sigmoide para as regiões -2,00 a 2,00 (linha azul da Figura 59), 6,00 a 11,5 (linhas vermelha ou verde da Figura 60) e -2,00 a 2,00 (linha verde da Figura 59) foram determinados os seus valores de pK_a 0,08 (desprotonação do nitrogênio), 8,61 (desprotonação do HBT neutro) e 1,11 (pK_a^* , HBT neutro no estado excitado), respectivamente. A Equação de Förster (Equação 18) foi utilizada para calcular o valor do pK_a^* do HBT e compará-lo com o obtido com o ajuste sigmoide. A Tabela 12 resume todos os valores de pK_a e pK_a^* calculados e apresenta os comprimentos de onda de absorção e emissão das espécies de HBT identificadas.

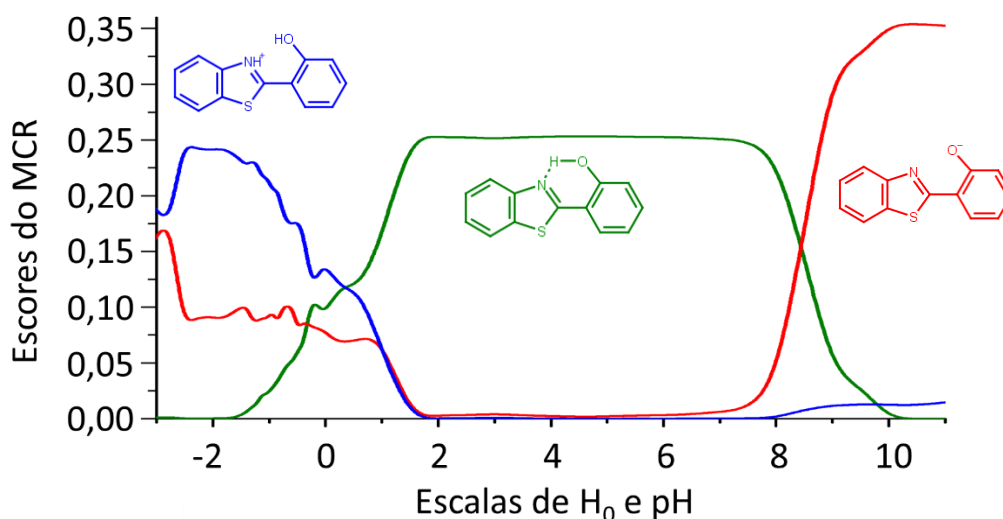


Figura 59. Escores do MCR – ALS em função de H_0 e pH. As linhas em azul, verde e vermelho representam as estruturas do HBT protonado, HBT neutro e HBT desprotonado, respectivamente.

Tabela 12. Valores dos comprimentos de onda de excitação e emissão e pK_a calculados pelo ajuste sigmoide dos escores do MCR – ALS (pK_a e pK_a^*) e estimado pela Equação de Förster.

Espécie	λ_{abs} (nm)	λ_{em} (nm)	pK_a	$pK_{aFörster}^*$	pK_a^*
Monocátion	363	460	0,08	-	-
Neutro	358	415	8,61	5,96	1,11
Ânion	375	460	-	-	-
Ceto	-	513	-	-	-

Os resultados apresentados na Tabela 13 demonstram que os valores de pK_a^* e $pK_{aFörster}^*$ são discrepantes. Como demonstrado na seção 8.1, a Equação de Förster possui algumas aproximações que são responsáveis por essas diferenças.

Normalmente, a determinação do pK_a^* é realizada com base em titulações fluorimétricas. Entretanto, essas titulações não resultam nos espectros de excitação e emissão das espécies presentes nas soluções. Apenas para fator comparativo, foi escolhido um comprimento de onda de excitação (350 nm) fixo para todas as superfícies e foram analisados os seus espectros de emissão, a partir da decomposição pelo MCR-ALS da matriz de dados formado pelos espectros de emissão a cada pH. Esse estudo foi realizado afim de determinar a possibilidade de encontrar os valores de pK_a , pK_a^* e $pK_{aFörster}^*$ sem utilização de dados trilineares.

As Figuras 60 e 61 apresentam respectivamente os resultados das matrizes **C** e **S** do MCR – ALS para essa matriz de dados.

Os resultados apresentados pelas Figuras 59 e 60 são similares, sendo que foi observada uma melhora nos valores dos escores do HBT desprotonado (linha vermelho) na região de -3,78 até 1,00. Os resultados para os valores de pK_a do monocátion, pK_a do neutro e pK_a^* do neutro foram -0,09 (antes era de 0,08), 8,67 (antes era de 8,61) e 0,87 (antes era de 1,11), respectivamente. Levando-se em conta a variação da escala de H_0 é em torno de 0,20 perto de zero e que o pH varia em torno de 0,5, pode-se dizer que esses resultados são equivalentes. Esses resultados indicam que para a determinação dessas constantes podem ser utilizados dados trilineares ou não trilineares (Tabela 13). Entretanto, os dados não trilineares não fornecem uma resposta para os espectros de excitação, e assim somente foram obtidos os espectros de emissão, como apresentado na Figura 61.

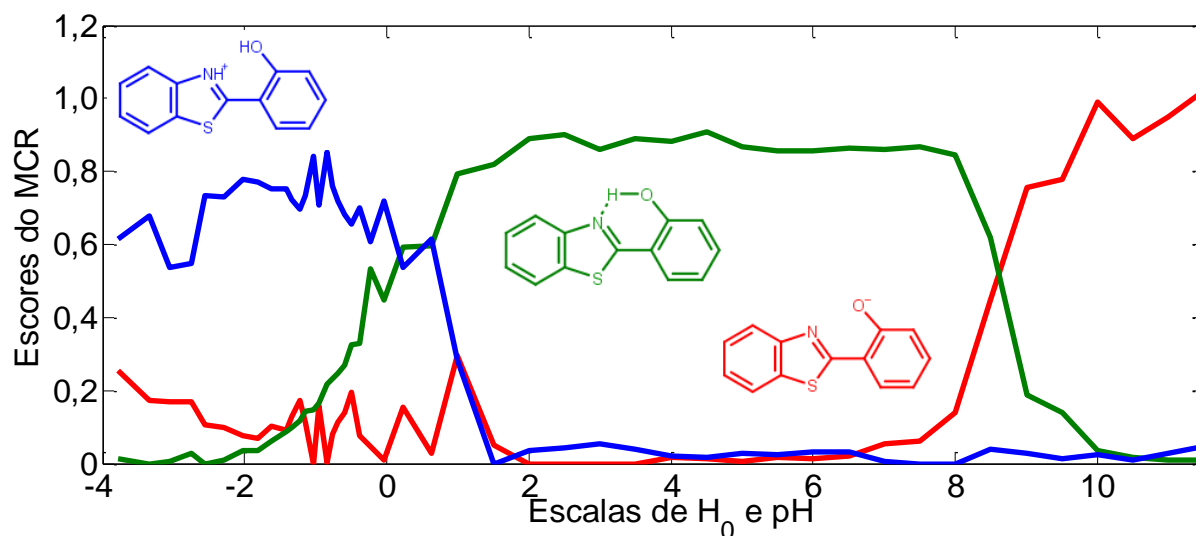


Figura 60. Escores do MCR – ALS em função de H_0 e pH para o conjunto de dados não trilineares. As linhas em azul, verde e vermelho representam as estruturas do HBT protonado, HBT neutro e HBT desprotonado, respectivamente (Figura 58).

Tabela 13. Comparação entre os resultados obtidos com os dados trilineares, bilineares e encontrados na literatura¹¹⁰.

Espécie	Trilinear		Bilinear		Literatura	
	pK_a	pK_a^*	pK_a	pK_a^*	pK_a	pK_a^*
Monocátion	0,08	-	-0,09	-	1,16*	-
Neutro	8,61	1,11	8,67	0,87	8,21	1,30

*Valor de pK_a obtido em uma solução de etanol/propanol¹¹⁰.

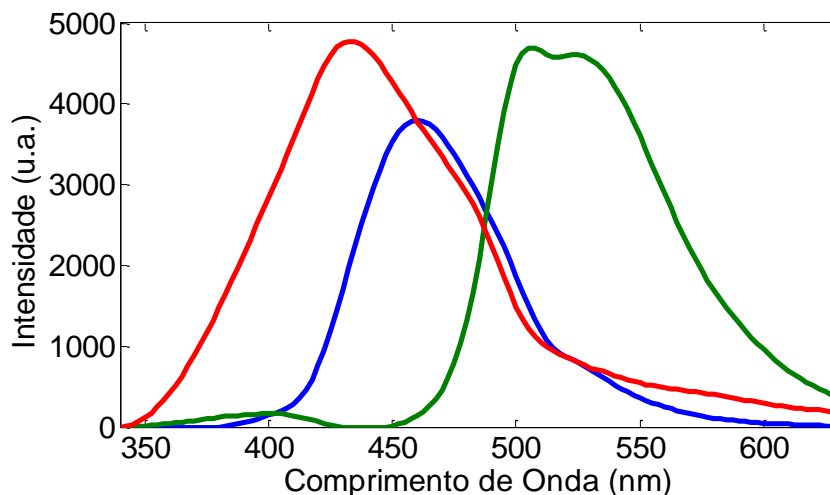


Figura 61. Espectros de emissão recuperados pelo MCR – ALS para o conjunto de dados não trilineares. As linhas em azul, verde e vermelho representam as estruturas do HBT protonado, HBT neutro e HBT desprotonado, respectivamente (Figura 58).

7.5. Conclusões Parciais

A combinação das superfícies de EEM com o MCR – ALS forneceu resultados promissores para a previsão do perfil químico das espécies presentes nos estados fundamental (S_0) e excitado (S_1). Essas espécies, por sua vez, podem estar relacionadas com processos de desativação devido a reações de transferência de prótons nos estados S_0 e S_1 .

O ajuste sigmoide proposto para os escores do MCR – ALS proporcionaram uma estimativa para os valores de pK_a para as formas neutra e protonada do HBT e pK_a^* para o ânion do HBT. O pK_a^* foi comparado com o $pK_{a\text{Förster}}^*$, sendo a diferença entre esses valores relacionada com as aproximações realizadas pela equação de Förster.

Os espectros recuperados pelo MCR – ALS foram comparados com os obtidos na literatura e essa comparação permitiu a identificação das estruturas de obtidas com a variação da concentração de prótons na solução.

**CAPÍTULO 8 –
CONCLUSÕES
GERAIS**

A análise de amostras forenses representa um grande desafio, pois as conclusões levam a determinar a direção que um processo poderá seguir. Em vista disso, torna-se de grande valia o uso de técnicas analíticas capazes de identificar corretamente a composição das amostras que serão analisadas.

Além disso, amostras que possuem propriedades fotoquímicas e fotofísicas ainda não estão sendo amplamente estudadas juntamente com os métodos quimiométricos. Como apresentado no Capítulo 8 essa combinação pode proporcionar em algumas vantagens, como a obtenção de espectros de excitação e emissão desses fluoróforos. E, também, é possível obter uma estimativa para os valores de pK_a e pK_a^* com base nos escores do MCR – ALS.

Muitos métodos baseados em medidas espectroscópicas possuem algumas vantagens em relação à outras metodologias analíticas, como, por exemplo, mantêm a integridade física das amostras e não necessitam de etapas laboriosas de preparo de amostra. Essa vantagem é a mais importante do ponto de vista forense uma vez que a mesma amostra pode ser analisada por diversas técnicas analíticas.

Os métodos de resolução de curvas possibilitam a comparação entre os espectros dos componentes puros das amostras e os espectros recuperados pelos métodos escolhidos. Esses métodos também possibilitam a identificação de componentes desconhecidos. Essa vantagem pode levar a correta identificação dos componentes presentes nas amostras. Além disso, um segundo resultado desses métodos está diretamente relacionado com a determinação da proporção química de cada um desses compostos nas amostras analisadas.

A combinação das técnicas espectroscópicas com os métodos de resolução de curvas proporcionou resultados importantes do ponto de vista forense. Foram analisadas amostras adulteradas em laboratório (óleo de gergelim e combustível) e amostras apreendidas pela polícia (cocaína e crack) e em ambos os casos foi possível a identificação dos compostos presentes (com exceção das amostras de óleo de gergelim adulteradas).

Referências Bibliográficas

- (1) Houck, M. M., *Forensic Science : Modern Methods Of Solving Crime*, First ed.; Praeger, 2007.
- (2) Bell, S., *Annu. Rev. Anal. Chem.* **2009**, 2 (1), 297–319.
- (3) Bueno, J.; Sikirzhytski, V.; Lednev, I. K., *Anal. Chem.* **2013**, 85 (15), 7287–7294.
- (4) McLaughlin, G.; Lednev, I. K., *Anal. Bioanal. Chem.* **2011**, 401 (8), 2511–2518.
- (5) Sylvia, J. M.; Janni, J. a.; Klein, J. D.; Spencer, K. M., *Anal. Chem.* **2000**, 72 (23), 5834–5840.
- (6) Yu, W. W.; White, I. M., *Analyst* **2013**, 138 (4), 1020–1025.
- (7) Sikirzhytski, V.; Sikirzhytskaya, A.; Lednev, I. K., *Forensic Sci. Int.* **2012**, 222 (1–3), 259–265.
- (8) McLaughlin, G.; Sikirzhytski, V.; Lednev, I. K., *Forensic Sci. Int.* **2013**, 231 (1–3), 157–166.
- (9) McLaughlin, G.; Lednev, I. K., *J. Forensic Sci.* **2015**, 60 (3), 595–604.
- (10) Muro, C. K.; Doty, K. C.; Bueno, J.; Halámková, L.; Lednev, I. K., *Anal. Chem.* **2015**, 87 (1), 306–327.
- (11) Mc Naught, A. D.; Wilkinson, A. Reference Material. In *IUPAC Compendium of Chemical Terminology*; IUPAC: Research Triangle Park, NC, 2014; p 1598.
- (12) Rohatgi-Mukherjee, K. K. *Fundamentals of Photochemistry*, 1st editio.; John Wiley & Sons, Ltd: New Delhi, 1978.
- (13) Sena, M. M.; Trevisan, M. G.; Poppi, R. J., *Quim. Nova* **2005**, 28 (5), 910–920.
- (14) de Juan, A.; Tauler, R., *Crit. Rev. Anal. Chem.* **2006**, 36 (3–4), 163–176.
- (15) Harshman, R. A.; Lundy, M. E., *Comput. Stat. Data Anal.* **1994**, 18 (1), 39–72.
- (16) Hyvärinen, A.; Oja, E., *Neural Networks* **2000**, 13 (4–5), 411–430.
- (17) de Oliveira Penido, C. A. F.; Pacheco, M. T. T.; Lednev, I. K.; Silveira, L., *J. Raman Spectrosc.* **2016**, 47 (1), 28–38.
- (18) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Lakowicz, J. R., Ed.; Springer US: Boston, MA, 2006.
- (19) Long, D. A. *The Raman Effect*; John Wiley & Sons, Ltd: Chichester, UK, 2002.
- (20) Lohumi, S.; Lee, S.; Lee, H.; Cho, B.-K., *Trends Food Sci. Technol.* **2015**, 46 (1), 85–98.
- (21) Qin, J.; Chao, K.; Kim, M. S., *Food Chem.* **2013**, 138 (2–3), 998–1007.
- (22) Oliveira, F. C. C.; Brandão, C. R. R.; Ramalho, H. F.; da Costa, L. A. F.; Suarez,

- P. A. Z.; Rubim, J. C., *Anal. Chim. Acta* **2007**, *587* (2), 194–199.
- (23) Vankeirsbilck, T.; Vercauteren, A.; Baeyens, W.; Van der Weken, G.; Verpoort, F.; Vergote, G.; Remon, J., *TrAC Trends Anal. Chem.* **2002**, *21* (12), 869–877.
- (24) Hargreaves, M. D.; Page, K.; Munshi, T.; Tomsett, R.; Lynch, G.; Edwards, H. G. M., *J. Raman Spectrosc.* **2008**, *39* (7), 873–880.
- (25) Moreno, V. M.; López-López, M.; Atoche, J.; García-Ruiz, C., *Sci. Justice* **2014**, *54* (2), 164–169.
- (26) Almeida, M. R.; Correa, D. N.; Zacca, J. J.; Logrado, L. P. L.; Poppi, R. J., *Anal. Chim. Acta* **2015**, *860* (2015), 15–22.
- (27) Firdous, S.; Anwar, S.; Waheed, A.; Maraj, M., *Laser Phys.* **2016**, *26* (4), 046001.
- (28) Miranda, A. M.; Castilho-Almeida, E. W.; Martins Ferreira, E. H.; Moreira, G. F.; Achete, C. A.; Armond, R. A. S. Z.; Dos Santos, H. F.; Jorio, A., *Fuel* **2014**, *115* (16), 118–125.
- (29) *New Trends in Fluorescence Spectroscopy*; Valeur, B., Brochon, J.-C., Eds.; Springer Series on Fluorescence; Springer Berlin Heidelberg: Berlin, Heidelberg, 2001; Vol. 1.
- (30) Jablonski, A., *Nature* **1933**, *131* (3319), 839–840.
- (31) Valeur, B.; Berberan-Santos, M. N. *Molecular Fluorescence*, 2nd ed.; Wiley-VCH, Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2012.
- (32) Skoog, D. A.; West, D. M.; Holler, F. J.; Crouch, S. R. *Fundamentals of Analytical Chemistry*, 9th ed.; Cengage Learning, 2014.
- (33) Escandar, G. M.; Goicoechea, H. C.; Muñoz de la Peña, A.; Olivieri, A. C., *Anal. Chim. Acta* **2014**, *806*, 8–26.
- (34) Kumar, K.; Mishra, A. K., *Talanta* **2013**, *117*, 209–220.
- (35) Goetz, A. F. H.; Vane, G.; Solomon, J. E.; Rock, B. N., *Science* (80). **1985**, *228* (4704), 1147–1153.
- (36) Firkala, T.; Farkas, A.; Vajna, B.; Nagy, Z. K.; Pokol, G.; Marosi, G.; Szilágyi, I. M., *J. Pharm. Biomed. Anal.* **2015**, *107* (2015), 318–324.
- (37) Mosca, S.; Frizzi, T.; Pontone, M.; Alberti, R.; Bombelli, L.; Capogrosso, V.; Nevin, A.; Valentini, G.; Comelli, D., *Microchem. J.* **2016**, *124* (2016), 775–784.
- (38) Stewart, S.; Priore, R. J.; Nelson, M. P.; Treado, P. J., *Annu. Rev. Anal. Chem.* **2012**, *5* (1), 337–360.
- (39) Edelman, G.; van Leeuwen, T. G.; Aalders, M. C. G., *Forensic Sci. Int.* **2012**, *223*

- (1–3), 72–77.
- (40) Wang, W.; Li, C.; Tollner, E. W.; Rains, G. C.; Gitaitis, R. D., *Comput. Electron. Agric.* **2012**, *80* (2012), 126–134.
- (41) *Visualizing Chemistry: The Progress and Promise of Advanced Chemical Imaging*, 1st edit.; National Academies Press: Washington (DC), 2006.
- (42) Sabin, G. P.; Souza, A. M. de; Breikreitz, M. C.; Poppi, R. J., *Quim. Nova* **2012**, *35* (3), 612–615.
- (43) Savitzky, A.; Golay, M. J. E., *Anal. Chem.* **1964**, *36* (8), 1627–1639.
- (44) Susi, H.; Michael Byler, D., *Biochem. Biophys. Res. Commun.* **1983**, *115* (1), 391–397.
- (45) Brereton, R. G., *Analyst* **2000**, *125* (11), 2125–2154.
- (46) Trefethen, L. N.; Bau III, D. *Numerical Linear Algebra*, 1st ed.; SIAM: Society for Industrial and Applied Mathematics, 1997.
- (47) de Juan, A.; Jaumot, J.; Tauler, R., *Anal. Methods* **2014**, *6* (14), 4964–4976.
- (48) Jaumot, J.; de Juan, A.; Tauler, R., *Chemom. Intell. Lab. Syst.* **2015**, *140*, 1–12.
- (49) Março, P. H.; Valderrama, P.; Alexandrino, G. L.; Poppi, R. J.; Tauler, R., *Quim. Nova* **2014**, *37* (9), 1525–1532.
- (50) de Juan, A.; Jaumot, J.; Tauler, R., *Anal. Methods* **2014**, *6* (14), 4964–4976.
- (51) Comon, P., *Signal Processing* **1994**, *36* (3), 287–314.
- (52) Hyvärinen, A.; Oja, E., *Neural Networks* **2000**, *13* (4–5), 411–430.
- (53) Rutledge, D. N.; Jouan-Rimbaud Bouveresse, D., *TrAC - Trends Anal. Chem.* **2013**, *50*, 22–32.
- (54) Jouan-Rimbaud Bouveresse, D.; Moya-González, A.; Ammari, F.; Rutledge, D. N., *Chemom. Intell. Lab. Syst.* **2012**, *112*, 24–32.
- (55) Durbin, J.; Watson, G. S., *Biometrika* **1950**, *37* (3/4), 409.
- (56) Bro, R., *Chemom. Intell. Lab. Syst.* **1997**, *38* (2), 149–171.
- (57) Andersson, C. A.; Bro, R., *Chemom. Intell. Lab. Syst.* **2000**, *52* (1), 1–4.
- (58) Murphy, K. R.; Stedmon, C. A.; Graeber, D.; Bro, R., *Anal. Methods* **2013**, *5* (23).
- (59) Pousa, G. P. A. G.; Santos, A. L. F.; Suarez, P. A. Z., *Energy Policy* **2007**, *35* (11), 5393–5398.
- (60) Robota, H. J.; Alger, J. C.; Shafer, L., *Energy & Fuels* **2013**, *27* (2), 985–996.
- (61) Alves, J. C. L.; Poppi, R. J., *Fuel* **2016**, *165*, 379–388.
- (62) Song, C. *Chemistry of Diesel Fuels*, 1st ed.; Mochida, I., Ed.; CRC Press, 2000.

- (63) Alves, J. C. L.; Henriques, C. B.; Poppi, R. J., *Fuel* **2012**, *97*, 710–717.
- (64) Mazivila, S. J.; Gontijo, L. C.; Santana, F. B. De; Mitsutake, H.; Santos, D. Q.; Borges Neto, W., *Energy & Fuels* **2015**, *29* (1), 227–232.
- (65) Alves, J. C. L.; Poppi, R. J., *Talanta* **2013**, *104*, 155–161.
- (66) Meira, M.; Quintella, C. M.; Ribeiro, E. M. O.; Silva, H. R. G.; Guimarães, A. K.; Luna, S.; Silva, W. L.; Brito, I. J. de, *J. Braz. Chem. Soc.* **2015**, *26* (7), 1351.
- (67) Soares, I. P.; Russo, R. M. de O.; Prates, R. G. D.; Augusti, R.; Fortes, I. C. P.; Pasa, V. M. D., *Quim. Nova* **2011**, *34* (8), 1439–1442.
- (68) de Vasconcelos, F. V. C.; de Souza, P. F. B.; Pimentel, M. F.; Pontes, M. J. C.; Pereira, C. F., *Anal. Chim. Acta* **2012**, *716*, 101–107.
- (69) Gaydou, V.; Kister, J.; Dupuy, N., *Chemom. Intell. Lab. Syst.* **2011**, *106* (2), 190–197.
- (70) Valderrama, P.; Braga, J. W. B.; Poppi, R. J., *Quim. Nova* **2009**, *32* (5), 1278–1287.
- (71) Goicoechea, H. C.; Olivieri, A. C.; Tauler, R., *Analyst* **2010**, *135* (3), 636–642.
- (72) Olivieri, A. C., *J. Chemom.* **2005**, *19* (4), 253–265.
- (73) *ASTM Standard D6866. Standard Test Methods for Determining the Biobased Content of Solid, Liquid, and Gaseous Samples Using Radiocarbon Analysis*, [Www.Astm.Org](http://www.astm.org).
- (74) *Norma Brasileira ABNT NBR 15568. Biodiesel – Determinação Do Teor de Biodiesel Em Óleo Diesel Por Espectroscopia Na Região Do Infravermelho Médio*, Associação Brasileira de Normas Técnicas, Rio de Janeiro, RJ; 2008, [Www.Abnt.Org.Br](http://www.abnt.org.br).
- (75) *ASTM Standard D7371. Standard Test Method for Determination of Biodiesel (Fatty Acid Methyl Esters) Content in Diesel Fuel Oil Using Mid Infrared Spectroscopy (FTIR- ATR-PLS Method)*. West Conshohocken (PA): ASTM International; 2014, [Www.Astm.Org](http://www.astm.org).
- (76) Shimanouchi, T. *NSRDS-NBS 39. Tables of Molecular Vibrational Frequencies Consolidated Volume I*; National Standard Reference Data System, 1972; p 161.
- (77) Tauler, R.; Smilde, A.; Kowalski, B., *J. Chemom.* **1995**, *9* (1), 31–58.
- (78) Paiva, E. M.; Rohwedder, J. J. R.; Pasquini, C.; Pimentel, M. F.; Pereira, C. F., *Fuel* **2015**, *160*, 57–63.
- (79) Santos, V. O.; Oliveira, F. C. C.; Lima, D. G.; Petry, A. C.; Garcia, E.; Suarez, P.

- A. Z.; Rubim, J. C. A., *Anal. Chim. Acta* **2005**, *547* (2), 188–196.
- (80) West, M. J.; Went, M. J., *Drug Test. Anal.* **2011**, *3* (9), 532–538.
- (81) Pérez-Alfonso, C.; Galipienso, N.; Garrigues, S.; de la Guardia, M. A., *Forensic Sci. Int.* **2014**, *237*, 70–77.
- (82) Penido, C. A. F. O.; Pacheco, M. T. T.; Zângaro, R. A.; Silveira, L., *J. Forensic Sci.* **2015**, *60* (1), 171–178.
- (83) Tallarida, C. S.; Egan, E.; Alejo, G. D.; Raffa, R.; Tallarida, R. J.; Rawls, S. M., *Neuropharmacology* **2014**, *79*, 590–595.
- (84) Lanaro, R.; Costa, J. L.; Cazenave, S. O. S.; Alves, E.; Barros, M. C.; Polesel, D. N., In *Brazilian Journal of Toxicology*; Sociedade Brasileira de Toxicologia: Ribeirão Preto, 2011; Vol. 24, p 388.
- (85) Bertucci, C.; Tedesco, D.; Fabini, E.; Di Pietra, A. M.; Rossi, F.; Garagnani, M.; Del Borrello, E.; Andrisano, V., *J. Chromatogr. A* **2014**, *1363*, 150–154.
- (86) Tong, L.; Ding, L.; Li, Y.; Wang, Z.; Wang, J.; Liu, Y.; Yang, L.; Wen, A., *J. Chromatogr. B* **2011**, *879* (5–6), 299–303.
- (87) Penido, C. A. F. de O.; Silveira, L.; Pacheco, M. T. T., *Instrum. Sci. Technol.* **2012**, *40* (5), 441–456.
- (88) Botelho, É. D.; Cunha, R. B.; Campos, A. F. C.; Maldaner, A. O., *J. Braz. Chem. Soc.* **2014**, *25* (4), 611–618.
- (89) Lapachinske, S. F.; Okai, G. G.; dos Santos, A.; de Bairros, A. V.; Yonamine, M., *Forensic Sci. Int.* **2015**, *247*, 48–53.
- (90) Fukushima, A. R.; Carvalho, V. M.; Carvalho, D. G.; Diaz, E.; Bustillos, J. O. W. V.; Spinosa, H. de S.; Chasin, A. A. M., *Forensic Sci. Int.* **2014**, *243*, 95–98.
- (91) Burgina, E. B.; Baltakhinov, V. P.; Boldyreva, E. V; Shakhtschneider, T. P., *J. Struct. Chem.* **2004**, *45* (1), 64–73.
- (92) Bio-Rad Laboratories, Inc. SpectraBase; SpectraBase Compound ID=YcCeSaut19 SpectraBase Spectrum ID=CGxLjPdgaVS <http://spectrabase.com/spectrum/CGxLjPdgaVS> (acessado em 18 de Janeiro de 2019).
- (93) Xu, J.; Liu, X.-F.; Wang, Y.-T., *Food Chem.* **2016**, *212*, 72–77.
- (94) Zhang, L.; Shuai, Q.; Li, P.; Zhang, Q.; Ma, F.; Zhang, W.; Ding, X., *Food Chem.* **2016**, *192*, 60–66.
- (95) Zhang, L.; Huang, X.; Li, P.; Na, W.; Jiang, J.; Mao, J.; Ding, X.; Zhang, Q.,

- Chemom. Intell. Lab. Syst.* **2017**, *161*, 147–150.
- (96) Ulberth, F.; Buchgraber, M., *Eur. J. Lipid Sci. Technol.* **2000**, *102* (11), 687–694.
- (97) Sikorska, E.; Romaniuk, A.; Khmelinskii, I. V.; Herance, R.; Bourdelande, J. L.; Sikorski, M.; Koziol, J., *J. Fluoresc.* **2004**, *14* (1), 25–35.
- (98) Bahram, M.; Bro, R.; Stedmon, C.; Afkhami, A., *J. Chemom.* **2006**, *20* (3–4), 99–105.
- (99) Harshman, R. a; De Sarbo, W. S., In *Research methods for multimode data analysis*; Law, H. G., Snyder, C. W., McDonald, R. P., Hattie, J., Eds.; Praeger Publishers Inc, 1984; pp 602–642.
- (100) Neves, A. C. de O.; Tauler, R.; de Lima, K. M. G., *Anal. Chim. Acta* **2016**, *937*, 21–28.
- (101) Kirby, A. J., *Acc. Chem. Res.* **1997**, *30* (7), 290–296.
- (102) Oliveira, T. C. F.; Carmo, L. F. V.; Murta, B.; Duarte, L. G. T. A.; Nome, R. A.; Rocha, W. R.; Brandão, T. A. S., *Phys. Chem. Chem. Phys.* **2015**, *17* (4), 2404–2415.
- (103) Zhao, J.; Ji, S.; Chen, Y.; Guo, H.; Yang, P., *Phys. Chem. Chem. Phys.* **2012**, *14* (25), 8803–8817.
- (104) Williams, D. L.; Heller, A., *J. Phys. Chem.* **1970**, *74* (26), 4473–4480.
- (105) Hyman, H. H.; Kilpatrick, M.; Katz, J. J., *J. Am. Chem. Soc.* **1957**, *79* (14), 3668–3671.
- (106) Hammett, L. P.; Deyrup, A. J., *J. Am. Chem. Soc.* **1932**, *54* (7), 2721–2739.
- (107) Hammett, L. P.; Deyrup, A. J., *J. Am. Chem. Soc.* **1932**, *54* (11), 4239–4247.
- (108) Potter, C. A. S.; Brown, R. G., *Chem. Phys. Lett.* **1988**, *153* (1), 7–12.
- (109) Elsaesser, T.; Schmetzer, B., *Chem. Phys. Lett.* **1987**, *140* (3), 293–299.
- (110) Potter, C. A. S.; Brown, R. G.; Vollmer, F.; Rettig, W., *J. Chem. Soc., Faraday Trans.* **1994**, *90* (1), 59–67.

ANEXOS

Anexo 1 – Autorização para análise das amostras de cocaína e crack



TESTE DE UMA NOVA TÉCNICA DE ANÁLISE DE DROGAS DE RUA



Este documento mostra o cronograma das análises experimentais realizadas no Laboratório de Quimiometria em Química Analítica do Instituto de Química da UNICAMP em amostras de droga de rua recebidas de apreensões pelo laboratório de toxicologia do NPC Campinas, e sob a responsabilidade da Perita Criminal Larissa Dantas. Os experimentos foram feitos na forma de prospecção de nova técnica de eventual interesse científico para identificação de substâncias proscritas ilegalmente comercializadas no município de Campinas e região. Foram utilizados os aparelhos espectrômetro dispersivo Raman Station 400F e espectrômetro Spectrum 100 N FT-NIR, acoplado a um sistema de imagem Spotlight 400N, ambos da Perking-Elmer. Tais equipamentos realizam coleta de dados da forma de imagens químicas. Ressalvamos que as tecnologias são do tipo não destrutiva, e que, portanto não prejudicarão no presente ou no futuro quaisquer materiais por nós analisados.

N	DATA DA ANÁLISE	NPC AMOSTRA	Técnica utilizada	
			Raman	NIR
1	01/11/2016	9954/16	X	X
2	02/11/2016	9936/16	X	X
3	03/11/2016	9892/16	X	X
4	06/02/2017	14949/16	-	X
5	07/02/2017	14863/16	X	X
6	08/02/2017	14425/16	X	X
7	09/02/2017	14423/16	X	X
8	10/02/2017	13527/16	X	X
9	13/02/2017	13960/16	X	X
10	14/02/2017	13018/16	X	X
11	15/02/2017	10200/16	X	X
12	16/02/2017	312/17	-	X
13	17/02/2017	317/17	-	X
14	20/02/2017	332/17	X	X
15	21/02/2017	359/17	X	X
16	22/02/2017	360/17	-	X
17	23/02/2017	406/17	-	X
18	24/02/2017	430/17	X	X
19	06/03/2017	453/17	X	X
20	07/03/2017	4789-80/17	X	X
21	08/03/2017	4789-80/17	X	X
22	09/03/2017	489/17	X	X
23	10/03/2017	494/17	X	X
24	13/03/2017	502/17	X	X
25	14/03/2017	504/17	X	X
26	15/03/2017	507/17	X	X
27	16/03/2017	540/17	X	X
28	17/03/2017	609-611/17	X	X

Envolvidos

William F. Cadore Dantas

William Dantas
Doutorando Instituto de Química UNICAMP

Larissa Dantas
Larissa Dantas
Perita Criminal

Ronei Jesus Poppi

Dr Ronei Jesus Poppi
Professor Titular do Instituto de Química UNICAMP

Edvaldo Messias Barrós
Edvaldo Messias Barrós
Perito Criminal Diretor do NPC Campinas

Anexo 2 – Identificação e substituição de Outliers

```
function data_novo = identificacao_de_outliers(data,nx,ny)

% Na etapa da PCA sempre plotar o T^2 contra o Q para identificar quais
% seriam os espectros a serem substituídos.

figure
plot(data')
model = pca(data,3);

n = input('Número dos Pixels a serem substituídos');

if n == zeros(1);
    data_novo = data;
    return
end

[~,g] = size(n);

y_temp = data;

for i=1:g
    y = media_vizinhos(y_temp,n(i),nx,ny);
    y_temp = y;
    figure
    plot(data(n(i),:),'.-')
    hold on
    plot(y(n(i),:),'r')
end

data_novo = y_temp;

end
```

```
function [novo_espectro] = media_vizinhos(data,n,nx,ny)

% Sempre deve ser indicado os valores da linha correspondente ao
% espectro que será alterada pela rotina (i), bem como os valores
% de nx e ny para que seja calculado corretamente a média dos
% espectros vizinhos.

[~,nz] = size(data);
data_novo = nan(1,nz);

a = (1:1:nx*ny);
a = reshape(a,nx,ny);
a = a';
```

```

b = a(1,1);           % Primeiro pixel da imagem
c = a(1,nx);         % Pixel nx da primeira linha da imagem
d = a(ny,1);        % Primeiro pixel da última linha da imagem
e = a(ny,nx);       % Último pixel da imagem, pixel nx*ny
f = a(2:ny-1,1);    % Pixels da primeira linha da imagem, com
                    % exceção dos pixels 1 e nx
g = a(2:ny-1,nx);   % Pixels da primeira coluna da imagem, com
                    % exceção dos pixels 1 e primeiro pixel da última
                    % linha
h = a(1,2:nx-1);    % Pixels da última coluna da imagem, com exceção
                    % dos pixels nx e nx*ny
i = a(ny,2:nx-1);   % Pixels da última linha da imagem, com exceção
                    % dos pixels nx*ny-nx+1 e nx*ny
j = a(2:ny-1,2:nx-1); % Todos os outros pixels da imagem, com exceção
                    % dos pixels anteriores

if n == b
    data_novo = primeiro_pixel(data,n,nx);
end

if n == c
    data_novo = pixel_nx(data,n,nx);
end

if n == d
    data_novo = pixel_nxny_nx(data,n,nx);
end

if n == e
    data_novo = pixel_nxny(data,n,nx);
end

if ismember(n,f)
    data_novo = primeira_coluna(data,n,nx);
end

if ismember(n,g)
    data_novo = ultima_coluna(data,n,nx);
end

if ismember(n,h)
    data_novo = primeira_linha(data,n,nx);
end

if ismember(n,i)
    data_novo = ultima_linha(data,n,nx);
end

if ismember(n,j)
    data_novo = pixels_meio(data,n,nx);
end

novo_espectro = data;
novo_espectro(n,:) = data_novo;

```

```

%% Equações
% Primeiro pixel da imagem
function [data_novo] = primeiro_pixel(data,n,nx)
    data_novo = (data(n+nx,:)+data(n+nx+1,:)+data(n+1,:))./3;
end

% Pixel nx da primeira linha da imagem
function [data_novo] = pixel_nx(data,n,nx)
    data_novo = (data(n+nx,:)+data(n+nx-1,:)+data(n-1,:))./3;
end

% Primeiro pixel da última linha da imagem
function [data_novo] = pixel_nxny_nx(data,n,nx)
    data_novo = (data(n-nx,:)+data(n-nx+1,:)+data(n+1,:))./3;
end

% Último pixel da imagem, pixel nx*ny
function [data_novo] = pixel_nxny(data,n,nx)
    data_novo = (data(n-nx,:)+data(n-nx-1,:)+data(n-1,:))./3;
end

% Pixels da primeira coluna da imagem, com exceção dos pixels 1 e nx
function [data_novo] = primeira_coluna(data,n,nx)
    data_novo = (data(n-nx+1,:)+data(n-
nx,:)+data(n+1,:)+data(n+nx,:)+data(n+nx+1,:))./5;
end

% Pixels da última coluna da imagem, com exceção dos pixels 1 e primeiro
pixel da última linha
function [data_novo] = ultima_coluna(data,n,nx)
    data_novo = (data(n-nx-1,:)+data(n-nx,:)+data(n-1,:)+data(n+nx-
1,:)+data(n+nx,:))./5;
end

% Pixels da primeira linha da imagem, com exceção dos pixels nx e nx*ny
function [data_novo] = primeira_linha(data,n,nx)
    data_novo = (data(n+nx-1,:)+data(n+nx,:)+data(n+nx+1,:)+data(n-
1,:)+data(n+1,:))./5;
end

% Pixels da última linha da imagem, com exceção dos pixels nx*ny-nx+1 e
nx*ny
function [data_novo] = ultima_linha(data,n,nx)

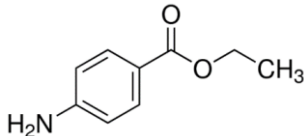
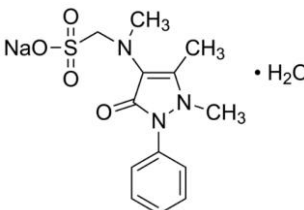
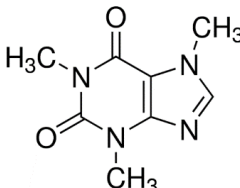
    data_novo = (data(n-nx-1,:)+data(n-nx,:)+data(n-nx+1,:)+data(n-
1,:)+data(n+1,:))./5;

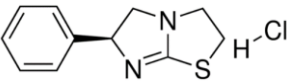
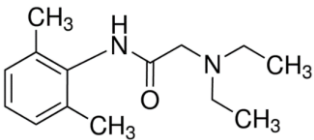
end

```

```
% Todos os outros pixels da imagem, com exceção dos pixels anteriores
function [data_novo] = pixels_meio(data,n,nx)
    data_novo = (data(n-nx-1,:)+data(n-nx,:)+data(n-
nx+1,:)+data(n+nx-1,:)+...
                data(n+nx,:)+data(n+nx+1,:)+data(n-1,:)+data(n+1,:))./8;
    end
end
```

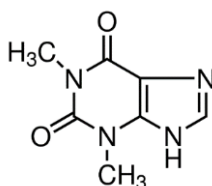

Anexo 3 – Modos de vibração dos aditivos da cocaína

Composto	Número de onda (cm ⁻¹)	Atribuição ^{17,82,91}
Benzocaína 	864	Estiramento simétrico C-O
	1172	Deformação no plano H-C-H
	1282	Estiramento simétrico das ligações C-C, C-N e C-O e estiramento do anel aromático
	1604	Estiramento e deformação do anel aromático e tesoura NH ₂
	1682	Estiramento simétrico C=O
Dipirona 	602	Deformação N-C-O
	790	Deformação C-N-C
	1062	Estiramento simétrico N=C=O
	1596	Deformação N-H
	1630	Estiramento C-C do anel aromático
1666	Estiramento simétrico C=O	
Cafeína 	644	Deformação C=C-N
	742	Deformação O=C-C
	802	Deformação N-C-H
	930	Estiramento simétrico N-CH ₃
	1070	Estiramento simétrico C-C
	1240	Deformação H-C=N
	1286	Estiramento C-N
	1330	Estiramento C-N
	1362	Estiramento C-N
	1600	Estiramento C=C
1660	Estiramento assimétrico C=O	
1698	Estiramento simétrico C=O	

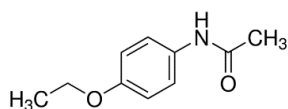
	618	Estiramento C-S e deformação C-N-H
Levamisol	636	Estiramento C-S, deformação assimétrico do anel fenílico
	996	Deformação simétrico do anel fenílico
	1038	Torção do grupo CH ₂
	1466	Deformação CH ₂
	1606	Deformação C-H
	780	Deformação fora do plano C-H do anel aromático
	954	Estiramento C-C
	994	Estiramento C-C
	1046	Estiramento C-C do anel aromático e estiramento C-N
	1094	Estiramento C-N
Lidocaína	1274	Estiramentos C-N, C-C e do anel aromático
	1388	Estiramento C-N
	1440	Estiramento C-N
	1452	Estiramento C-N e deformação N-H
	1472	Estiramento C-N
	1548	Estiramento C=C do anel aromático
	1594	Estiramento e deformação do anel aromático
	1656	Estiramento C=O da amida primária

	614	Deformação C=C-N
	668	Deformação O=C-N
	764	Deformação O=C-C
	930	Deformação N-C-H
	950	Estiramento C-N
	970	Deformação N=C-H
	980	Estiramento assimétrico N-CH ₃
	1088	Estiramento simétrico C-N
	1212	Estiramento assimétrico C-N
	1248	Estiramento C-N
	1286	Estiramento C-N
	1316	Estiramento C-N
	1570	Estiramento simétrico C=O
	1610	Estiramento C=C
	1663	Estiramento assimétrico C=O
	1708	Estiramento C=N
	634	Deformações do fenil e amida
	648	Deformação do anel aromático
	858	Deformação fora do plano C-H
	969	Estiramento C-O
	1178	Deformação no plano C-H
	1236	Deformação no plano C-H
	1270	Estiramento C-CH ₃
	1307	Deformações CH ₂ e CH ₃
	1328	Estiramentos C-N e fenil-N-C, deformação C-N-H
	1440	Estiramento C-C
	1456	Estiramento C-C
	1540	Estiramento C=C
	1600	Estiramento do fenil
	1660	Deformação N-H
	1674	Estiramento C=O

Teofilina



Fenacetina



Anexo 4 – Soluções de Ácido Clorídrico e Soluções Tampão

- Soluções de ácido clorídrico: Escala H_0

A escala de acidez de Hammett (H_0) foi desenvolvida por Louis P. Hammett e Alden J. Deyrup em 1932^{A4.1,A4.2}. Essa escala foi baseada no equilíbrio de protonação da anilina, a Equação A4.1 apresenta de forma esquemática o equilíbrio de protonação genérico, onde B representa a anilina. Com base nesse equilíbrio eles obtiveram a Equação A4.2 que é utilizada para determinar os valores de H_0 ^{A4.3}.



$$H_0 = pK_{BH^+} - \log \frac{C_{BH^+}}{C_B} \quad \text{Equação A4.2}$$

K_{BH^+} é a constante de ionização do ácido conjugado, e $\frac{C_{BH^+}}{C_B}$ é a razão entre as concentrações das espécies envolvidas. A escala H_0 permite que haja valores negativos para a acidez de um meio, sendo útil para soluções concentradas de ácido forte. Com base em dados da literatura, obteve-se soluções de HCl concentrado na faixa de 0,10 a 8,00 mol/L (pH 1 até H_0 -3,78)^{A4.4}. A Tabela A4.1 apresenta os volumes de cada uma das soluções utilizadas para serem obtidos esses valores de pH e H_0 . Em todos os casos a concentração inicial de HCl foi de 8,00 mol/L, a concentração de HBT foi 8,05 $\mu\text{mol/L}$ e para ajustar a força iônica foi utilizada uma solução de KCl de 0,10 mol/L. Todas as soluções foram preparadas em cubetas de quartzo, sendo que o volume final de 2,00 mL.

- Soluções Tampão: Escala pH

A Tabela A4.2 apresenta os volumes de cada uma das soluções adicionadas para a obtenção dos valores de pH esperados^{A4.5}. Foram utilizadas soluções de citrato de sódio (0,10 mol/L), HCl (0,10 mol/L), NaOH (0,10 mol/L), KH_2PO_4 (0,10 mol/L), H_3BO_3 (0,10 mol/L), NaHCO_3 (0,05 mol/L), Na_2HPO_4 (0,05 mol/L), KCl (0,20 mol/L) e NaOH (0,20 mol/L). Todas os tampões foram preparados em cubetas de quartzo, sendo que o volume final de 2,00 mL.

Tabela A4.1. Volumes de HCl e KCl adicionados em cada uma das soluções para serem obtidos os valores de H₀ pretendidos.

H ₀	Volume de HCl (mL)	Volume de KCl (mL)	H ₀	Volume de HCl (mL)	Volume de KCl (mL)
-3,78	2,0000	0,0000	-1,04	0,7500	1,2500
-3,30	1,8750	0,1250	-0,95	0,6875	1,3125
-3,05	1,8000	0,2000	-0,85	0,6250	1,3750
-2,75	1,7000	0,3000	-0,77	0,5625	1,4375
-2,55	1,6250	0,3750	-0,68	0,5000	1,5000
-2,26	1,5000	0,5000	-0,59	0,4375	1,5625
-2,02	1,3750	0,6250	-0,49	0,3750	1,6250
-1,81	1,2500	0,7500	-0,38	0,3125	1,6875
-1,62	1,1250	0,8750	-0,23	0,2500	1,7500
-1,42	1,0000	1,0000	-0,04	0,1875	1,8125
-1,33	0,9375	1,0625	0,23	0,1250	1,8750
-1,23	0,8750	1,1250	0,62	0,0625	1,9375
-1,14	0,8125	1,1875	0,92	0,0250	1,9750

Tabela A4.2. Soluções tampão para as amostras que estiveram na faixa pH 1,1 até 11,6.

Faixa de pH	pH	Volume de Citrato (mL)	Volume de HCl (mL)
	1,5	0,4440	1,5560
	2,0	0,6120	1,3880
	2,5	0,7080	1,2920
	3,0	0,8060	1,1940
	3,5	0,9360	1,0640
	4,0	1,1200	0,880
	4,5	1,4380	0,5620
Faixa de pH	pH	Volume de NaOH (mL)	Volume de Citrato (mL)
	5,0	0,0720	1,9280
5,0-6,6	5,5	0,5540	1,4460
	6,0	0,8080	1,1920
	6,5	0,9260	1,0740
Faixa de pH	pH	Volume de NaOH (mL)	Volume de KH₂PO₄ (mL)
	7,0	0,5820	1,0000
5,8-8	7,6	0,8480	1,0000
	8,0	0,9220	1,0000
Faixa de pH	pH	Volume de NaOH (mL)	Volume de H₃BO₃ e KCl (mL)
	8,6	0,2360	1,0000
8-10,2	9,2	0,5280	1,0000
	9,5	0,6920	1,000
	9,8	0,8120	1,0000
Faixa de pH	pH	Volume de NaOH (mL)	Volume de NaHCO₃(mL)
10,6-11	10,6	0,3820	1,0000
	11,0	0,4540	1,0000
Faixa de pH	pH	Volume de NaOH (mL)	Volume de Na₂HPO₄ (mL)
11-11,9	11,6	0,0257	1,0000

- (A4.1) Hammett, L. P., Deyrup, A. J., *J. Am. Chem. Soc.*, **1932**, 54, 7, 2721-2739
- (A4.2) Hammett, L. P., Deyrup, A. J., *J. Am. Chem. Soc.*, **1932**, 54, 11, 4239-4247
- (A4.3) Oliveira, T. C. F., Carmo, L. F. V., Murta, B., Duarte, L. G. T. A., Nome, A. R., Rocha, W. R., Brandão, T. A. S., *Phys. Chem. Chem. Phys.*, **2015**, 17, 2404-2415
- (A4.4) Paul, M. A., Long, F. A., *Chem. Rev.*, **1957**, 57, 1–45.
- (A4.5) Perrin, D. D., Dempsey, B., *Buffer for pH and Metal Ion Control*, 1st ed., John Wiley & Sons, **1974**.