



Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza - CCEN
Departamento de Química Fundamental - DQF

LÍVIA RODRIGUES E BRITO

Determinação de aditivos detergentes dispersantes em gasolina utilizando a técnica do *ring-oven* e imagens hiperespectrais na região do infravermelho próximo

Recife
2014

LÍVIA RODRIGUES E BRITO

Determinação de aditivos detergentes dispersantes em gasolina utilizando a técnica do *ring-oven* e imagens hiperspectrais na região do infravermelho próximo

Dissertação de Mestrado submetida ao programa de Pós-Graduação em Química da Universidade Federal de Pernambuco, como parte dos requisitos para a obtenção do título de Mestre em Química.

Área de concentração: Química Analítica.

Orientadora: Maria Fernanda Pimentel
Co-orientador: Jarbas José Rodrigues Rohwedder

Recife
2014

Catálogo na fonte
Bibliotecário Jefferson Luiz Alves Nazareno CRB4-1758

B862d Brito, Livia Rodrigues e.
Determinação de aditivos detergentes dispersantes em gasolina utilizando a técnica do ring-oven e imagens hiperespectrais na região do infravermelho próximo. / Livia Rodrigues e Brito. – Recife: O Autor, 2014.
153 f.: fig., tab.

Orientadora: Maria Fernanda Pimentel.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN. Química Fundamental, 2014.
Inclui referências e apêndices.

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

543 CDD (22. ed.) UFPE-FQ 2015-36

Dissertação submetida ao Corpo Docente do Programa de Pós-Graduação em Química do Departamento de Química Fundamental da Universidade Federal de Pernambuco como parte dos requisitos necessários para a obtenção do Grau de Mestre em Química.

Aprovada:

Profa. Maria Fernanda Pimentel Avelar
(Orientadora)
Departamento de Engenharia Química
Universidade Federal de Pernambuco

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

Prof. Edenir Rodrigues Pereira Filho
Departamento de Química
Universidade Federal de São Carlos

Determinação de aditivos detergentes dispersantes em gasolina utilizando a técnica do Ring-Oven e imagens hiperespectrais na região do infravermelho próximo.

Por

Lívia Rodrigues e Brito

Departamento de Química Fundamental
Centro de Ciências Exatas e da Natureza
Universidade Federal de Pernambuco
Recife-PE-Brasil
25 de agosto de 2014

Por toda a ajuda direta e indireta e pela enorme paciência,
dedico este trabalho à minha família, amigos e orientadores.

AGRADECIMENTOS

Agradeço primeiramente à minha orientadora, Prof^a Dr.^a Maria Fernanda Pimentel, pela excelente orientação, dedicação e paciência. Por me entender quando nem eu parecia saber o que queria fazer, por confiar que eu conseguiria e, principalmente, por me dar conselhos e me incentivar a continuar na pesquisa.

Ao meu co-orientador, Prof. Dr. Jarbas José Rodrigues Rohwedder, por todo apoio em todas as idas à Unicamp, pelos e-mails trocados e pelas excelentes ideias. Foram poucos momentos de convivência direta, mas eles foram suficientes para eu me tornar uma fã incondicional do seu caráter e humildade.

A Prof.^a Dr.^a Fernanda Araújo Honorato, não só por me ajudar a dar meus primeiros passos na pesquisa, mas por ser sempre tão presente e atenciosa.

Ao Prof. Dr. Célio Pasquini pelas discussões sempre produtivas e pelo apoio em diferentes etapas deste trabalho.

Aos meus pais, Jorge e Márcia, e minha irmã, Júlia, por me ajudarem indiretamente, especialmente à minha mãe por me ajudar mesmo quando isso significava me ouvir explicar alguma coisa que ela não entendia para tentar me ajudar de alguma forma.

Aos meus avós, Cleto e Albanita, pela compreensão de tantas ausências e pelo apoio e amor incondicional.

Ao meu namorado e amigo, Luiz Henrique, por estar presente em todas as horas, mesmo quando eu só precisava de companhia, e por me ajudar até lendo o que eu escrevia (sem entender) simplesmente para me ajudar.

Aos meus amigos do colégio, presentes há mais de 7 anos na minha vida, com um agradecimento especial a Luís Felipe por ter me ajudado diretamente neste trabalho, sem ele eu teria perdido muitas noites de sono aprendendo programação e talvez nem tivesse conseguido fazer tudo sozinha.

Aos grandes amigos que fiz na Unicamp, especialmente a Benedito, cujas conversas sempre me alegravam quando eu estava longe de casa, e à minha amiga Aline que me abrigou e, em pouco tempo, se tornou uma amiga muito especial.

Aos meus amigos do laboratório de combustíveis, com um agradecimento especial aos meus amigos Marcones e Vivi, com quem compartilhei tantas alegrias, aflições e que tornaram todo o trabalho mais divertido e à Michelle, minha companheira de experimentos e de tantas idas à Unicamp.

Aos meus amigos do DQF pelas conversa, caronas e dias/noites de estudo, especialmente a Ronaldo e Thaisa que me ajudaram quando eu mais precisei e se tornaram amigos mais do que especiais.

À toda a equipe do LAC, sempre disposta a ajudar, fazendo o possível e o impossível para que eu pudesse ir sempre para os congressos.

Ao CNPq pela minha bolsa de mestrado, ao INCTAA e NUQAAPE pelos recursos concedidos e ao laboratório de combustíveis (LAC) e grupo de instrumentação analítica (GIA) por permitirem que eu realizasse este trabalho.

RESUMO

A adição de aditivos detergentes dispersantes nas gasolinas brasileiras será obrigatória a partir de julho de 2015. É necessário, portanto, desenvolver uma metodologia que permita quantificar esses aditivos para verificar o cumprimento da lei. Neste trabalho, é proposto um método que associa a técnica do *ring-oven* com as imagens hiperespectrais no infravermelho próximo (NIR-HI). Como os aditivos são adicionados em baixas concentrações, a técnica do *ring-oven* foi empregada para concentrá-los previamente à análise por NIR-HI. Anéis foram produzidos a partir de amostras de gasolinas comum adicionadas dos aditivos (denominados G, T, W e Y) fornecidos pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) e as imagens adquiridas utilizando uma câmera hiperespectral (SisuCHEMA). Três estratégias de extração dos espectros do anel foram testadas a fim de se escolher a mais rápida e objetiva. A estratégia escolhida se baseia nos histogramas dos escores da primeira componente principal das imagens analisadas individualmente. Modelos de calibração individuais para cada aditivo foram construídos empregando a regressão por mínimos quadrados parciais (PLS), por isso, fez-se necessária uma etapa prévia de classificação. O melhor resultado para classificação foi obtido empregando a análise discriminante linear (LDA) associada ao algoritmo genético (GA) para seleção de variáveis, o qual apresentou uma taxa de classificações corretas de 92,31 %. Observou-se que a maioria dos erros de classificação envolveram amostras dos aditivos G e W. Um único modelo de regressão para esses dois aditivos foi, então, construído e seu erro foi equivalente aos dos modelos individuais. Os modelos de regressão apresentaram erros médios de predição entre 2 e 15 %. Esses resultados mostram que a metodologia proposta pode ser utilizada para determinar as concentrações dos aditivos com confiabilidade e garantir que eles estão sendo adicionados conforme a lei.

Palavras chave: Gasolina. Aditivos. *Ring-oven*. Imagem hiperespectral. Infravermelho próximo.

ABSTRACT

The addition of detergent dispersant additives to Brazilian gasoline will be mandatory from July 2015. It is necessary, therefore, to develop a methodology that allows quantifying these additives to verify their compliance with the law. In this work, a method that associates the *ring-oven* technique with near infrared hyperspectral images (NIR-HI) is proposed. Because the additives are added in low concentrations, the *ring-oven* technique was employed to concentrate them prior to the NIR-HI analysis. Rings were produced from samples of gasolines without additives spiked with additives (called G, T, W and Y) provided by the National Agency of Petroleum, Natural Gas and Biofuels (ANP) and the images were acquired using a hyperspectral camera (SisuCHEMA). Three strategies for extraction of the ring spectra were tested in order to select the faster and most objective. The chosen strategy is based on the histograms of the first principal component scores of the images analyzed individually. Regression models were built for each additive using partial least squares (PLS) regression, so it was necessary to have a previous classification stage. The best classification result was obtained using the linear discriminant analysis (LDA) associated with the genetic algorithm (GA) for variable selection, which showed a correct classification rate of 92.31 %. It was observed that most of the misclassification errors involved the samples of the G and W additives. A single regression model was then built for these two additives and its error was equivalent to the errors of the individual models. The regression models showed average prediction errors between 2 and 15 %. These results show that the proposed methodology can be used to determine the additive concentrations with reliability and to ensure that they are been added according to the law.

Key-words: Gasoline. Additives. Ring-oven. Hyperspectral image. Near infrared.

LISTA DE FIGURAS

- Figura 1** – Comparação dos depósitos formados na válvula de admissão (a) antes e (b) depois da utilização de aditivos detergentes dispersantes derivados do poliisobuteno. 26
- Figura 2** – Estruturas de alguns compostos utilizados como aditivos detergentes dispersantes para gasolinas: (a) e (b) derivados de succinimida, (c) sal de amônio quaternário e (d) poliéter amina. 27
- Figura 3** – (a) Aparato original do ring-oven e (b) equipamento comercial disponível em 1970. 33
- Figura 4** – Anel formado a partir da reação de cloreto de ferro (III) e ferrocianeto de potássio. 34
- Figura 5** – Tipos de vibrações moleculares. 37
- Figura 6** – Representação esquemática do modelo do oscilador (a) harmônico e (b) anarmônico para a variação de energia potencial (U) de moléculas diatômicas. 37
- Figura 7** – Tipos de imagens: (a) imagem em escala de cinza, (b) imagem RGB e (c) imagem multiespectral ou hiperespectral. 39
- Figura 8** – Desdobramento da matriz de dados segundo a função reshape. 42
- Figura 9** – Gráfico de dispersão da PC 3 versus a PC 2. 46
- Figura 10** – (a) Imagem RGB original, (b) gráfico de dispersão da PC 2 versus PC 3 com 6 regiões selecionadas por polígonos de cores diferentes e (c) imagem com os pixels das regiões selecionadas marcados com as mesmas cores. 47
- Figura 11** – Sistema utilizado na destilação dos aditivos. 55
- Figura 12** – Sistema utilizado para o preparo dos anéis. 58
- Figura 13** – Forno utilizado para o preparo dos anéis desmontado. 58
- Figura 14** – Câmera hiperespectral SisuCHEMA SWIR utilizada na aquisição das imagens. 59
- Figura 15** – (a) Produção do anel e (b) papel de filtro com anel produzido a partir de um blend de gasolinas contendo 100 ppm do aditivo G. 64
- Figura 16** – Espectros de uma amostra contendo 50 ppm do aditivo G: (a) espectros completos e (b) espectros após a remoção das regiões ruidosas. 64
- Figura 17** – Imagens dos escores da PCA de um anel de uma amostra contendo 50 ppm do aditivo G (a) sem pré-processamento, pré-processada com (b) SNV, (c) MSC e (d) 1° derivada com filtro de Savitzky-Golay (polinômio de 2° grau e janela de 7 pontos). 66

Figura 18 – Espectros de um anel de uma amostra contendo 50 ppm do aditivo G corrigidos com SNV. 66

Figura 19 – Imagens dos escores da PCA de um anel contendo 50 ppm do aditivo G produzido a partir da condição (a) 1 (600 μ L de amostra), (b) 2 (1200 μ L de amostra), (c) 3 (1800 μ L de amostra) e (d) 4 (2400 μ L de amostra). 68

Figura 20 – Imagens dos escores da PCA de anéis contendo (a) 50 ppm, (b) 500 ppm e (c) 1.000 ppm do aditivo G. 69

Figura 21 – Gráficos dos pesos da PCA de um anel contendo (a) 50 ppm e (c) 1.000 ppm do aditivo G. 70

Figura 22 – Mapas de concentração de um anel contendo 50 ppm do aditivo G para a MCR-ALS (a) individual e (b) aumentada (todas as imagens de diferentes concentrações dos aditivos juntas). 71

Figura 23 – Espectros otimizados pela MCR-ALS dos aditivos, do blend de gasolinas comuns e do papel com as regiões características dos aditivos destacadas em vermelho. 73

Figura 24 – Espectros (a) médios e (b) somados e subtraídos dos espectros do papel extraídos através dos mapas de concentração da MCR-ALS dos anéis produzidos a partir das diferentes concentrações do aditivo G na gasolina comum. 74

Figura 25 – Imagens dos escores da PCA de um anel de uma amostra contendo 50 ppm do aditivo G utilizando (a) os espectros completos e (b) apenas as regiões características dos aditivos. 77

Figura 26 – Imagens dos escores da primeira componente da PCA de um anel de uma amostra contendo 50 ppm do aditivo G utilizando apenas as regiões características dos aditivos e agrupando-se as imagens pelas (a) triplicatas, (b) concentrações de 50 a 300 ppm e (b) concentrações de 50 a 1.000 ppm. 77

Figura 27 – Imagens dos escores da PCA de um anel de uma amostra contendo (a) 50 ppm e (b) 1.000 ppm do aditivo G utilizando apenas as regiões características dos aditivos e agrupando-se as imagens de todas as concentrações (50 a 1.000 ppm). 78

Figura 28 – Histogramas dos escores da primeira componente principal da PCA de anéis de amostras contendo (a) 100 ppm, (b) 500 ppm e (c) 1.000 ppm do aditivo G agrupando todas as concentrações. 81

Figura 29 – Imagens de uma amostra de um anel contendo 50 ppm do aditivo G com pixels selecionados utilizando (a) um desvio padrão, (b) dois desvios-padrões, (c) três desvios padrões e (d) quatro desvios padrões no cálculo dos limites (os pixels em preto são os que estão acima do limite). 81

Figura 30 – Imagens de um anel contendo 100 ppm do aditivo G com pixels selecionados utilizando três desvios padrões com a PCA construída (a) agrupando-se as imagens de todas as concentrações e (b) individualmente para cada imagem. 84

Figura 31 – Imagens dos escores da primeira componente principal da PCA das regiões características dos aditivos de anéis preparados a partir de gasolinas comuns de diferentes distribuidores. Anel (a) não identificado, (b) pouco definido e (c) bem definido. 86

Figura 32 – Espectros de todos os anéis do aditivo G somados e subtraídos dos espectros do papel extraídos através dos histogramas da primeira componente da PCA. 87

Figura 33 – Espectros de todos os anéis do aditivo G somados e subtraídos dos espectros do papel extraídos através dos histogramas da primeira componente da PCA. Subtração dos espectros do papel (a) dos pixels das bordas da imagem e (b) dos pixels com valores de escores abaixo do máximo do histograma. 89

Figura 34 – Gráficos dos valores preditos das concentrações em função das concentrações definidas no preparo das amostras para os aditivos (a) G, (b) T, (c) W e (d) Y. 91

Figura 35 – Gráficos dos coeficientes de regressão para os modelos dos aditivos (a) G, (b) T, (c) W e (d) Y. 91

Figura 36 – Espectro médio pré-processado com suavização com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa com as variáveis selecionadas pelo SPA. 94

Figura 37 – Funções discriminantes do modelo SPA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa. 95

Figura 38 – Espectros médios pré-processado com (a) 2º derivada com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa e (b) suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa com as variáveis selecionadas pelo GA apenas nas regiões características do espectro. 97

Figura 39 – Funções discriminantes do modelo GA-LDA dos espectros pré-processados com a 2º derivada com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa. 98

Figura 40 – Funções discriminantes do modelo GA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa. 99

Figura 41 – Funções discriminantes do modelo LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa utilizando todas as variáveis das regiões características dos aditivos. 100

Figura 42 – (a) Gráficos dos valores preditos das concentrações em função das concentrações definidas no preparo das amostras do modelo conjunto dos aditivos G e W e (b) Gráficos dos coeficientes de regressão do modelo. 103

LISTA DE TABELAS

Tabela 1 – Alguns tipos de aditivos e suas principais funções.	24
Tabela 2 – Concentrações das amostras preparadas utilizando o blend de gasolinas e as gasolinas de diferentes distribuidores para cada aditivo.	57
Tabela 3 – Quantidade de amostras nos conjuntos de treinamento, validação e predição.	61
Tabela 4 – Avaliação da estabilidade térmica dos aditivos.	63
Tabela 5 – Condições de preparo dos anéis analisadas por imagem.	67
Tabela 6 – Melhores resultados obtidos para os espectros extraídos pela MCR-ALS com base nos mapas de concentração dos aditivos.	75
Tabela 7 – Melhores resultados obtidos para os espectros extraídos pela PCA com base nas imagens dos escores da primeira componente.	79
Tabela 8 – Melhores resultados obtidos para os espectros extraídos pela PCA com base nos histogramas dos valores dos escores da primeira componente (análise das imagens de todas as concentrações agrupadas).	83
Tabela 9 – Melhores resultados obtidos para os espectros extraídos pela PCA com base nos histogramas dos valores dos escores da primeira componente (análise individual das imagens).	85
Tabela 10 – Resultados obtidos para os espectros somados e subtraídos dos espectros (1) dos pixels das bordas da imagem e (2) dos pixels com valores de escores abaixo do máximo do histograma.	90
Tabela 11 – Resultados obtidos para os espectros somados e subtraídos dos espectros dos pixels com valores de escores abaixo do máximo do histograma.	90
Tabela 12 – Resultados de classificação (SPA-LDA).	93
Tabela 13 – Resultados do modelo SPA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa, onde N é a quantidade de amostras de cada classe.	95
Tabela 14 – Resultados de classificação (GA-LDA).	96
Tabela 15 – Resultados do modelo GA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa, onde N é a quantidade de amostras de cada classe.	99
Tabela 16 – Resultados de classificação (LDA) obtidos para as variáveis das regiões características dos aditivos.	100

Tabela 17 – Resultados do modelo LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa utilizando todas as variáveis das regiões características dos aditivos. 101

Tabela 18 – Resultados de previsão obtidos utilizando os quatro modelos de regressão dos aditivos na previsão das amostras dos quatro aditivos. 101

Tabela 19 – Resultados obtidos para os espectros somados subtraídos dos espectros dos pixels com valores de escores abaixo do máximo do histograma. 102

Tabela 20 – Limites de detecção e quantificação dos melhores modelos de regressão. 103

LISTA DE ABREVIATURAS E SIGLAS

ALS	Mínimos quadrados alternados, do inglês <i>alternating least squares</i>
ANP	Agência Nacional de Petróleo, Gás Natural e Biocombustíveis
ASTM	<i>American society for testing and materials</i>
CCD	Cromatografia em camada delgada
GA	Algoritmo genético, do inglês <i>genetic algorithm</i>
GPC	Cromatografia por permeação em gel, do inglês <i>gel permeation chromatography</i>
IV	Infravermelho
JK	<i>Jack-knife</i>
KS	<i>Kennard-Stone</i>
LDA	Análise discriminante linear, do inglês <i>linear discriminant analysis</i>
LIBS	Espectroscopia de emissão em plasma induzido por laser, do inglês <i>laser induced breakdown spectroscopy</i>
MCR	Resolução multivariada de curvas, do inglês <i>multivariate curve resolution</i>
MSC	Correção do espalhamento multiplicativo, do inglês <i>multiplicative scatter correction</i>
NIR	Infravermelho Próximo, do inglês <i>near infrared spectroscopy</i>
PCA	Análise de componentes principais, do inglês <i>principal components analysis</i>
PLS	Mínimos quadrados parciais, do inglês <i>partial least squares</i>
RMSEC	Erro quadrático médio de calibração, do inglês <i>root mean square error of calibration</i>
RMSECV	Raiz quadrada do erro médio quadrático de validação cruzada, do inglês <i>root mean square error of cross-validation</i>
RMSEP	Raiz quadrada do erro médio quadrático de previsão, do inglês <i>root mean square error of prediction</i>
ROI	Região de interesse, do inglês <i>region of interest</i>
SEC	Cromatografia por exclusão de tamanho, do inglês <i>size exclusion chromatography</i>
SNV	Variação normal padrão, do inglês <i>standard normal variate</i>
SPA	Successive projections algorithm

SUMÁRIO

1. INTRODUÇÃO	18
2. OBJETIVOS	22
3. FUNDAMENTAÇÃO TEÓRICA	23
3.1. Gasolina	23
3.2. Aditivos	24
3.3. <i>Ring-oven</i>	32
3.4. Infravermelho próximo	35
3.5. Imagens hiperespectrais	38
3.5.1. Técnicas de pré-processamentos	42
3.5.2. Análise de componentes principais	45
3.5.3. Resolução multivariada de curvas	47
3.5.4. Regressão por mínimos quadrados parciais	48
3.5.5. Análise discriminante linear	52
4. MATERIAIS E MÉTODOS	55
4.1. Aditivos	55
4.2. Preparo das amostras	56
4.3. Preparo dos anéis	57
4.4. Aquisição das imagens	59
4.5. Processamento das imagens e construção dos modelos	60
5. RESULTADOS E DISCUSSÃO	62
5.1. Resultados preliminares	62
5.1.1. Extração dos espectros do anel	71
5.1.1.1. Mapas de concentração da resolução multivariada de curvas	72
5.1.1.2. Imagens dos escores das componentes principais	76
5.1.1.2.1. Determinação visual dos limites	78
5.1.1.2.2. Determinação dos limites com base nos histogramas	80
5.2. Resultados consolidados	86
5.2.1. Modelos de classificação	92
6. CONCLUSÃO	105
REFERÊNCIAS BIBLIOGRÁFICAS	107
APÊNDICES	111
APÊNDICE A - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos manualmente dos mapas de concentração da MCR-ALS aumentada.	111
APÊNDICE B - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos manualmente das imagens dos escores da PCA aumentada.	112
APÊNDICE C - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos através dos histogramas dos escores da PCA aumentada.	113

APÊNDICE D - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos através dos histogramas dos escores da PCA individual de cada imagem. 114

APÊNDICE E - Modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 115

APÊNDICE F - Resultados de previsão utilizando os modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 116

APÊNDICE G - Modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 119

APÊNDICE H - Resultados de previsão utilizando os modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 120

APÊNDICE I - Modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 123

APÊNDICE J - Resultados de previsão utilizando os modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 124

APÊNDICE K - Modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 127

APÊNDICE L - Resultados de previsão utilizando os modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 128

APÊNDICE M - Modelos de regressão do aditivo G construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 131

APÊNDICE N - Resultados de previsão utilizando os modelos de regressão do aditivo G construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 132

APÊNDICE O - Modelos de regressão do aditivo Y construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 135

APÊNDICE P - Resultados de previsão utilizando os modelos de regressão do aditivo Y construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 136

APÊNDICE Q - Modelos de classificação SPA-LDA construídos com variáveis selecionadas no espectro completo. 139

APÊNDICE R - Modelos de classificação SPA-LDA construídos com variáveis selecionadas apenas nas regiões características dos aditivos. 140

APÊNDICE S - Modelos de classificação GA-LDA construídos com variáveis selecionadas no espectro completo. 141

APÊNDICE T - Modelos de classificação GA-LDA construídos com variáveis selecionadas apenas nas regiões características dos aditivos. 143

APÊNDICE U - Modelos de classificação LDA construídos com variáveis as 53 variáveis das regiões características dos aditivos. 145

APÊNDICE V - Resultados de previsão utilizando os melhores modelos de regressão construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 146

APÊNDICE W - Modelos de regressão dos aditivos G e W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 148

APÊNDICE X - Resultados de previsão utilizando os modelos de regressão dos aditivos G e W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem). 149

1. INTRODUÇÃO

Os gases de exaustão dos automóveis, emitidos durante a queima do combustível, são uma das principais fontes de poluição atmosférica. A emissão desses gases é um problema crescente, principalmente nas grandes cidades, devido ao número cada vez maior de automóveis trafegando nas ruas (KARPOV, 2007). A composição e toxicidade dos gases depende, dentre outros fatores, das características do motor e da composição do combustível utilizado.

O aumento da conscientização ambiental tem levado ao surgimento de regulamentações ambientais cada vez mais rígidas quanto à emissão de poluentes na atmosfera. O resultado é a necessidade de desenvolvimento de tecnologias mais limpas que satisfaçam essas regulamentações. No que diz respeito aos automóveis e combustíveis, os produtores passaram a se preocupar mais com a composição dos gases de exaustão, buscando soluções técnicas para adequar os motores e combustíveis às exigências ambientais (KARPOV, 2007).

Com relação à composição dos combustíveis, mais especificamente das gasolinas, uma das formas de reduzir a toxicidade dos gases de exaustão é através da adição de aditivos detergentes dispersantes. Esses aditivos têm como principal função reduzir a formação de depósitos nos motores e válvulas, o que aumenta sua eficiência e, conseqüentemente, resulta em uma queima mais limpa. Isso porque a composição dos gases de exaustão é alterada e eles se tornam menos tóxicos devido à redução da concentração de monóxido de carbono (CO) e dos hidrocarbonetos em 50 a 60% e dos óxidos de nitrogênio (NO_x) em cerca de 20% (NIKITINA et al., 2006).

Reconhecendo os benefícios da adição desses aditivos nas gasolinas ou devido às legislações ambientais cada vez mais rígidas, alguns países como os Estados Unidos (EUA), Rússia e alguns países da Europa Ocidental já utilizam a gasolina contendo aditivos há alguns anos. Em alguns países, a adição de aditivos se tornou obrigatória, como é o caso dos EUA, nos quais desde 1996 as gasolinas comercializadas contêm aditivos detergentes. No caso dos países da Europa, apesar do uso de aditivos não ser obrigatório, em 2006, cerca de 80% das gasolinas já continham detergentes (NIKITINA et al., 2006).

No Brasil, a gasolina aditivada (contendo aditivos detergentes dispersantes) já é comercializada, contudo apresenta um preço superior ao da gasolina comum (sem aditivo). Esse fato, aliado à falta de informação sobre os seus benefícios e por não

estarem disponíveis em todos os postos revendedores, faz com que o seu consumo no país ainda seja baixo. Contudo, visando à redução da toxicidade dos gases de exaustão, a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), órgão responsável por implementar a política nacional dos combustíveis e especificar a sua qualidade, instituiu que a partir de 1º de julho de 2015 toda gasolina comercializada no Brasil deverá conter aditivos detergentes dispersantes na sua composição (BRASIL, 2014). A concentração mínima desses aditivos irá variar dependendo do tipo do aditivo e da sua eficiência, mas a Resolução ANP N°1, de 6 de janeiro de 2014, define que a concentração máxima permitida será de 5.000 ppm (0,5% (m/m)).

Apesar de haver um grande número de patentes relacionadas à síntese de novos aditivos detergentes dispersantes registradas nas últimas décadas, não existem muitos métodos para sua identificação e determinação em gasolina na literatura. Esses métodos são essenciais para as agências reguladoras verificarem se a adição está sendo realizada conforme estabelecido. Uma vez que alguns países, como os Estados Unidos, instituíram o uso de gasolinas contendo aditivos detergentes, estudos são realizados no sentido de desenvolver métodos para identificá-los e quantificá-los nos combustíveis desde os anos 90. Um exemplo é a patente canadense CA 2132806 A1, que propõe o uso da cromatografia por permeação em gel utilizando um detector por índice de refração (BRAUER; MIIN, 1995). Esse método permite a determinação tanto qualitativa quanto quantitativa dos aditivos realizando-se uma etapa prévia de calibração com soluções padrão de composição conhecidas.

Métodos baseados em técnicas espectroscópicas, como a espectroscopia na região do infravermelho, também vêm sendo desenvolvidos e avaliados. Uma vez que a concentração dos aditivos na gasolina é baixa (no máximo 0,5%), a análise por infravermelho (próximo e médio), proposta por De Medeiros (2009), necessita de uma etapa prévia de concentração dos aditivos. Uma das formas utilizadas por Medeiros (2009) para pré-concentrar os aditivos, também utilizada por Dos Santos (2012) e Da Silva (2014), foi destilar a gasolina e analisar o resíduo da destilação, no qual estão os aditivos e outros componentes de elevada massa molar da gasolina. Embora essa estratégia de pré-concentração tenha se mostrado eficiente para a análise qualitativa (DA SILVA et al., 2014), ela não permite quantificar os aditivos em concentrações inferiores a 150 ppm (DE MEDEIROS, 2009).

Uma técnica de análise qualitativa, na qual analitos que estavam muito diluídos eram concentrados em papéis de filtro na forma de anéis para serem posteriormente identificados com a utilização de indicadores ou reveladores, foi desenvolvida na década de 50 por Herbert Weisz. Nessa técnica, o papel de filtro é mantido sob aquecimento e a amostra é gotejada sobre ele. Assim, os componentes voláteis evaporam e os demais migram, por capilaridade, do centro do papel para suas bordas. Essa técnica foi desenvolvida para ser aplicada em análises qualitativas onde se dispunha de quantidades reduzidas de amostras, as quais não poderiam ser analisadas por métodos convencionais devido à baixa concentração dos analitos. Contudo, ao longo dos anos, ela foi utilizada em diversos ramos da química analítica, sendo aplicada não só na análise qualitativa mas também quantitativa, na determinação de íons inorgânicos, substâncias orgânicas, radioativas, etc (WEISZ, 1970).

Cerca de 60 anos depois do desenvolvimento da técnica do *ring-oven*, ela foi revisitada por Cortez e Pasquini (2012) e utilizada na pré-concentração de cátions de sódio, ferro e cobre em etanol combustível para posterior determinação utilizando espectroscopia de plasma induzido por laser (LIBS, do inglês *laser induced breakdown spectroscopy*). Ao associar a técnica do *ring-oven* com LIBS, esse trabalho mostrou a possibilidade de utilizá-la com outras técnicas modernas de análise, como a imagem hiperespectral na região do infravermelho próximo (NIR, do inglês *near infrared spectroscopy*), buscando solucionar problemas que envolvem baixas concentrações dos analitos, como é o caso da determinação dos aditivos nas gasolinas.

As imagens hiperespectrais são imagens nas quais cada *pixel* (menor elemento da imagem) é representado por um espectro completo. Esse tipo de imagem possui informação espacial e espectral, o que permite adquirir informações sobre a composição e distribuição de componentes em uma amostra. Com o desenvolvimento dos equipamentos de aquisição de imagens hiperespectrais, tornou-se possível obter imagens com espectros Raman, infravermelho (IV), ultravioleta dentre outros, dependendo do tipo de informação espectral que se deseja (PRATS-MONTALBÁN; DE JUAN; FERRER, 2011).

Por serem constituídas de três dimensões, duas espaciais e uma espectral, essas imagens contêm grande quantidade de informação. Elas podem ter milhares ou até milhões de pontos e, normalmente, apresentam elevada correlação entre os dados. Sua análise só foi possível com a adequação das técnicas quimiométricas

clássicas como, por exemplo, análise de componentes principais (PCA, do inglês *principal components analysis*) e resolução multivariada de curvas (MCR, do inglês *multivariate curve resolution*) (VIDAL; AMIGO, 2012).

Muitas áreas, como a farmacêutica, médica e forense, têm se beneficiado do uso deste tipo de imagens. Um exemplo são os trabalhos que têm sido desenvolvidos na área farmacêutica para avaliar a distribuição e homogeneidade dos componentes (princípios ativos e excipientes) (AMIGO et al., 2008; FRANCH-LAGE et al., 2011).

Na área de combustíveis, há vários trabalhos que utilizam a espectroscopia NIR associada a técnicas quimiométricas, tanto na determinação das propriedades dos combustíveis (HONORATO et al., 2008) quanto em análises qualitativas (SILVA et al., 2012). Contudo, ainda não há relatos na literatura de trabalhos envolvendo imagens hiperespectrais na região NIR, possivelmente, porque o tipo de estudo realizado até o momento não se beneficia do uso de informações espaciais.

No presente trabalho, é proposto um novo método para a determinação de aditivos detergentes dispersantes em gasolinas associando a técnica do *ring-oven* com a análise de imagens hiperespectrais no NIR.

2. OBJETIVOS

2.1. Objetivo geral

- Desenvolver um método para a determinação de aditivos detergentes dispersantes em gasolina utilizando a técnica do *ring-oven* e imagens hiperespectrais no infravermelho próximo.

2.2. Objetivos específicos

- Avaliar a técnica do *ring-oven* como estratégia de pré-concentração dos aditivos
- Avaliar o uso das imagens hiperespectrais na identificação e quantificação dos aditivos detergentes dispersantes em gasolinas;
- Comparar diferentes estratégias de seleção e extração dos espectros da região de interesse das imagens;
- Construir modelos de classificação para identificar o tipo do aditivo presente nas amostras;
- Construir modelos de regressão para quantificar os aditivos.

3. FUNDAMENTAÇÃO TEÓRICA

3.1. Gasolina

A gasolina é um combustível derivado do petróleo que consiste em uma mistura complexa de compostos, na qual os hidrocarbonetos são os principais constituintes. Além dos hidrocarbonetos, os quais podem ser saturados (parafinas e cicloparafinas) ou insaturados (olefinas e aromáticos), ela também pode conter compostos sulfurados, oxigenados, nitrogenados e metálicos em sua composição, a qual varia dependendo, por exemplo, do modo de produção e da origem do petróleo.

Segundo a Resolução ANP N° 40, de 25 de outubro de 2013, as gasolinas automotivas classificam-se em: gasolina A e gasolina C. A gasolina A é aquela produzida nas refinarias, petroquímicas ou nos formuladores e que é isenta de compostos oxigenados. Já a gasolina C, a qual é comercializada nos postos, é a gasolina A acrescida de etanol anidro combustível (BRASIL, 2013b). A concentração do etanol pode variar dependendo da sua disponibilidade e preço, estando atualmente fixada em 25% (v/v) (BRASIL, 2013a). Ao ser adicionado à gasolina, ele aumenta a sua resistência à detonação e reduz a emissão de monóxido de carbono na atmosfera.

Ainda segundo a Resolução ANP N° 40, as gasolinas A e C podem ser de dois tipos: Comum ou Premium. A diferença é que a gasolina Premium possui uma resistência à detonação maior, com índice antidetonante mínimo de 91, enquanto que o da gasolina comum é de 87 (BRASIL, 2013b).

Além do etanol anidro combustível, a partir de 1° de julho de 2015, as gasolinas C comercializadas no Brasil deverão conter, obrigatoriamente, aditivos detergentes dispersantes em concentrações que irão variar dependendo da sua eficiência (BRASIL, 2014). Os aditivos são produtos adicionados tanto para adequar o combustível às especificações técnicas (ex. aditivos antidetonantes) quanto para agregar características benéficas a ele (ex. aditivos detergentes dispersantes) (SILIN et al., 2012).

Após a adição dos aditivos pelo distribuidor, a gasolina C deixa de ser chamada comum para ser chamada aditivada. A gasolina aditivada se diferencia visualmente da comum pela sua coloração, pois enquanto a comum varia de incolor a amarelada, a aditivada recebe um corante que pode ser de qualquer cor, com exceção de azul (reservada para a gasolina de aviação) (BRASIL, 2013b).

3.2. Aditivos

Os aditivos, além do componente ativo, responsável pelas suas características benéficas, podem conter um diluente, adicionado para facilitar a mistura com o combustível e seu fluxo no motor, e/ou um fluido carregador, adicionado para viabilizar ou potencializar a redução da formação dos depósitos pelo componente ativo (BRASIL, 2014). É importante observar que essa mistura de componentes pode conter mais de um componente ativo.

Existem diferentes tipos de aditivos para combustíveis, os quais se diferenciam quanto às suas funções. Alguns desses aditivos e suas principais funções estão listados na Tabela 1.

Tabela 1 – Alguns tipos de aditivos e suas principais funções.

Aditivo	Principais funções
Detergente dispersante	Remover os depósitos e minimizar a sua formação
Antioxidante	Reduzir a oxidação dos compostos orgânicos
Antidetonação	Aumentar a resistência à detonação
Inibidor de corrosão	Minimizar a corrosão do sistema de combustível
Desativador de metais	Reduzir o efeito catalítico dos metais na oxidação
Anticongelante	Evitar o congelamento do combustível no motor

Fonte: adaptada de De Medeiros, 2009.

Durante o funcionamento normal de um automóvel, ocorre uma série de reações de combustão envolvendo os diferentes componentes presentes na gasolina. Além dos gases produzidos nessas reações, ocorre também a formação de alguns compostos que podem se depositar nas partes internas do motor. Gasolinas com quantidade elevada de hidrocarbonetos insaturados (olefinas e aromáticos), compostos oxigenados e aditivos antidetonantes tendem a apresentar uma quantidade maior de depósitos, principalmente, nas válvulas de admissão, as quais controlam a entrada e saída dos gases na câmara de explosão. Depósitos formados nessa região do motor promovem um aumento da combustão incompleta do combustível e, conseqüentemente, uma redução da potência do motor. Outras conseqüências diretas são a alteração na composição dos gases de exaustão

(aumento da sua toxicidade) e o aumento do consumo de combustível (NIKITINA et al., 2006).

Os aditivos detergentes dispersantes tanto removem os depósitos previamente formados (aditivos *clean up*) quanto minimizam a sua formação (aditivos *keep clean*). Os aditivos *clean up* removem os depósitos através da sua dispersão e posterior solubilização no combustível.

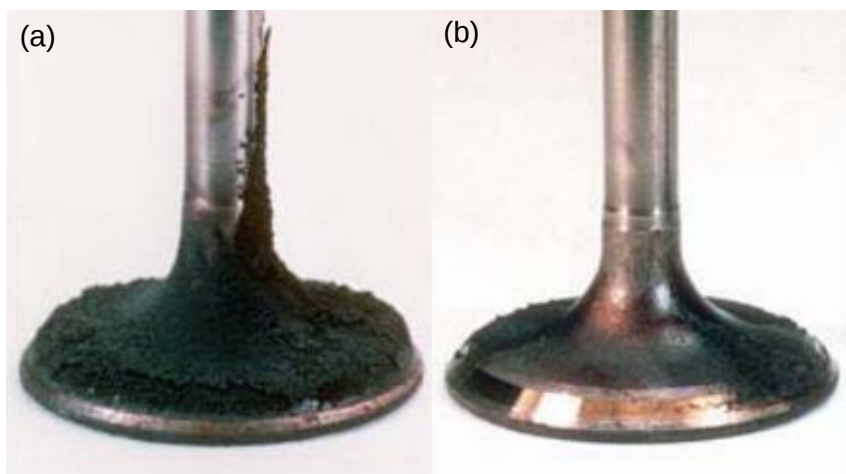
Os aditivos tanto podem ser adicionados diretamente na gasolina pelo distribuidor, para ser posteriormente comercializada (gasolina aditivada), quanto podem ser adicionados no tanque de combustível pelo próprio consumidor, o qual adquire o aditivo em frascos nos postos revendedores. No primeiro caso, o aditivo tem a função de manter o motor limpo (*keep clean*), enquanto que, no segundo, sua função é limpar o motor (*clean up*).

Os aditivos detergentes dispersantes removem os depósitos em cerca de 60 a 70% e minimizam sua formação posterior, resultando em uma queima mais eficiente e melhorando o desempenho do motor. Além disso, causam a redução de cerca de 3% no consumo de combustível, 50 a 60% da emissão de monóxido de carbono (CO) e hidrocarbonetos e 20 % da emissão de óxidos de nitrogênio (NO_x) (NIKITINA et al., 2006; KARPOV, 2007).

A eficiência do aditivo quanto à limpeza de depósitos já formados deve ser comprovada pelo fabricante ou agente importador. Um exemplo de demonstração de eficiência na remoção de depósitos na válvula de admissão, após o uso de aditivos detergentes dispersantes derivados do poliisobuteno, pode ser observado na Figura 1 (ORONITE, 2014).

Reconhecendo os benefícios da adição dos aditivos detergentes dispersantes ou de forma a atender às legislações ambientais cada vez mais rígidas, muitos países como os Estados Unidos, alguns países da Europa Ocidental e a Rússia já adotaram o seu uso (NIKITINA et al., 2006; KARPOV, 2007). Devido aos benefícios ambientais, o Brasil vai seguir o exemplo dos EUA e tornar obrigatória a adição dos aditivos, o que faz necessário o desenvolvimento de metodologias para a verificação do cumprimento da lei.

Figura 1 – Comparação dos depósitos formados na válvula de admissão (a) antes e (b) depois da utilização de aditivos detergentes dispersantes derivados do poliisobuteno.



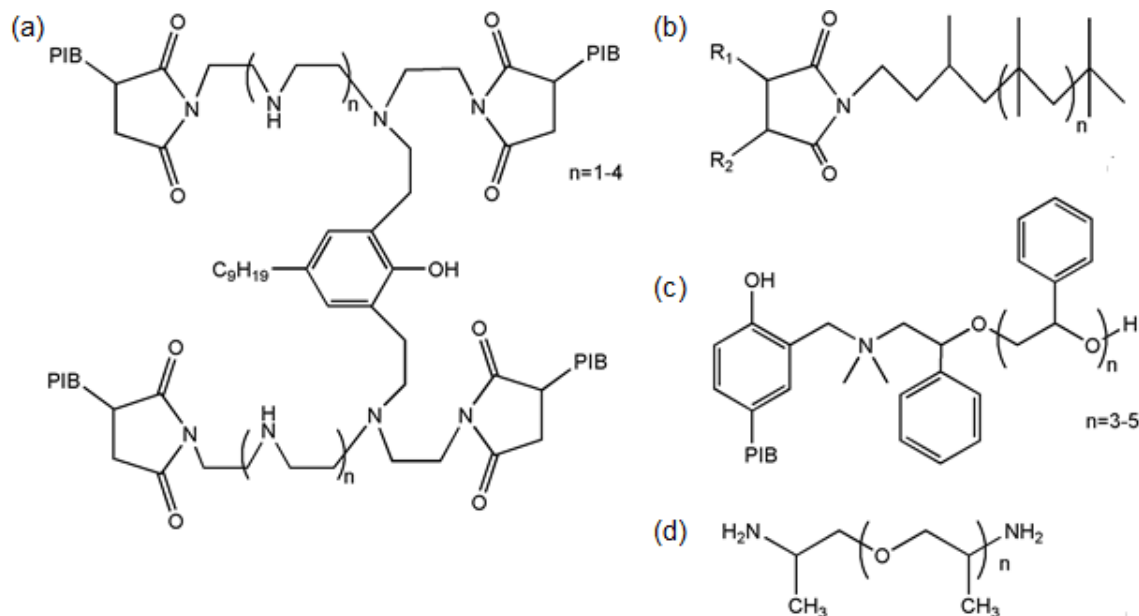
Fonte: adaptada de Oronite, 2014

Por se tratar de segredo industrial, as empresas fornecedoras desses aditivos não revelam a sua composição ao consumidor. Contudo, tem-se um grande número de patentes relacionadas ao desenvolvimento e desempenho de novos aditivos detergentes dispersantes permitindo que se conheça alguns tipos de compostos que são utilizados. Duas características comuns aos aditivos são seu peso molecular e ponto de ebulição elevados. Como exemplos, tem-se os derivados de poliisobutileno (PIB), poliéter aminas, derivados de succinimidas (HERBSTMAN et al., 1991; KELEMEN et al., 2003), sais de amônio quaternário (STEVENSON; THETFORD; VILARDO, 2011; DEROSA, 2012) e aminas (GRAUPNER et al., 2005). Algumas estruturas correspondentes a esses aditivos estão representadas na Figura 2.

Em alguns casos opta-se por utilizar misturas de compostos, como a associação de monoaminas primárias de cadeia carbônica linear e saturada de baixo peso molecular (ex. dodecilamina) com aminas de elevado peso molecular (ex. poliisobutileno etilenodiamina) (GRAUPNER et al., 2005). Isso é feito porque alguns compostos são específicos para a limpeza das válvulas (de injeção e/ou de admissão) enquanto outros são específicos para a limpeza das câmaras de combustão, devendo ser adicionados conjuntamente de modo a promover a limpeza de todas as partes internas do motor. Além disso, alguns compostos como poliisobutileno aminas não devem ser adicionados individualmente, pois, apesar de reduzirem a quantidade de depósitos nas válvulas, contribuem para o aumento dos depósitos na câmara de combustão. Dessa forma, esses compostos devem ser associados a alcóxilatos ou

derivados de succinimidas para que seja observada a remoção dos depósitos tanto nas válvulas quanto na câmara de combustão (KELEMEN et al., 2003).

Figura 2 – Estruturas de alguns compostos utilizados como aditivos detergentes dispersantes para gasolinas: (a) e (b) derivados de succinimida, (c) sal de amônio quaternário e (d) poliéter amina.



Fonte: da autora

Algumas das principais características que os aditivos devem ter para possibilitar sua adição na gasolina são a sua solubilidade nessa e estabilidade em elevadas temperaturas, durante a estocagem (não formar depósitos) e em contato com a água (não formar emulsões) (NIKITINA et al., 2006; KARPOV, 2007). Quanto à estabilidade em altas temperaturas, eles devem permanecer estáveis a cerca de $250\text{ }^{\circ}\text{C}$ e apresentar uma decomposição limpa (sem resíduos) acima da temperatura de operação do motor, cerca de $300\text{ }^{\circ}\text{C}$ (LIN et al., 2002).

No que diz respeito à concentração ideal de aditivo que deve ser adicionada para se obter uma limpeza eficiente, tem-se que esse valor irá variar de um composto para outro, dependendo da sua eficiência. Além disso, a quantidade de aditivo adicionada irá variar se seu uso é regular, para manter o motor limpo (aditivos *keep clean*), ou ocasional, para remover os depósitos existentes (aditivos *clean up*), pois nesse último caso ela costuma ser maior (GRAUPNER et al., 2005).

Pelo que foi observado para alguns compostos patenteados, as concentrações variam de 5 a 5.000 ppm. No caso de alguns derivados de succinimidas e aminas de peso molecular mais baixo a concentração ideal é de até 1.000 ppm, enquanto no

caso de aminas de peso molecular mais elevado ela é de até 2.000 ppm (KELEMEN et al., 2003; GRAUPNER et al., 2005). No Brasil, a concentração adequada de cada aditivo é de conhecimento apenas da ANP, responsável pelo seu registro.

A eficiência dos aditivos, que permite determinar sua faixa de concentração ideal, deve ser comprovada através de ensaios de referência. A ANP instituiu que no ensaio de medição de depósitos nas válvulas de admissão (ABNT NBR 16038) o limite máximo é de 100 mg, valor médio considerando as quatro válvulas, e que a redução do fluxo da gasolina resultante da formação de depósitos nos bicos injetores deve ser no máximo de 5% (ASTM D5598 ou ASTM D6421) (BRASIL, 2014).

Apesar de haver um grande número de registros de aditivos, não são muitos os trabalhos e patentes que descrevem métodos para a identificação e determinação de aditivos detergentes dispersantes em combustíveis. Como a concentração dos aditivos presentes em gasolinas é, em geral, pequena e o aditivo propriamente dito (componente ativo) corresponde a uma parcela de apenas 30 a 60 % da massa, alguns métodos necessitam de uma etapa de pré-concentração da amostra (BRAUER; MIIN, 1995).

Alguns dos métodos propostos que foram encontrados na literatura utilizam a cromatografia por permeação em gel (GPC, do inglês *gel permeation chromatography*), também conhecida como cromatografia por exclusão de tamanho (SEC, do inglês *size exclusion chromatography*). O uso dessa técnica se deve ao fato da massa molar dos aditivos ser muito elevada comparada aos demais constituintes das gasolinas, o que permite separá-los. Outros métodos propostos se baseiam no espalhamento de luz, na reação do aditivo com um indicador que resulte em uma mudança espectral ou de cor e na espectroscopia no infravermelho.

A patente US 5244809 A descreve um método para a quantificação de aditivos detergentes em gasolina ou outros combustíveis derivados do petróleo através da evaporação da amostra seguida pela sua passagem através de um detector evaporativo por espalhamento de luz. Trata-se de um método rápido e que necessita de pouca quantidade de combustível, contudo, não é específico, pois detecta a presença de todas as partículas de elevada massa molecular presentes na amostra. Esse método foi utilizado na análise de amostras contendo até 1,25 % (v/v) do aditivo concentrado (NOWAK, 1993).

A patente CA 2132806 A1 descreve um método para a identificação e quantificação de aditivos detergentes em combustíveis utilizando cromatografia por

permeação em gel com um detector por índice de refração diferencial. Nesse método, são usadas 2 ou 3 colunas, devendo a amostra permanecer cerca de 13 minutos em cada; não é necessário realizar nenhum preparo da amostra, ela é injetada diretamente; e deve-se utilizar uma fase móvel de viscosidade semelhante à da gasolina. Como, nesse método, ocorre a separação dos aditivos dos demais constituintes da gasolina, torna-se possível identificá-los e quantificá-los. Esse método foi utilizado na quantificação de amostras contendo entre 100 e 1.500 ppm de uma mistura de aditivos, a qual é composta por cerca de 30 a 60 % (m/m) do componente ativo (BRAUER; MIIN, 1995).

O trabalho de Colaiocco e Lattanzio descreve um método para a quantificação de aditivos detergentes do tipo amida com peso molecular entre 700 e 800 g.mol⁻¹ em gasolinas. Esse método utiliza a cromatografia por exclusão de tamanho com um detector evaporativo por espalhamento de luz e não necessita de preparo de amostra. Esse método foi utilizado na quantificação de amostras contendo entre 50 e 500 ppm do aditivo (COLAIOCCO; LATTANZIO, 1995).

A patente US 20040214341 A1 descreve um método para a detecção e quantificação de aditivos em matrizes de hidrocarbonetos como a gasolina e o diesel. O método utiliza, pelo menos, duas colunas de cromatografia por permeação em gel e três detectores, sendo o primeiro um detector de absorção no ultravioleta-visível, o segundo um detector por fluorescência e o terceiro um detector por índice de refração, evaporativo por espalhamento de luz ou espectrométrico de massas. Dentre os aditivos que foram analisados estão os detergentes dispersantes derivados de aminas, poliéter aminas, succinimidas, etc. Uma vez que as razões entre os sinais de cada detector são únicas para cada tipo de aditivo, é possível identificá-los. A quantificação foi realizada utilizando as áreas dos picos. Esse método foi utilizado na quantificação de amostras contendo entre 50 e 750 mg/L de misturas de aditivos (FEDEROVA, 2004).

A patente CN 1673737 A descreve um método para a identificação de aditivos detergentes dos tipos aminas e amidas poliolefínicas e poliéter aminas utilizando cromatografia em camada delgada (CCD). Esse método se baseia na pré-concentração da amostra através da destilação em pressão atmosférica ou à vácuo (pré-concentração de 10 a 200 vezes), seguida da filtração e descoloração através da extração em fase sólida e análise por CCD. O método mostrou-se efetivo na análise

de amostras de concentração variando de 20 a 500 mg/L (GUOQING; LILI; MINGXING, 2005).

A patente US 20080190354 A1 descreve a construção e utilização de um kit que permite a identificação de aditivos detergentes dispersantes ou outros compostos de interesse em combustíveis como gasolina e diesel. Os compostos que podem ser identificados são aqueles que possuem grupos básicos como aminas, iminas, amidas, imidas e succinimidas. Nesse método, um indicador com grupos ácidos deve ser adicionado a um recipiente contendo o combustível e gerar uma resposta, a qual pode ser uma mudança no espectro do composto (espectros de absorção, transmissão, emissão) ou, preferencialmente, na cor da solução, para permitir a realização de análises em campo através da comparação da cor com uma cartela de cores. Também através da comparação da cor, esse kit permite a realização de determinações semiquantitativas em campo. Esse método foi utilizado na análise de amostras de diesel contendo 250 e 500 ppm de aditivo (MALPAS et al., 2008).

O trabalho de De Medeiros (2009) descreve um método para a quantificação de três tipos de aditivos detergentes dispersantes em gasolina utilizando a espectroscopia no IV (médio e próximo) associada a técnicas quimiométricas e de pré-concentração das amostras. Duas técnicas de pré-concentração foram utilizadas: destilação atmosférica (com coleta do resíduo para análise) e ensaio da goma não lavada (com coleta da goma para análise). Essa etapa de pré-concentração foi necessária pelo fato da espectroscopia no IR não ter limites de detecção tão baixos quanto as concentrações nas quais os aditivos podem estar presentes (abaixo de 0,1%) (DE MEDEIROS, 2009). Os modelos construídos nesse trabalho apresentaram erros de previsão elevados, entre 81,6 e 248 ppm, e capacidade de detecção mínima (quantidade mínima do aditivo que pode ser quantificado com um erro aceitável) de 153 ppm. Uma vez que os modelos construídos para cada aditivo foram diferentes (valores de R^2 e quantidade de variáveis latentes diferentes), é necessário conhecer previamente o tipo de aditivo contido na gasolina para que se possa quantificá-lo.

O trabalho de Dos Santos (2012) descreve um método para quantificação de três tipos de aditivos detergentes dispersantes, cedidos pela ANP e nomeados G, T e W, em gasolina utilizando a cromatografia por exclusão de tamanho com um detector por índice de refração associada a uma etapa prévia de pré-concentração. A pré-concentração foi realizada destilando-se a gasolina, assim como feito por De Medeiros (2009), e diluindo-se o resíduo obtido em 5 mL de tetrahidrofurano. Nesse trabalho,

os erros percentuais de previsão das amostras variaram entre 6,1 % para o aditivo G e 19,2 % para o aditivo T (DOS SANTOS, 2012). Uma vez que as curvas de calibração de cada aditivo foram diferentes (diferentes coeficientes angulares), esse método também necessita de um conhecimento prévio do tipo de aditivo contido na gasolina para que se possa quantificá-lo.

O trabalho de Da Silva et al. (2014), assim como o De Medeiros (2009), descreve um método utilizando a espectroscopia no IV (médio e próximo) associada a técnicas de análise multivariada e pré-concentração através da destilação atmosférica. Esse trabalho, o qual foi desenvolvido pelo nosso grupo de pesquisa, teve como objetivo desenvolver modelos de classificação para identificar se as gasolinas continham ou não aditivos detergentes dispersantes. O fator de pré-concentração obtido foi de cerca de 100 vezes, pois se partiu de um volume inicial de 100 mL para um volume final de cerca de 1 mL. A porcentagem de classificações corretas das amostras de validação externa foi de 92 % para o NIR e 100 % para o infravermelho médio (DA SILVA et al., 2014).

A patente US 20140038303 A1, a qual é uma melhoria da US 20080190354 A1, também está relacionada à construção de um kit para identificação de aditivos detergentes dispersantes ou outros compostos de interesse pela reação dos seus grupos básicos com os grupos ácidos dos indicadores. A diferença é que ao invés de se adicionar o indicador à gasolina ele está contido em um suporte sólido, como por exemplo papel de filtro de grau cromatográfico, e esse é mergulhado na gasolina. As vantagens desse kit em relação ao anterior é que não se tem problemas quanto à estabilidade das soluções dos indicadores e não é preciso ter frascos para as reações, os quais precisariam ser lavados a cada nova amostra de gasolina. Esse método foi utilizado na análise de amostras de diesel contendo 250, 500 e 1.000 ppm de aditivo (READING; MCFARLANE, 2014).

No Brasil, ainda não há um método padrão para a quantificação dos aditivos detergentes dispersantes na gasolina, mas o desenvolvimento de tal método é necessário para o controle de qualidade das gasolinas comercializadas. Como não há informações a respeito dos tipos de compostos e da concentração que os mesmos serão adicionados na gasolina a partir de 2015, o uso de uma etapa de pré-concentração pode ser importante, uma vez que deve facilitar tanto a análise qualitativa quanto a quantitativa.

3.3. *Ring-oven*

O método do *ring-oven* foi desenvolvido por Herbert Weisz na década de 50 e consiste em um método de análise no qual a amostra é concentrada na forma de um anel fino e bem definido (WEISZ, 1970). Esse método permite a identificação e determinação de componentes presentes em amostras bastante diluídas, as quais não poderiam ser analisadas na forma de *spot tests*, análises simples que permitem identificar compostos através da sua reação com outros compostos conhecidos.

O primeiro artigo sobre o *ring-oven* data de 1954 e foi publicado por Weisz na *Microchimica Acta* (WEISZ, 1970). Após esse primeiro artigo, uma série de outros foram publicados com diferentes aplicações e em diferentes áreas da química analítica, por se tratar de um método simples e sensível.

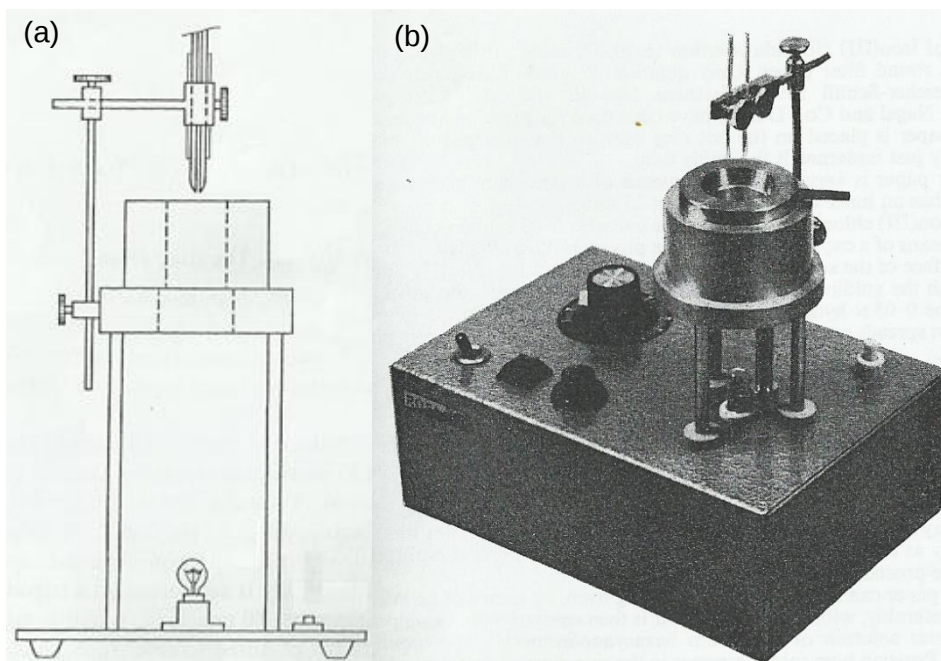
O método foi originalmente proposto para análises de quantidades bastante reduzidas de amostras relacionadas a obras de arte, mas sua simplicidade e rapidez fizeram com que ele fosse aplicado em situações nas quais a quantidade de amostra não era necessariamente um problema (WEISZ, 1970; 1987).

A princípio, ele foi aplicado para determinações qualitativas e semiquantitativas apenas de compostos inorgânicos. Contudo, com o passar do tempo, compostos orgânicos também passaram a ser determinados. Acredita-se que a primeira aplicação do *ring-oven* para a análise de compostos orgânicos tenha sido a identificação e determinação semiquantitativa de compostos nitrogenados em frações de petróleo. Isso foi possível reagindo-se a amostra contendo os compostos nitrogenados com o tetracianoetileno, já que a sua reação com esses compostos forma um produto de cor estável a 110 °C, enquanto que a sua reação com hidrocarbonetos forma produtos que são estáveis apenas à temperatura ambiente (WEISZ, 1970).

O *ring-oven* foi usado para a determinação de diferentes tipos de constituintes presentes em amostras, entre elas: proteínas, corantes, pesticidas, conservantes de alimentos, amostras biológicas, medicamentos, drogas, resíduos de disparos com armas, gorduras, óleos, etc (WEISZ, 1970; 1987). A sociedade americana para testes e materiais (ASTM, do inglês *american society for testing and materials*) chegou, inclusive, a propor o uso do método para a determinação de partículas de metais contaminantes em componentes eletrônicos e microeletrônicos (WEISZ, 1970).

Quando foi proposto, o aparato (Figura 3a) consistia em um bloco cilíndrico de alumínio de 35 mm de altura e 55 mm de diâmetro com um orifício central de 22 mm de diâmetro colocado sobre um bloco de diâmetro maior com o mesmo orifício central. O primeiro bloco possuía uma resistência na forma de um fio em seu interior e acima dele ficava um pequeno tubo de vidro de 60 mm de comprimento, o qual servia como guia para o capilar e podia ter sua posição ajustada. Além desses componentes, havia uma pequena lâmpada elétrica posicionada abaixo dos blocos de modo a permitir a visualização das bordas do anel à medida que ele fosse sendo formado (WEISZ, 1970). O papel de filtro ficava preso no bloco de aquecimento por um anel metálico, o qual não é mostrado na figura. Um exemplo de um equipamento disponível comercialmente, em 1970, pode ser visto na Figura 3b.

Figura 3 – (a) Aparato original do *ring-oven* e (b) equipamento comercial disponível em 1970.



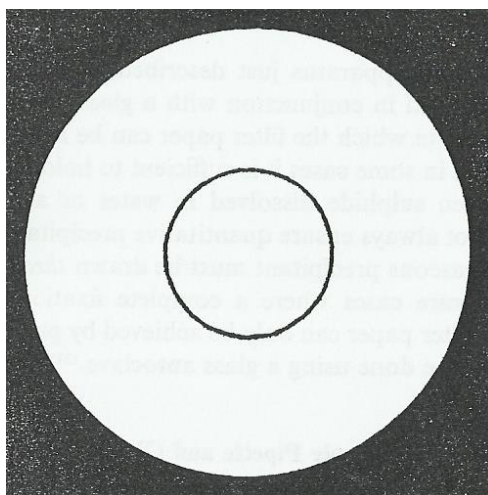
Fonte: adaptada de Weisz, 1970

Nesse método, uma gota da amostra (cerca de 1 μL) é colocada em um papel de filtro, a fim de formar uma mancha, e o papel é colocado no centro do *ring-oven*. O solvente é, então, gotejado e a mancha é lavada a partir do seu centro de modo a arrastar os componentes solúveis da amostra para as extremidades do anel. O anel se forma porque, ao atingir as extremidades aquecidas do bloco, o solvente de lavagem evapora e os componentes da amostra que foram carregados por ele ficam

retidos no papel. É importante observar que, antes da lavagem, é possível fixar alguns componentes da amostra no papel através de reações químicas específicas. Essa fixação permite separar esses componentes daqueles que serão lavados pelo solvente, a fim de permitir sua posterior identificação.

Após a formação dos anéis, é possível identificar os componentes da amostra realizando testes com reagentes que provoquem a formação de compostos insolúveis e, preferencialmente, coloridos. Um exemplo de reação que forma um precipitado colorido é a de uma solução de cloreto de ferro (III) com ferrocianeto de potássio, a qual resulta em um precipitado de coloração azul (Figura 4). Em alguns casos, o precipitado formado não é colorido, mas apresenta, por exemplo, emissão de fluorescência, o que permite identificá-lo (WEISZ, 1970). No caso de se desejar identificar mais de um componente, o anel pode ser cortado em setores e cada setor usado para identificar um componente distinto.

Figura 4 – Anel formado a partir da reação de cloreto de ferro (III) e ferrocianeto de potássio.



Fonte: Weiz, 1970

Ao longo dos anos, novas aplicações foram surgindo para o *ring-oven* e algumas modificações foram sendo feitas na sua estrutura original. Foram desenvolvidos equipamentos portáteis de vidro para análises de poluentes no ar; equipamentos com vários orifícios no bloco de aquecimento, para permitir a concentração de mais de uma amostra em um mesmo papel; e equipamentos rotatórios que permitem a combinação do *ring-oven* com a cromatografia em coluna, funcionando como um coletor de frações da coluna (WEISZ, 1987). Além de

modificações no equipamento, alguns autores propuseram o desenvolvimento de acessórios para reduzir o tamanho do orifício de 22 para 14 mm de diâmetro, possibilitando a formação de dois anéis concêntricos em um só papel (WEISZ, 1970). Com esse acessório, o primeiro anel era formado, alguns componentes eram precipitados e os outros eram submetidos a uma segunda lavagem usando o diâmetro usual, para formar um anel maior.

Quando o *ring-oven* foi desenvolvido, a análise dos anéis era feita, principalmente, através da comparação colorimétrica de anéis obtidos a partir das amostras desconhecidas com anéis de soluções padrão. No caso de produtos que não eram coloridos, usavam-se propriedades como a emissão de fluorescência e também era feita uma análise visual. Atualmente, diante do avanço das técnicas instrumentais, é possível analisar os anéis tanto qualitativamente quanto quantitativamente empregando técnicas espectroscópicas que permitam a análise de pequenas áreas. Algumas técnicas que podem ser usadas são a espectroscopia de plasma induzido por laser, a espectroscopia no infravermelho (com análise pontual) e as imagens hiperespectrais.

Um exemplo dessa associação do *ring-oven* com novas técnicas espectroscópicas é o trabalho de Cortez e Pasquini (2013), no qual realizou-se a determinação de metais (sódio, ferro e cobre) em etanol combustível. Nesse trabalho as amostras foram pré-concentradas nos anéis e esses foram, em seguida, analisados usando a LIBS. O *ring-oven* foi usado para pré-concentrar os metais porque os valores máximos de concentração estabelecidos pela ANP (2 e 5 µg/g) são muito baixos para possibilitar a análise direta das amostras (CORTEZ; PASQUINI, 2013).

Esse trabalho mostrou a possibilidade de se associar o *ring-oven* às novas técnicas analíticas disponíveis, permitindo a resolução de problemas que envolvam amostras contendo analitos em baixas concentrações. A evolução das técnicas analíticas traz, então, novas possibilidades para o emprego desse método, permitindo sua aplicação em amostras dos mais variados tipos e em diversas áreas.

3.4. Infravermelho próximo

A radiação na região do infravermelho próximo, a qual está compreendida entre 750 e 2.500 nm ou 4.000 e 13.300 cm⁻¹, é um tipo de radiação menos energética que a ultravioleta e visível e que é responsável, quando absorvida, pelas transições entre

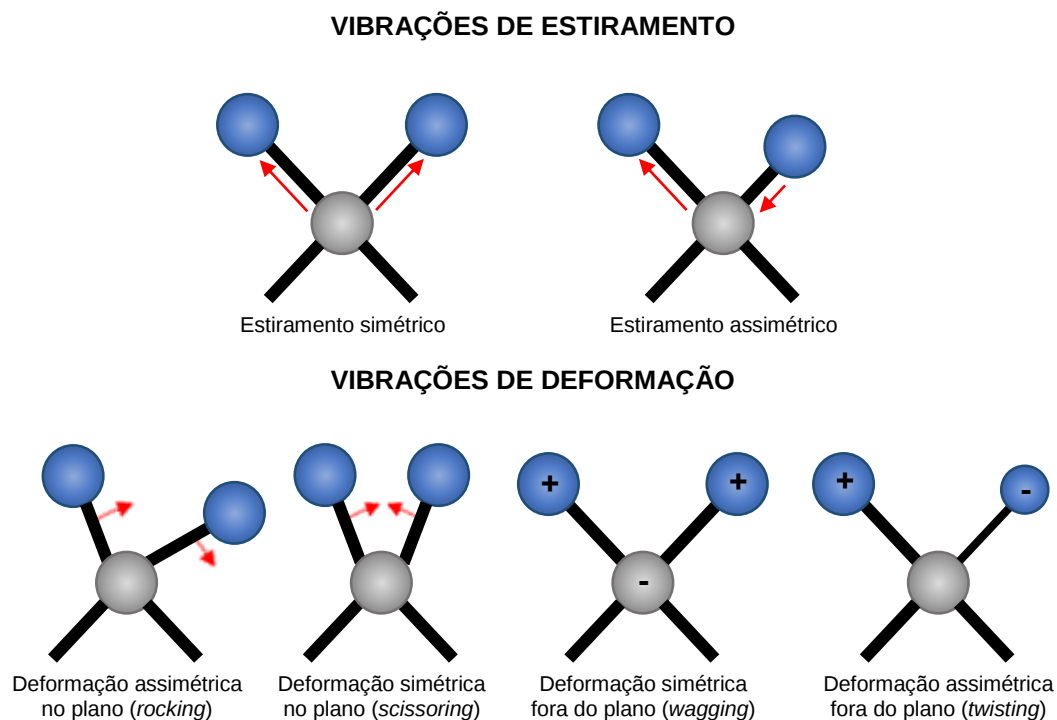
estados vibracionais de uma molécula (PASQUINI, 2003). Pelo fato das bandas serem largas e muito sobrepostas, a interpretação espectral não é direta como no caso do infravermelho médio. Por isso, essa região do espectro foi negligenciada por muito tempo até que fossem desenvolvidas ferramentas matemáticas capazes de extrair informações qualitativas e quantitativas desses dados (técnicas de análise multivariada).

A espectroscopia na região do NIR é uma técnica de análise rápida, não destrutiva, o feixe de luz da radiação tem uma penetração elevada e demanda um preparo mínimo ou nenhum da amostra, pois em muitos casos é possível realizar a análise direta (PASQUINI, 2003). É importante observar que em algumas situações, como na análise de amostras cujo analito está em baixa concentração (abaixo do limite de detecção da técnica), uma etapa de pré-concentração precisa ser empregada para permitir sua análise.

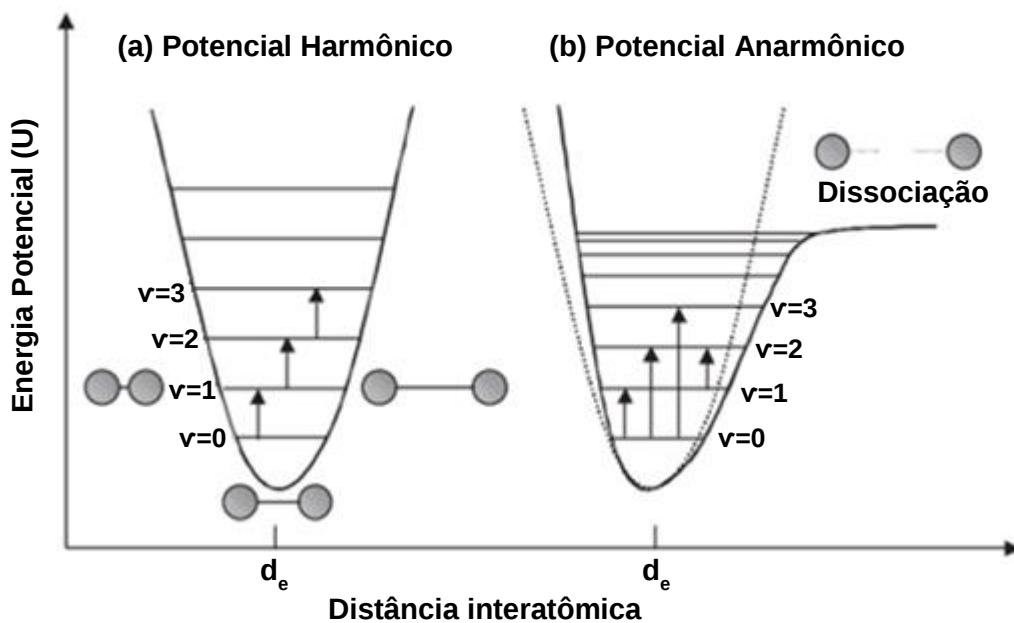
Os espectros NIR se originam de variações de energia resultantes de transições entre níveis vibracionais. Essas transições só ocorrem, entretanto quando a energia da radiação incidente corresponde à diferença de energia entre os níveis. Além disso, a radiação incidente só interage com a molécula quando a vibração é ativa, o que ocorre quando essa vibração provoca uma mudança no momento de dipolo da molécula (PASQUINI, 2003).

As vibrações podem ser de dois tipos: de estiramento, quando envolve variação na distância interatômica ao longo do eixo da ligação, ou de deformação, quando envolve variação no ângulo entre duas ligações. As vibrações de estiramento e deformação podem ser simétricas ou assimétricas, sendo que as últimas podem ser no plano ou fora dele, como pode ser visto na Figura 5 (HOLLER; SKOOG; CROUCH, 2009). Nesta Figura, o símbolo (+) indica o movimento para fora da página, em direção ao leitor, e o (-) para dentro da página, se afastando do leitor.

As vibrações de estiramento moleculares podem ser aproximadas por um modelo simplificado de duas massas conectadas por uma mola, o qual apresenta níveis discretos de energia (HOLLER; SKOOG; CROUCH, 2009). O movimento de duas massas (movimento harmônico simples) pode ser descrito pelo modelo do oscilador harmônico (Figura 6a), contudo, no caso das moléculas esse modelo apresenta inconsistências.

Figura 5 – Tipos de vibrações moleculares.

Fonte: da autora

Figura 6 – Representação esquemática do modelo do oscilador (a) harmônico e (b) anarmônico para a variação de energia potencial (U) de moléculas diatômicas.

Fonte: adaptada de Pasquini, 2003

No modelo harmônico, tem-se que apenas transições entre níveis vibracionais adjacentes ($\Delta v=1$) são observadas, as diferenças de energia entre esses níveis

adjacentes são sempre iguais e as vibrações são independentes, não sendo possível haver combinações de vibrações. As inconsistências nesse modelo, no que diz respeito à descrição das ligações, são que transições com $\Delta v=2$ ou maiores (sobretons) e combinações de vibrações, ambas proibidas pelo modelo, são observadas nos espectros NIR. Além disso, as repulsões das nuvens eletrônicas, no momento da aproximação dos átomos, fazem com que a energia potencial do sistema aumente mais rapidamente do que o descrito no modelo e o afastamento desses átomos acaba resultando na dissociação dos mesmos, levando a uma estabilização da energia (PASQUINI, 2003).

Um modelo mais adequado para descrever os sistemas moleculares é o do oscilador anarmônico (Figura 6b). Esse modelo, entre outras coisas, prevê a ocorrência das bandas de sobretons e de combinações observadas nos espectros NIR e descreve uma diminuição da diferença de energia entre níveis vibracionais à medida que o número quântico vibracional (v) aumenta.

As bandas que são observadas nos espectros NIR são aquelas resultantes de vibrações envolvendo ligações entre átomos de hidrogênio e outros átomos mais pesados (ligações O-H, C-H, N-H e S-H). Comparando-se as intensidades das bandas de combinação e dos sobretons com as das transições fundamentais, observadas no infravermelho médio, tem-se que elas são 10 e 1000 vezes menores, respectivamente (PASQUINI, 2003).

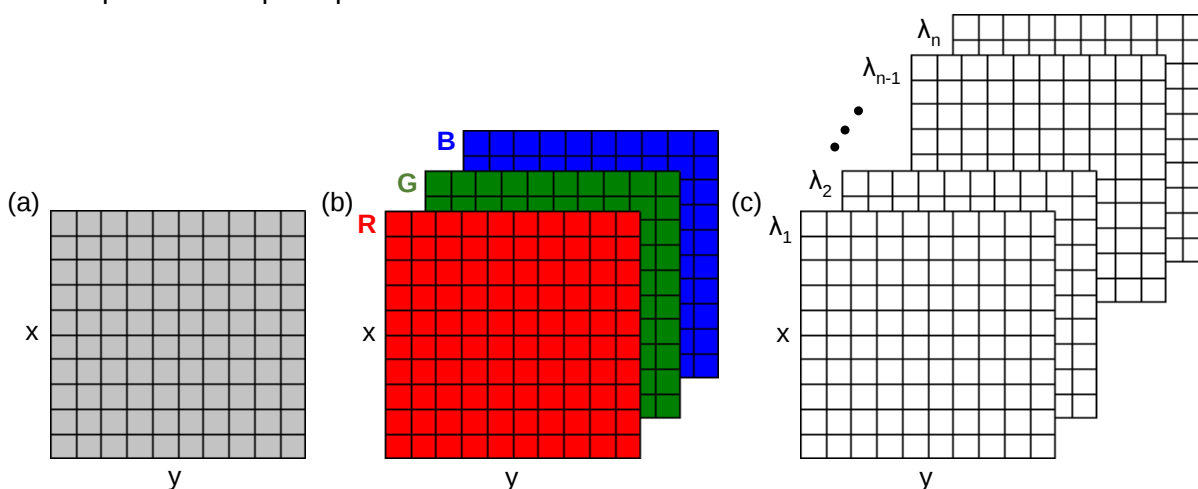
Várias são as formas de obtenção de espectros NIR, entre elas a transmitância ou absorbância, reflectância difusa, transflectância e interactância (PASQUINI, 2003). Uma forma de aquisição de dados na região do NIR que tem ganhado destaque nos últimos anos é a imagem hiperespectral, na qual para cada *pixel* se tem um espectro completo.

3.5. Imagens hiperespectrais

As imagens são representações espaciais de objetos e são compostas por elementos chamados *pixels*, no caso de imagens com duas dimensões espaciais, ou *voxels*, no caso de imagens com três dimensões espaciais, que são as imagens de múltiplas camadas de uma amostra (PRATS-MONTALBÁN; DE JUAN; FERRER, 2011). Como pode ser visto na Figura 7, além das dimensões x e y dos *pixels*, ou x , y e z dos *voxels*, as imagens são formadas por uma dimensão adicional, a qual pode

ser representada por um único valor ou canal (ex. imagens em escala de cinza), três valores (ex. imagens RGB que são compostas por valores relacionados às cores vermelha, verde e azul) ou diversos valores (ex. imagens multiespectrais ou hiperespectrais dependendo da quantidade de canais espectrais). Em geral, considera-se que as imagens hiperespectrais são aquelas com mais de 100 canais espectrais (GELADI; GRAHN; BURGUER, 2007).

Figura 7 – Tipos de imagens: (a) imagem em escala de cinza, (b) imagem RGB e (c) imagem multiespectral ou hiperespectral.



Fonte: da autora

Nas imagens hiperespectrais, cada *pixel* é representado por um espectro que pode ser, por exemplo, um espectro NIR ou Raman. Assim, as imagens hiperespectrais têm uma estrutura cúbica e são compostas por milhares ou até milhões de pontos, como é o caso das imagens analisadas neste trabalho, as quais possuem mais de 9 milhões de pontos (200 *pixels* x 200 *pixels* x 226 canais espectrais). No caso das imagens de múltiplas camadas, a estrutura é de um hipercubo e o número de pontos é ainda maior.

Com o desenvolvimento de novos equipamentos para a aquisição de imagens hiperespectrais, computadores, softwares e algoritmos capazes de lidar com os dados das imagens e extrair suas informações, tem-se observado uma expansão do seu uso em diversas áreas de conhecimento (VIDAL; AMIGO, 2012). A principal vantagem das imagens com relação à espectroscopia convencional é a obtenção de informações sobre a distribuição dos compostos na superfície da amostra, o que é possível devido à associação da informação espectral com a espacial (PRATS-MONTALBÁN; DE JUAN; FERRER, 2011). A obtenção desse tipo de informação possibilitou o uso das

imagens em aplicações farmacêuticas relacionadas à distribuição e homogeneidade de componentes em comprimidos (AMIGO et al., 2008; FRANCH-LAGE et al., 2011). Outras áreas como a médica (PIQUERAS et al., 2011) e forense (BRAZ; LÓPEZ-LÓPEZ; GARCÍA-RUIZ, 2013) também têm se beneficiado das imagens hiperespectrais.

Por se tratar de uma grande quantidade de dados espectrais, o uso de ferramentas quimiométricas é indispensável na extração das informações das imagens (DE JUAN et al., 2009). Muitas das técnicas e algoritmos utilizados na análise multivariada clássica foram adaptadas para a análise de imagens e, adequando-se às novas necessidades, outros foram desenvolvidos. A escolha da técnica e algoritmo depende do objetivo da análise. Caso se deseje, por exemplo, realizar uma análise exploratória para identificar similaridades entre os *pixels* é possível utilizar a análise de componentes principais.

É importante ressaltar que, algumas vezes, a imagem obtida não compreende só a região de interesse (ROI, do inglês *region of interest*) como também partes do fundo e outras regiões que não interessam. Isso pode ocorrer, por exemplo, na análise de comprimidos, uma vez que as amostras têm um formato circular e a área onde são obtidos os dados espectrais é, normalmente, quadrada ou retangular, o que resulta na obtenção de informações que não estão associadas à amostra. Uma vez que nem toda a imagem possui informações da amostra, a porção indesejada deve ser removida previamente à análise.

Muitas vezes, a seleção da ROI não é trivial, por se ter regiões bastante semelhantes, por exemplo. Algumas maneiras de se realizar essa seleção são: seleção manual (na imagem global, em um comprimento de onda ou na imagem dos escores da PCA), seleção através do uso de histogramas e seleção manual de valores limites na imagem dos escores da PCA (VIDAL; AMIGO, 2012). A seleção manual, apesar de garantir a seleção apenas da ROI, é subjetiva e muito demorada, o que é uma grande desvantagem quando se deseja analisar muitas imagens. A seleção através dos histogramas e a determinação manual dos limites também não são isentas de problemas, pois diferenças sutis de composição podem fazer com que não existam limites bem definidos que permitam uma boa separação entre a ROI e o restante da imagem. Isso pode resultar na inclusão de partes de outras regiões que não interessam ou perda de parte da ROI. Nesse caso, é importante definir o que é mais importante, incluir toda a ROI, mesmo que isso resulte na seleção de partes de

outras regiões; ou garantir que apenas a ROI seja selecionada, mesmo que isso resulte na perda de partes dela.

O processamento de imagens pode ser dividido em três subáreas: pré-processamento, compressão e análise ou processamento de imagens propriamente dito (PRATS-MONTALBÁN; DE JUAN; FERRER, 2011). A etapa de pré-processamento consiste na correção dos defeitos das imagens e dos espectros, como a presença de *pixels* anômalos ou espectros com ruídos e desvios de linha de base. A seleção da ROI pode ser considerada parte da etapa de pré-processamento das imagens. A compressão, por sua vez, consiste na redução da imagem em termos de memória utilizada. Por fim, a análise ou processamento de imagens consiste na extração de informações relevantes. Os resultados dos processamentos podem ser expressos na forma de números (ex. valores de concentração), vetores (ex. histogramas) ou novas imagens (ex. imagens dos escores ou mapas de concentração) (GELADI; GRAHN; BURGUER, 2007)

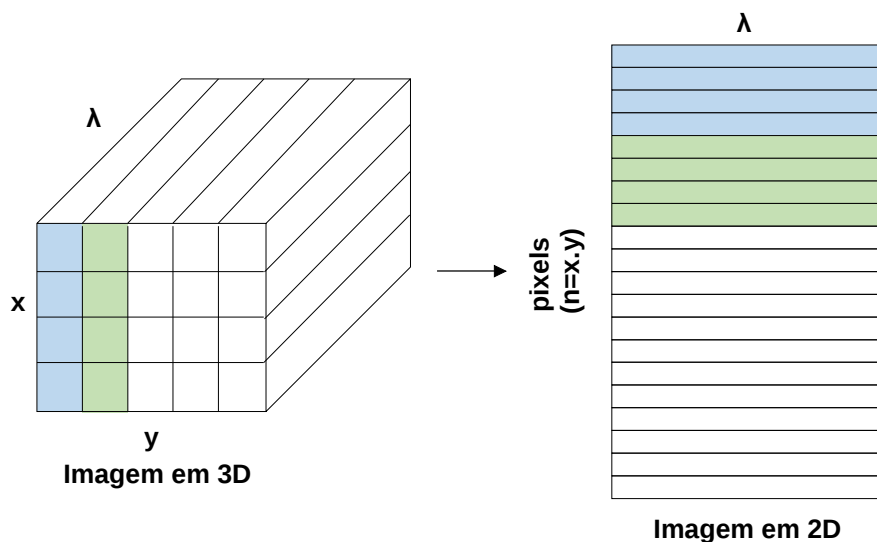
Além dos problemas que afetam os dados espectrais usuais, as imagens são afetadas pela presença de *pixels* anômalos (chamados de *dead pixels* ou *outliers*) (DE JUAN et al., 2009). Esses *pixels* são representados por espectros com uma determinada quantidade de valores zero ou ausentes (chamados de *missing values*) e costumam resultar de anomalias nos detectores. A presença desses *pixels*, os quais podem estar sozinhos, em grupos ou linhas, afeta os resultados de algumas técnicas como a PCA, que é fortemente influenciada pela presença de *dead pixels* (VIDAL; AMIGO, 2012). Assim, a identificação e substituição desses *pixels* por valores médios ou medianos dos *pixels* vizinhos, a qual pode ser feita através de alguns algoritmos disponíveis, deve ser realizada antes da compressão ou processamento da imagem.

Apesar do desenvolvimento de computadores cada vez mais potentes, a compressão dos dados pode ser uma etapa importante quando se trabalha com uma grande quantidade de imagens ou elas possuem elevada resolução (VIDAL; AMIGO, 2012). A compressão pode ser feita de várias maneiras, algumas delas são: mudança do formato da imagem (*bytes encoding*), seleção de variáveis e redução da resolução espacial ou espectral através da substituição dos valores de um intervalo por um valor representativo (*data binning*) (VIDAL; AMIGO, 2012) (VIDAL; AMIGO, 2012) (VIDAL; AMIGO, 2012) (VIDAL; AMIGO, 2012) (VIDAL; AMIGO, 2012) (VIDAL; AMIGO, 2012). A compressão utilizando *data binning* resulta em perda de resolução, contudo, muitas imagens podem ser comprimidas sem perder informações relevantes (DE JUAN et al., 2009).

A escolha do tipo de compressão depende do tipo de dado que se tem e do objetivo da análise.

Antes de aplicar as técnicas de pré-processamento e de análise nas imagens, uma prática comum é desdobrar a matriz de dados (cubo) em uma matriz de duas dimensões, na qual os *pixels* são dispostos nas linhas e os canais espectrais (números ou comprimentos de onda) nas colunas. No software MATLAB (MathWorks) a função usada para este propósito é a *reshape*, a qual decompõe a matriz de dados de cima para baixo, como mostrado na Figura 8.

Figura 8 – Desdobramento da matriz de dados segundo a função *reshape*.



Fonte: da autora

3.5.1. Técnicas de pré-processamentos

Após o desdobramento da matriz de dados é possível aplicar as técnicas de pré-processamento espectral utilizadas na análise multivariada clássica, como a normalização, suavização com filtro de Savitzky-Golay, variação normal padrão (SNV, do inglês *standard normal variate*), correção do espalhamento multiplicativo (MSC, do inglês *multiplicative scatter correction*) e derivadas espectrais, como a derivada com filtro de Savitzky-Golay. No caso dos espectros de refletância difusa no NIR, é comum se observar a presença de desvios de linha de base decorrentes do espalhamento da radiação. Assim, as técnicas de pré-processamento mais usadas, nesse caso, são as que corrigem esses desvios: SNV, MSC e derivadas espectrais. A vantagem de se

usar SNV ou MSC é que essas técnicas não alteram significativamente o formato original dos espectros.

A normalização é utilizada quando se deseja eliminar efeitos de variação resultantes de diferenças de concentração ou quantidade de amostra. Ela é realizada dividindo-se as variáveis espectrais por uma constante, a qual pode ser, por exemplo, o valor máximo do espectro (normalização pelo máximo). Outro tipo de normalização é a pela faixa, que divide as variáveis pelo valor do intervalo entre o valor máximo e mínimo do espectro. Esse tipo de normalização é usado no lugar da normalização pelo máximo quando os espectros apresentam valores positivos e negativos, como é o caso dos espectros derivados. Além dessas, há a normalização pela área, pela média, pelo vetor unitário, etc.

Na suavização com filtro de Savitzky-Golay é feito o ajuste de um polinômio a uma janela ou intervalo simétrico de pontos do espectro e o ponto central dessa janela tem seu valor substituído pelo valor do polinômio nesse ponto. Tanto a ordem do polinômio que é ajustado quanto o tamanho da janela (número de pontos que ela inclui) podem ser variados. O objetivo dessa técnica de pré-processamento é minimizar os ruídos espectrais e a escolha do tamanho da janela é uma etapa importante, pois janelas muito pequenas podem não eliminar os ruídos e janelas muito grandes podem eliminar informações espectrais importantes, presentes em bandas estreitas.

Na SNV é feito um autoescalamento nas amostras, cada variável tem seu valor corrigido subtraindo-se a média dos valores do espectro (m_i) e dividindo-se o valor resultante pelo desvio padrão do espectro (s_i), como pode ser visto na Equação 1 (FEARN et al., 2009).

$$X_{SNV} = \frac{X_{original} - m_i}{s_i} \quad (1)$$

Na MSC, por outro lado, o espectro é corrigido com base em um espectro de referência que é, em geral, o espectro médio. Nesse caso, são calculados os coeficientes (b_0 e b_1) e o erro (e) que relacionam o espectro original ao de referência (Equação 2) e o espectro corrigido é obtido através da subtração e divisão por esses coeficientes (Equação 3) (RINNAN; BERG; ENGELSEN, 2009).

$$X_{original} = b_0 + b_1.X_{referência} + e \quad (2)$$

$$X_{MSC} = \frac{x_{original} - b_0}{b_1} \quad (3)$$

Comparando-se essas duas técnicas, tem-se que a vantagem da SNV é que não é preciso ter um espectro de referência, cada espectro é corrigido independentemente dos demais. Isso é importante, principalmente, quando o pré-processamento espectral é feito antes da seleção da ROI, pois os espectros que não sejam da ROI não irão influenciar os espectros da ROI. É importante observar que, para a maioria dos problemas, essas duas técnicas têm resultados semelhantes.

Na derivada com filtro de Savitzky-Golay é feito o ajuste de um polinômio a uma janela simétrica do espectro, como é feito na suavização, a derivada do polinômio é calculada e o valor obtido para o ponto central da janela substitui o valor desse ponto. O objetivo do ajuste do polinômio é promover a suavização do espectro, já que a razão sinal/ruído tende a diminuir com as derivadas. A ordem da derivada é escolhida dependendo do objetivo do pré-processamento, a primeira derivada corrige o desvio de linha de base (efeitos aditivos) enquanto que a segunda derivada remove não só o desvio de linha de base, mas também os efeitos multiplicativos (RINNAN; BERG; ENGELSEN, 2009).

Algumas vezes, a utilização de um só pré-processamento não é suficiente para corrigir os espectros e, portanto, é preciso combiná-los. A escolha do tipo de pré-processamento adequado para o problema em estudo deve levar em conta, basicamente, duas coisas: os erros dos modelos construídos após o pré-processamento e o número de variáveis latentes utilizadas nos mesmos, no caso da regressão por mínimos quadrados parciais (PLS, do inglês *partial least squares*) (RINNAN; BERG; ENGELSEN, 2009).

No que diz respeito à análise das imagens, algumas estratégias que podem ser usadas são: segmentação (identificação de *pixels* de composição semelhantes), resolução (obtenção dos espectros puros dos componentes e dos seus mapas de distribuição) e regressão (utilização das imagens para propósitos quantitativos) (PRATS-MONTALBÁN; DE JUAN; FERRER, 2011). A PCA é um exemplo de uma técnica de segmentação, MCR de uma técnica de resolução e PLS de uma técnica de regressão.

Algumas vezes, é possível utilizar espectros representativos de imagens (ex. espectro médio da ROI) e os algoritmos de análise multivariada clássica. A diferença de se utilizar as imagens hiperespectrais no lugar da espectroscopia convencional

pode ser, por exemplo, a necessidade de se utilizar a informação espacial da imagem na seleção dos espectros. Assim, é possível construir modelos de regressão ou de classificação usando os algoritmos clássicos, como PLS e análise discriminante linear (LDA, do inglês *linear discriminant analysis*), para analisar imagens.

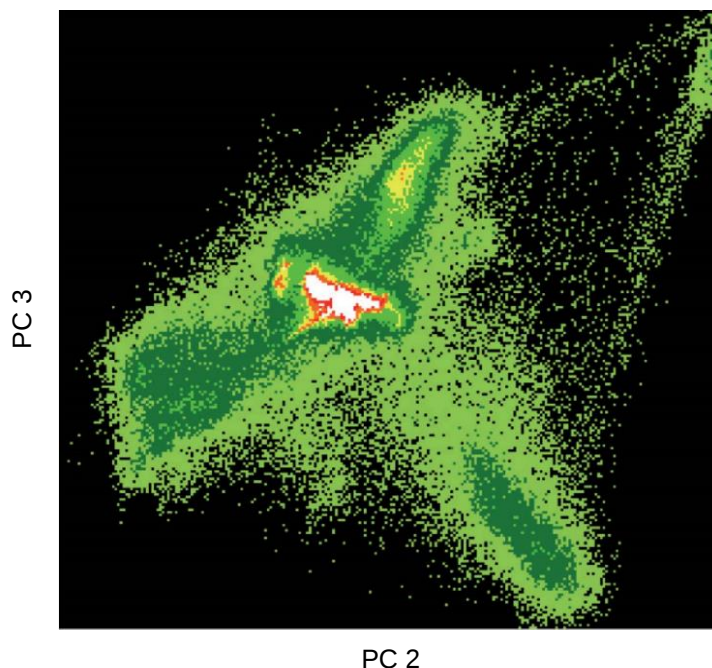
3.5.2. Análise de componentes principais

A PCA pode ser aplicada às imagens com vários objetivos, dentre eles a exploração visual, através das imagens dos escores; compressão, através redução da dimensionalidade dos dados (sem perdas significativas de informação); e identificação de agrupamentos (*clusters*). A PCA cria variáveis ortogonais, componentes principais (PCs, do inglês *principal components*), que explicam a maior variância dos dados. Ela decompõe a matriz dos dados em três matrizes (Equação 4): escores (matriz **T**), pesos (matriz **P**) e erros (matriz **E**). A matriz dos escores indica a posição das amostras ou *pixels*, no caso das imagens, nas novas variáveis (componentes principais) enquanto que a matriz de pesos indica a contribuição de cada variável original para as novas variáveis.

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E} \quad (4)$$

No caso da análise de imagens, a matriz dos escores pode ser reorganizada espacialmente na forma de novas imagens (imagens de escores), através das quais é possível identificar os diferentes agrupamentos. Analisando-se essas imagens juntamente com os pesos é possível identificar que regiões do espectro estão mais relacionadas com cada agrupamento. Quando se conhece a composição da amostra é possível identificar quais componentes estão em cada agrupamento. Outra forma de analisar os resultados da PCA é através da observação dos gráficos de dispersão (*score plots*), gráficos de uma PC versus outra no qual cada ponto corresponde a um *pixel* da imagem. Nesses gráficos, cores são usadas para definir a densidade de *pixels* nos escores, como pode ser visto na Figura 9, na qual as cores branca, vermelha e amarela foram usadas para identificar regiões de elevada densidade, enquanto que a cor verde (claro e escuro) foi usada para identificar regiões de baixa densidade (GELADI; GRAHN; BURGUER, 2007).

Figura 9 – Gráfico de dispersão da PC 3 versus a PC 2.

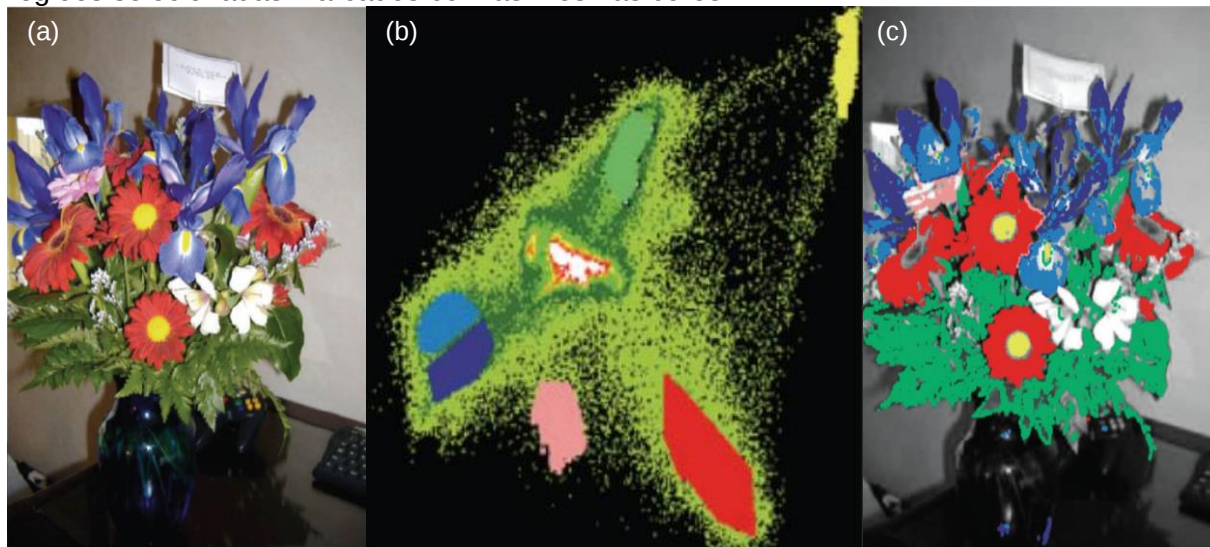


Fonte: adaptada de GELADI; GRAHN; BURGUER, 2007.

Alguns programas computacionais de análise de imagens como o Evince (UmBio) permitem selecionar *pixels* nos gráficos de dispersão e selecioná-los ou realçá-los nas imagens dos escores, de forma a observar sua posição na imagem. *Pixels* que se localizam próximos nos gráficos de dispersão têm propriedades semelhantes, assim, é possível aplicar máscaras a esses gráficos para selecionar *clusters*. Assim, pode-se, por exemplo, selecionar os *pixels* da ROI e eliminar os demais. Um exemplo da aplicação de máscaras para a seleção de diferentes *clusters* está representado na Figura 10. É importante observar que esses gráficos permitem, também, identificar a presença de *outliers*, *pixels* com valores de escores extremos e que se localizam, portanto, distantes dos grupos (GELADI; GRAHN; BURGUER, 2007).

Alguns algoritmos específicos para análise de imagens utilizam a matriz cúbica como entrada, mas também é possível construir os modelos utilizando os algoritmos da análise multivariada clássica, desde que se desdobre a matriz antes.

Figura 10 – (a) Imagem RGB original, (b) gráfico de dispersão da PC 2 versus PC 3 com 6 regiões selecionadas por polígonos de cores diferentes e (c) imagem com os *pixels* das regiões selecionadas marcados com as mesmas cores.



Fonte: adaptada de GELADI; GRAHN; BURGUER, 2007.

3.5.3. Resolução multivariada de curvas

Os métodos de resolução de curvas multivariadas se baseiam na lei de Beer-Lambert, a qual relaciona o valor de absorbância com a concentração de um determinado composto, e têm como objetivo separar as informações dos diferentes constituintes da amostra (PRATS-MONTALBÁN; DE JUAN; FERRER, 2011). Nesses métodos, a matriz de dados é decomposta em três novas matrizes (Equação 5): dos espectros puros (matriz **S**), das concentrações (matriz **C**) e dos erros (matriz **E**), que são os resíduos não explicados pelo modelo. A matriz **C** é composta pela concentração relativa de cada componente em cada *pixel* e pode ser reorganizada espacialmente na forma de um mapa de concentração, a fim de permitir a análise da distribuição dos componentes na imagem.

$$\mathbf{X} = \mathbf{C} \cdot \mathbf{S}^T + \mathbf{E} \quad (5)$$

A resolução de curvas multivariadas com os mínimos quadrados alternados (MCR-ALS, do inglês *multivariate curve resolution - alternating least squares*) é um algoritmo iterativo no qual são dadas estimativas iniciais dos espectros puros ou dos mapas de concentração e são impostas restrições (ex. não negatividade das concentrações) para a resolução. A partir das estimativas iniciais, o algoritmo otimiza

as matrizes das concentrações e espectros puros ajustando-as aos dados experimentais. A otimização é realizada até que a diferença nos desvios padrão de dois resultados consecutivos seja menor do que um determinado valor, o que significa que o resultado convergiu, ou quando o número de ciclos, previamente determinado, chega ao fim (JAUMOT et al., 2005).

A vantagem da MCR-ALS está relacionada com a diversidade e flexibilidade das restrições que podem ser usadas tanto na matriz de concentrações quanto dos espectros puros. Quando se conhece a concentração ou o espectro puro de algum dos compostos da amostra, por exemplo, é possível fixá-la, dessa forma ela não irá variar enquanto as demais variam. Outras restrições que podem ser usadas são a não negatividade, a qual força os valores das concentrações e, caso se deseje, dos espectros a serem zero ou positivos; unimodalidade, que permite apenas um máximo por perfil de concentração; e balanço de massas (do inglês *closure*), que permite, por exemplo, fixar um valor máximo para a concentração total (JAUMOT et al., 2005). Diferentes restrições podem ser aplicadas às matrizes de espectros e concentrações e, até mesmo, a diferentes imagens analisadas simultaneamente (MCR-ALS aumentado), de forma a obter os melhores resultados.

Algumas situações em que é vantajoso utilizar a forma aumentada da MCR-ALS são: análise de várias camadas de uma mesma amostra; de amostras diferentes, mas contendo componentes em comum (ex. comprimidos de diferentes lotes); e de uma amostra em função do tempo, temperatura ou outra variável (PRATS-MONTALBÁN; DE JUAN; FERRER, 2011). A análise simultânea de imagens resulta em melhores resultados, sendo, portanto, uma estratégia que pode ser usada, por exemplo, para melhorar a resolução de amostras com baixas concentrações dos componentes.

3.5.4. Regressão por mínimos quadrados parciais

As imagens podem ser usadas tanto para se obter informações qualitativas, no que diz respeito à composição química de uma amostra e à distribuição dos seus componentes, quanto quantitativas. Modelos baseados na regressão pelo PLS podem ser usados para prever a concentração de um determinado componente na imagem como um todo ou em cada *pixel* individualmente, o que permite a obtenção de mapas de concentração com valores reais.

Nas imagens, tem-se um valor de concentração relacionado a amostra como um todo (todos os *pixels* da ROI) e, por isso, na etapa de calibração e validação utiliza-se um espectro representativo da ROI. Os espectros representativos podem ser, por exemplo, um espectro médio ou a soma de todos os espectros, de forma a contabilizar os valores contidos em todos os *pixels* da ROI. Na etapa de previsão, contudo, tanto se pode estimar o valor da concentração da amostra, nesse caso se utiliza um espectro representativo, como se pode estimar o valor de concentração de cada *pixel*. Nesse último caso, os espectros de todos os *pixels* devem ser utilizados (PRATSMONTALBÁN; DE JUAN; FERRER, 2011).

No caso de se realizar as três etapas (calibração, validação e previsão) utilizando-se espectros representativos, é possível utilizar os algoritmos das análises clássicas e a análise é feita da mesma forma como nos dados espectroscópicos clássicos. No caso da previsão ser realizada para cada espectro individualmente, por outro lado, a análise deve ser feita através da observação dos mapas de concentração construídos a partir das matrizes dos valores preditos. A obtenção desse tipo de mapa permite obter informações de concentração a nível local e fornece informações para estudos de homogeneidade e pureza, semelhante aos mapas de concentração resultantes da MCR.

A regressão pelo PLS tem como uma de suas vantagens o fato de poder ser usada quando se tem um conhecimento parcial dos dados (não se conhece todos os componentes da amostra). Os modelos são bastante robustos desde que as amostras utilizadas na calibração sejam representativas daquelas que serão preditas. Caso as amostras de previsão sejam diferentes, os modelos podem falhar.

Ao contrário da PCA, que modela apenas a matriz dos dados (matriz **X**), a regressão pelo PLS (Equações 6 e 7) modela as matrizes dos dados e das concentrações (matriz **c**), de modo a maximizar a correlação entre elas. O vetor **q** da Equação 7 é semelhante a um vetor de pesos, a diferença é que ele não é normalizado (BRERETON, 2000). As variáveis da regressão pelo PLS são chamadas de variáveis latentes (VL) e não PCs.

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P} + \mathbf{E} \quad (6)$$

$$\mathbf{c} = \mathbf{T} \cdot \mathbf{q} + \mathbf{f} \quad (7)$$

Quando se deseja modelar mais de uma propriedade, é possível modelá-las uma de cada vez (PLS1) ou todas simultaneamente (PLS2). No caso da modelagem simultânea tem-se que os vetores de concentração e pesos da equação 7 são substituídos por matrizes.

Existem várias maneiras de se avaliar a qualidade de um modelo de regressão pelo PLS, uma delas é através da raiz quadrada do erro médio quadrático de calibração (RMSEC, do inglês *root mean square error of calibration*), que é uma medida do erro de autoprevisão do modelo (BRERETON, 2000). Nesse caso, as amostras usadas na modelagem têm suas propriedades preditas usando o modelo e seus valores (c_i) comparados com os obtidos na etapa de previsão ($\hat{c}_i$). O cálculo é feito utilizando a Equação 8, na qual N_{cal} é o número de amostras de calibração e n o número de variáveis latentes utilizadas pelo modelo. Quanto mais variáveis, menor o valor do RMSEC e melhor o ajuste do modelo aos dados. Assim, o RMSEC pode ser usado para determinar o número ideal de variáveis latentes.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^I \sum (c_i - \hat{c}_i)^2}{N_{cal} - n - 1}} \quad (8)$$

Outras maneiras de validar um modelo são através da validação cruzada (CV, do inglês *cross validation*) ou de um conjunto de teste. Na validação cruzada, amostras são sucessivamente deixadas de fora do conjunto de calibração para serem preditas e, em seguida, incluídas no conjunto de calibração, do qual novas amostras de previsão serão retiradas. O processo é repetido até que todas as amostras tenham sido usadas na previsão. Uma das formas de realizar a validação cruzada é tirando uma amostra de cada vez (validação *leave-one-out*). A raiz quadrada do erro médio quadrático de validação cruzada (RMSECV, do inglês *root mean square error of cross-validation*), calculado utilizando a Equação 9, não diminui com o aumento do número de variáveis latentes, mas atinge um mínimo quando o número correto de variáveis é encontrado. Na validação com um conjunto de teste, amostras diferentes daquelas utilizadas na calibração são previstas e o erro é calculado de maneira semelhante ao RMSECV, a diferença é que no lugar do número de amostras de calibração (N_{cal}) usa-se o número de amostras de teste (BRERETON, 2000).

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^I \sum (c_i - \hat{c}_{i(CV)})}{N_{cal}}} \quad (9)$$

Após a construção e validação do modelo tem-se a etapa de previsão, na qual as amostras têm suas propriedades preditas através do modelo. Quando os valores das propriedades são conhecidos, é possível calcular a raiz quadrada do erro médio quadrático de previsão (RMSEP, do inglês *root mean square error of prediction*) a partir da Equação 10, na qual N_{pred} é o número de amostras de previsão.

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^I \sum (c_i - \hat{c}_{i(pred)})}{N_{pred}}} \quad (10)$$

Além do RMSEP, é possível avaliar os erros de previsão calculando-se o erro relativo médio, média dos erros percentuais das concentrações, os quais podem ser calculados a partir da Equação 11, na qual c_i é o valor da concentração definido no preparo da amostra (valor de referência) e $\hat{c}_i$ o valor predito pelo modelo. Essa alternativa é importante quando se tem amostras com valores de concentrações muito distintos, pois um mesmo valor de erro absoluto pode corresponder, em termos percentuais, a um valor muito elevado para uma amostra de menor concentração.

$$\text{Erro percentual} = \frac{c_i - \hat{c}_i}{c_i} 100\% \quad (11)$$

Na regressão pelo PLS, a modelagem pode ser feita utilizando o espectro completo ou apenas as variáveis mais significativas, selecionadas, por exemplo, com o algoritmo *jack-knife*. A seleção das variáveis pelo *jack-knife* é feita construindo-se um modelo PLS com validação cruzada, usando todas as variáveis, e calculando-se os coeficientes de regressão das variáveis e seus intervalos de confiança (MARTENS; MARTENS, 2000; KARAMAN et al., 2013). As variáveis cujos coeficientes de regressão forem significativos (intervalo de confiança não passar pelo zero) são selecionadas. As demais variáveis, cujos coeficientes não forem significativos, passam a ter coeficiente igual a zero. O modelo construído apenas com as variáveis selecionadas é mais simples (utiliza um número menor de variáveis latentes) ou apresenta valores de erros menores do que os dos modelos iniciais.

3.5.5. Análise discriminante linear

A LDA é um método de classificação que define limites lineares entre as classes e se baseia nas medidas da distância de Mahalanobis das amostras para o centro de cada classe (BRERETON, 2009). A distância de Mahalanobis difere da Euclidiana por levar em conta as variâncias e covariâncias das variáveis (matriz de covariância). O quadrado da distância de Mahalanobis de uma amostra ($\mathbf{x}$) para o centro de sua classe (μ_j) é dada pela Equação 12, na qual Σ_j é a matriz de covariância da classe j (MOREIRA et al., 2009).

$$r^2(\mathbf{x}, \mu_j) = (\mathbf{x} - \mu_j)^T \Sigma_j^{-1} (\mathbf{x} - \mu_j) \quad (12)$$

Uma amostra desconhecida é classificada na classe cujo centro estiver mais próximo, assim, essa amostra só pode ser classificada em uma classe (MOREIRA et al., 2009).

Na LDA, o número de variáveis utilizadas deve ser menor do que o número de amostras de treinamento. Caso o número de variáveis seja maior, a matriz de covariância será singular (não invertível), impedindo o cálculo da distância de Mahalanobis (MOREIRA et al., 2009). Assim, a utilização da LDA para dados espectroscópicos, cujo número de variáveis é elevado, necessita de uma etapa prévia de seleção de variáveis.

A seleção de variáveis pode ser feita manualmente, quando se tem conhecimento das variáveis mais importantes, ou através do uso de algoritmos. Dois dos algoritmos que vem sendo utilizados com sucesso em problemas de classificação são: o algoritmo das projeções sucessivas (SPA, do inglês *successive projections algorithm*) e o algoritmo genético (GA, do inglês *genetic algorithm*).

A seleção de variáveis utilizando o SPA é guiada por uma função de custo G que calcula o risco de uma amostra ser classificada na classe errada quando uma cadeia de variáveis é selecionada. Essa função é calculada através da Equação 13, na qual K_v é o número de amostras de validação e g_k é o risco de classificações incorretas da k -ésima amostra de validação. A função g_k , calculada através da Equação 14, utiliza a distância de Mahalanobis para calcular as distâncias da amostra ($\mathbf{x}_k$) de índice de classe l_k para o centro de sua classe (numerador) e para o centro da

classe errada mais próxima (denominador), de índice de classe l_j (PONTES et al., 2005).

$$G = \frac{1}{K_v} \sum_{k=1}^{K_v} g_k \quad (13)$$

$$g_k = \frac{r^2(x_k, \mu_{l_k})}{\min_{l_j \neq l_k} r^2(x_k, \mu_{l_j})} \quad (14)$$

A cadeia de variáveis selecionada é aquela que apresenta menor valor da função de custo. Para que o valor dessa função seja mínimo é necessário minimizar a distância da amostra em relação ao centro da sua classe e maximizar a sua distância em relação ao centro da classe incorreta mais próxima. É importante ressaltar que os números mínimo e máximo de amostras que podem ser selecionadas para compor as cadeias de variáveis são definidos pelo operador.

A seleção de variáveis utilizando o GA é guiada por uma função de ajuste que é inversa à função de custo usada no SPA (PONTES et al., 2005). Esse algoritmo foi criado para simular o processo de evolução das espécies e, por isso, se baseia em cromossomos, gerações e evoluções baseadas em mutações e cruzamentos (CHIANG; PELL, 2004).

O algoritmo começa criando aleatoriamente n_p cromossomos, para os quais são selecionadas diferentes cadeias de variáveis, e calculando suas funções de ajuste. Como as cadeias de variáveis são criadas aleatoriamente nessa primeira etapa, as funções de ajuste tendem a ser pequenas. Ao longo da criação novas gerações de cromossomos (evolução), mutações e cruzamentos são realizadas para aumentar as funções de ajuste. Ao final de n_e evoluções, o cromossomo com maior função de ajuste é selecionado e o procedimento repetido n_r vezes, resultando na seleção de n_r cromossomos (CHIANG; PELL, 2004). A frequência de seleção das variáveis nos n_r cromossomos é, então, calculada e usada para selecionar as variáveis que melhor distinguem as classes (CHIANG; PELL, 2004).

O GA é um algoritmo estocástico, assim, as populações iniciais e, conseqüentemente, os resultados são diferentes a cada repetição. Por isso, a definição da melhor cadeia de variáveis é comumente feita a partir de várias repetições, selecionando-se aquela cujos resultados de classificação sejam melhores.

Tanto no SPA quanto no GA, a seleção da cadeia de variáveis é feita com base nas amostras de validação. Assim, tanto as amostras de treinamento quanto de validação são usadas na construção do modelo. As amostras de previsão, por sua vez, são usadas após a construção dos modelos a fim de comparar suas capacidades preditivas.

A divisão das amostras nos conjuntos de treinamento, validação e previsão pode ser feita manualmente ou utilizando-se um algoritmo de seleção de amostras como o Kennard-Stone (KS). O KS seleciona amostras representativas para compor o conjunto de treinamento e as amostras restantes são divididas entre os conjuntos de validação e previsão (PONTES et al., 2005). É importante ressaltar que a seleção é feita para cada classe individualmente e que o número de amostras em cada conjunto é definido pelo operador.

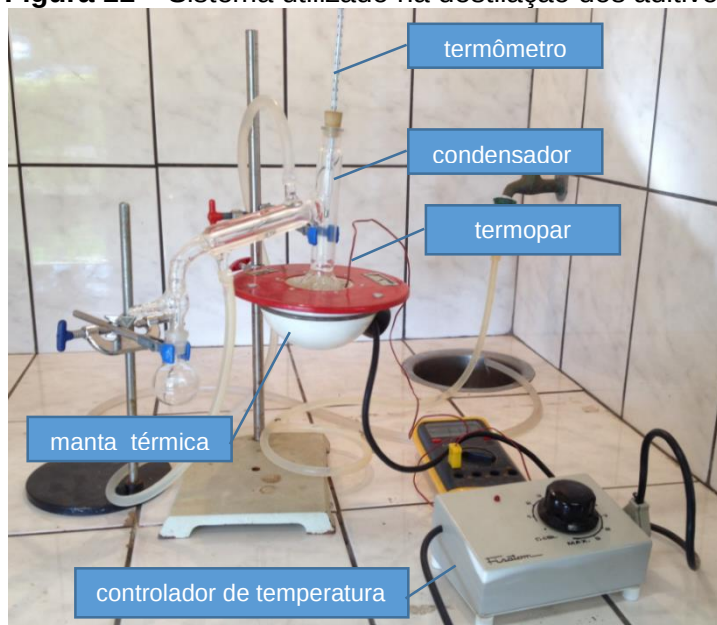
4. MATERIAIS E MÉTODOS

4.1. Aditivos

Dentre os cinco aditivos detergentes e dispersantes utilizados no Brasil, quatro foram fornecidos pela ANP para a realização do estudo. De forma a não identificar a composição dos mesmos ou dar indícios das empresas que os utilizam, a ANP os denominou de G, T, W e Y.

A fim de garantir a sua não-volatilidade nas condições de análise e avaliar sua solubilidade em diferentes solventes, os aditivos G, T e W (únicos disponíveis no momento da análise) foram submetidos a alguns testes preliminares. No que diz respeito à sua volatilidade, eles foram colocados em tubos de ensaio e submetidos a aquecimento em banho de areia (controlado com o auxílio de um termômetro) a fim de avaliar, com base em medidas em uma balança analítica, a ocorrência de perdas de massa. Outro teste realizado, também no sentido de analisar a volatilidade dos aditivos, foi a sua destilação à pressão atmosférica. Para a destilação, o aquecimento foi realizado utilizando-se uma manta térmica e o controle da temperatura foi feito utilizando-se um termopar, que mede a diferença de potencial (proporcional à temperatura). O sistema utilizado na destilação está apresentado na Figura 11.

Figura 11 – Sistema utilizado na destilação dos aditivos.



Fonte: da autora

No que diz respeito à sua solubilidade, os aditivos foram testados a fim de decidir o melhor solvente para o preparo das amostras. Assim, eles foram adicionados a tolueno, tetracloreto de carbono, álcool isopropílico, tetrahidrofurano, acetato de etila, 1,4-dioxano, metanol e éter etílico anidro e a sua solubilidade foi avaliada com base no aspecto visual das misturas (transparentes ou turvas).

4.2. Preparo das amostras

Uma vez que os aditivos são viscosos e é difícil transferir pequenas quantidades dos mesmos, necessárias para o preparo de soluções de baixas concentrações, optou-se por preparar uma solução estoque de aproximadamente 100.000 ppm (mg/Kg) em tolueno para cada um dos quatro aditivos. Para o preparo das soluções estoque, as massas dos aditivos e da solução final (aditivo + tolueno) foram medidas em uma balança analítica e os valores das concentrações calculados.

A partir de diluições dessas soluções estoques em gasolinas comuns (gasolinas base), foram preparadas soluções dos aditivos em concentrações variando de cerca de 50 a 1.000 ppm. As concentrações das soluções finais foram calculadas a partir das medidas das massas das soluções estoque adicionadas e da mistura final (solução estoque + gasolina). É importante observar que as gasolinas comuns utilizadas nesse trabalho foram todas obtidas na Central Analítica da UNICAMP, sendo, portanto, todas originárias do Estado de São Paulo.

Na primeira etapa do trabalho, na qual as amostras foram preparadas a partir de uma mistura (*blend*) de 12 gasolinas comuns diferentes (cerca de 1 litro de cada), foram preparadas 8 concentrações diferentes de cada aditivo. Na segunda etapa, na qual as amostras foram preparadas a partir de gasolinas comuns de distribuidores diferentes, foram preparadas outras 9 concentrações dos aditivos G, T e W. É importante ressaltar que, no caso do aditivo Y, foram preparadas 3 concentrações em uma mesma gasolina comum, outras 3 concentrações em uma gasolina comum diferente da primeira e, finalmente, cada uma das 3 concentrações restantes em três gasolinas comuns diferentes entre si e, também, das duas primeiras. Foram preparadas, então, 17 soluções para cada aditivo, totalizando 68 soluções.

Na Tabela 2 estão apresentadas as concentrações das amostras que foram preparadas para cada aditivo, utilizando o *blend* e as gasolinas de diferentes distribuidores. É importante observar que apesar dos valores das concentrações da

Tabela serem aproximados, os valores utilizados na construção dos modelos foram as concentrações corrigidas, as quais foram calculadas a partir das massas das solução estoque e final.

Tabela 2 – Concentrações das amostras preparadas utilizando o *blend* de gasolinas e as gasolinas de diferentes distribuidores para cada aditivo.

Aditivo	Gasolina base	Concentrações (mg.kg ⁻¹)										
		50	100	150	200	300	400	500	600	750	800	1.000
G	<i>Blend</i> de gasolinas	X	X	X	X	X	-	X	-	-	X	X
	Gasolinas diferentes	X	X	-	X	X	X	X	X	X	-	X
T	<i>Blend</i> de gasolinas	X	X	X	X	X	-	X	-	-	X	X
	Gasolinas diferentes	X	X	-	X	X	X	X	X	-	X	X
W	<i>Blend</i> de gasolinas	X	X	X	X	X	-	X	-	X	-	X
	Gasolinas diferentes	X	X	X	X	-	X	X	X	X	-	X
Y	<i>Blend</i> de gasolinas	X	X	X	X	X	-	X	-	X	-	X
	Gasolinas diferentes	X	X	X	X	X	X	X	X	-	X	-

4.3. Preparo dos anéis

Para o preparo dos anéis, utilizou-se um sistema construído pelo Grupo de Instrumentação e Automação em Química Analítica (GIA) da Universidade Estadual de Campinas (UNICAMP), apresentado nas Figuras 12 e 13. Esse sistema foi construído baseado no sistema proposto por Weisz, tendo-se introduzido nele uma bomba de seringa (Microlab® PSD/4, Hamilton) controlada por computador para a propulsão dos fluídos.

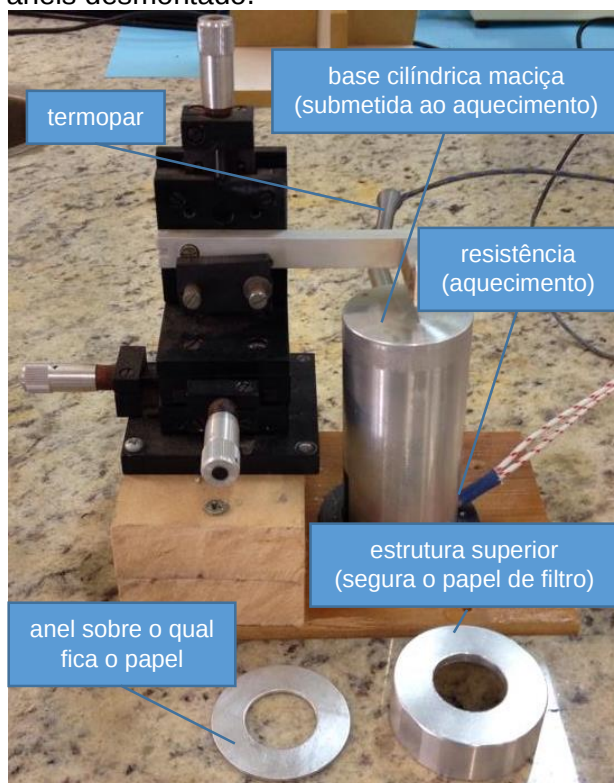
Como pode ser visto na Figura 13, a estrutura de alumínio que é submetida ao aquecimento é composta por três partes: uma base cilíndrica maciça; um anel, sobre o qual se coloca o papel de filtro (Whatman 42); e uma tampa anelar, a qual fixa o anel de alumínio e o papel, centralizando o conjunto à base cilíndrica de alumínio. É importante observar que o aquecimento é aplicado na base cilíndrica, através de uma resistência de 250 W (cartucho, Engethorm) colocada no interior dessa base, e se propaga para toda a estrutura. Um controlador de temperatura programável (TCM 45, Contemp), com auxílio de um termopar (tipo K, minipa), ajusta a temperatura do forno com resolução de 1 °C.

Figura 12 – Sistema utilizado para o preparo dos anéis.



Fonte: da autora

Figura 13 – Forno utilizado para o preparo dos anéis desmontado.



Fonte: da autora

É importante ressaltar que foram realizados estudos prévios no sentido de definir as melhores condições para o preparo dos anéis. Nesses experimentos, variou-se a temperatura de aquecimento, os volumes total e da alíquota da amostra adicionada ao papel e o intervalo de tempo entre as adições. Ao final da adição da

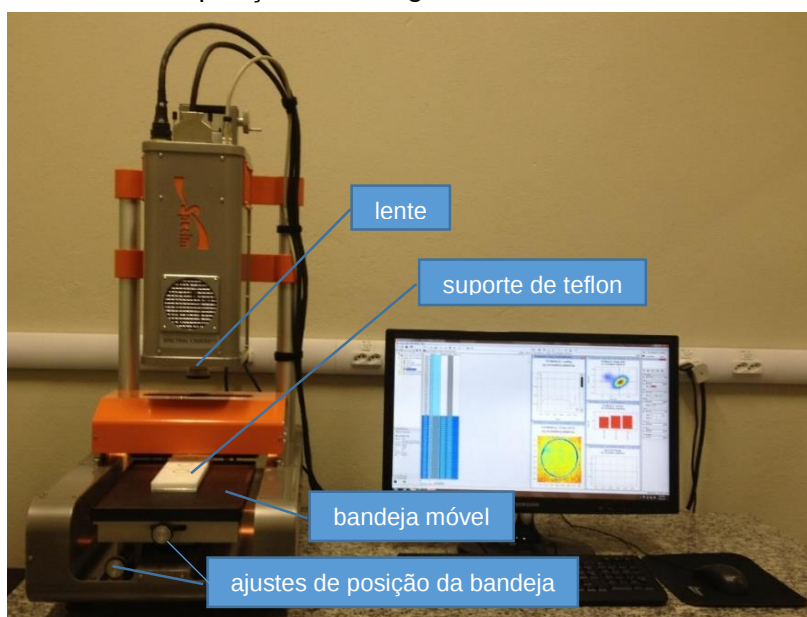
amostra foi introduzido tolueno como solvente de lavagem. O volume total e da alíquota adicionada de tolueno sobre o papel, bem como o intervalo de tempo entre as adições, também foram avaliados. Após a adição do tolueno, os anéis foram deixados por mais cerca de 5 minutos no sistema, sob aquecimento.

Para cada amostra, foram produzidos anéis em triplicata. Assim, foram produzidos 51 anéis de cada aditivo em concentrações variando de cerca de 50 a 1000 ppm (24 utilizando como gasolina base o *blend* de gasolinas comuns e 27 utilizando as gasolinas de diferentes distribuidores), totalizando 204 anéis; 39 anéis de gasolinas comuns, incluindo o *blend* de gasolinas; e 12 anéis de soluções de cerca de 1.000 ppm dos aditivos em tolueno (3 anéis de cada aditivo), totalizando 255 anéis.

4.4. Aquisição das imagens

Para a aquisição das imagens hiperespectrais no NIR utilizou-se a câmera hiperespectral SisuCHEMA SWIR da Specim Spectral Imaging Ltd. (Figura 14) com a lente de captura de 50 mm, cujos *pixels* tem dimensões de 150 x 150 μm . O intervalo espectral foi de 928 a 2.524 nm e a resolução espectral de 6,3 nm, totalizando 256 comprimentos de onda. As imagens foram adquiridas prendendo os papéis de filtro que continham os anéis em uma estrutura de teflon, de forma a mantê-los planos.

Figura 14 – Câmera hiperespectral SisuCHEMA SWIR utilizada na aquisição das imagens.



Fonte: da autora

4.5. Processamento das imagens e construção dos modelos

Após a aquisição das imagens, foi feito um corte inicial utilizando o programa computacional Evince 2.7.0 (UmBio), o que resultou em um cubo de dados com dimensões 200 *pixels* x 200 *pixels* x 256 canais espectrais. Posteriormente, os espectros foram pré-processados empregando SNV e a seleção dos pixels dos anéis (ROI) foi realizada empregando os mapas de concentração da MCR-ALS e as imagens dos escores e histogramas da PCA.

No caso da seleção da ROI a partir da MCR-ALS, a qual foi empregada apenas para as amostras preparadas a partir do *blend* de gasolinas comuns, utilizou-se como estimativas dos espectros puros dos aditivos as médias dos espectros extraídos dos anéis preparados a partir das soluções desses em tolueno. No que diz respeito aos espectros puros do papel e da gasolina comum, foram utilizados espectros médios de uma imagem apenas do papel e a média dos espectros extraídos dos anéis preparados a partir do *blend* de gasolinas comuns. Essa extração dos espectros foi feita manualmente no programa computacional Evince 2.7.0 (UmBio) a partir das imagens dos escores da PCA.

Após a extração dos espectros dos anéis, dois tipos de espectros representativos foram utilizados na construção dos modelos de regressão pelo PLS: espectros médios e espectros somados e subtraídos dos espectros do papel. Esses dois tipos de espectros representativos foram avaliados no que diz respeito aos valores dos erros (RMSEC e RMSECV) e R^2 das curvas construídas utilizando as amostras preparadas a partir do *blend* de gasolinas.

Os espectros representativos das imagens foram pré-processados empregando a suavização ou derivada com filtro de Savitzky-Golay, ambas com polinômios de 2° grau e diferentes tamanhos de janela (3, 7 ou 11 pontos). No caso dos modelos de classificação, a normalização pela faixa também foi empregada.

Na construção dos modelos PLS consolidados, foram utilizadas as médias das replicatas (excluindo-se os *outliers*). Assim, tendo-se um total de 54 anéis, esses modelos foram construídos com 12 ou 11 amostras (dependendo da exclusão ou não de amostras anômalas) e 5 amostras foram utilizadas na previsão. Das amostras de previsão, cujas concentrações variaram de 100 a 600 ppm, uma foi preparada a partir do *blend* de gasolinas comuns e quatro a partir de gasolinas comuns diferentes. Após

a construção dos modelos e definição dos melhores resultados, foram calculados os limites de detecção (LD) e de quantificação (LQ).

Na construção dos modelos de classificação, as replicatas foram consideradas como amostras distintas, contudo algumas foram excluídas por terem espectros representativos muito distintos dos espectros das outras duas replicatas. Para a construção desses modelos, utilizou-se a LDA juntamente com os algoritmos de seleção de variáveis SPA e GA. No GA a seleção foi feita utilizando-se 100 gerações de 200 cromossomos e porcentagens de mutação e cruzamento de 10% e 60%, respectivamente. Por se tratar de um algoritmo estocástico, ele foi repetido 5 vezes e o melhor resultado, em termos de erros de classificação, selecionado. A separação das amostras nos conjuntos de treinamento, validação e predição foi feita utilizando-se o algoritmo Kennard-Stone (KS) e atribuindo-se 60% das amostras ao conjunto de treinamento e 20% aos conjuntos de validação e predição (Tabela 3).

Tabela 3 – Quantidade de amostras nos conjuntos de treinamento, validação e predição.

Aditivo	Conjunto		
	Treinamento	Validação	Predição
G	28	10	10
T	29	10	10
W	27	9	9
Y	30	10	10
Total	114	39	39

Os programas computacionais utilizados no processamento dos dados foram o Evince 2.7.0 (UmBio), para a conversão do formato das imagens, corte inicial e extração dos espectros puros; MATLAB R2010a 7.10.0.499 (MathWorks), para o pré-processamento das imagens, seleção e extração dos espectros da ROI, construção dos modelos de classificação e cálculo dos limites de detecção e quantificação; e The Unscrambler X.1 (CAMO), para a construção dos modelos PLS e previsão das amostras de validação externa (amostras de previsão). Para o cálculo dos limites de detecção e quantificação, foi empregado o *toolbox* MVC1 (OLIVIERI; GOICOECHEA; IÑÓN, 2004).

5. RESULTADOS E DISCUSSÃO

Esse trabalho foi dividido em duas etapas, na primeira dispunha-se apenas de anéis produzidos a partir do *blend* de gasolinas comuns adicionado de diferentes concentrações dos aditivos, enquanto que na segunda dispunha-se de anéis produzidos a partir de diferentes gasolinas comuns (10 distribuidores do Estado de São Paulo). É importante observar que para o aditivo Y só se dispunha, na segunda etapa, de anéis produzidos a partir de gasolinas de 5 distribuidores diferentes porque foi observada a degradação da solução de 100.000 ppm do aditivo em tolueno, a qual se tornou turva. Não havendo quantidade suficiente do aditivo para preparar uma nova solução, 6 amostras que foram preparadas em duas gasolinas diferentes para a realização de testes preliminares foram utilizadas de forma a igualar a quantidade de anéis dos quatro aditivos.

Na primeira etapa do trabalho, avaliou-se a volatilidade e solubilidade dos aditivos e foram testadas diferentes condições para o preparo dos anéis e tratamento dos dados, a fim de definir a melhor estratégia de extração dos espectros representativos dos anéis. Na segunda etapa, após a definição da melhor condição para o tratamento dos dados e tendo-se anéis de diferentes gasolinas, foram testadas diferentes técnicas de pré-processamento e formas de compensar a informação espectral do papel. Além disso, foram construídos, nessa etapa, os modelos de regressão e classificação definitivos.

5.1. Resultados preliminares

Para testar a volatilidade dos aditivos, esses foram aquecidos em um banho de areia até 150 °C, temperatura máxima que o sistema do *ring-oven* poderia atingir sem danificar o papel. Até essa temperatura não foram observadas perdas de massa. Assim, garante-se que a 150 °C os aditivos estão presentes no papel.

Ainda para avaliar a volatilidade dos aditivos, eles foram submetidos à destilação atmosférica e mantidos sob aquecimento até que fosse observada a carbonização do material contido no balão. É importante observar que as temperaturas obtidas neste ensaio são valores aproximados e, por isso, as temperaturas de início da ebulição (Tabela 4) foram expressas como intervalos de temperatura. A análise dos resultados permite concluir que os aditivos são estáveis a

cerca de 250 °C, como era desejado, e se decompõem acima de 300°C, deixando um resíduo escuro e viscoso.

Tabela 4 – Avaliação da estabilidade térmica dos aditivos.

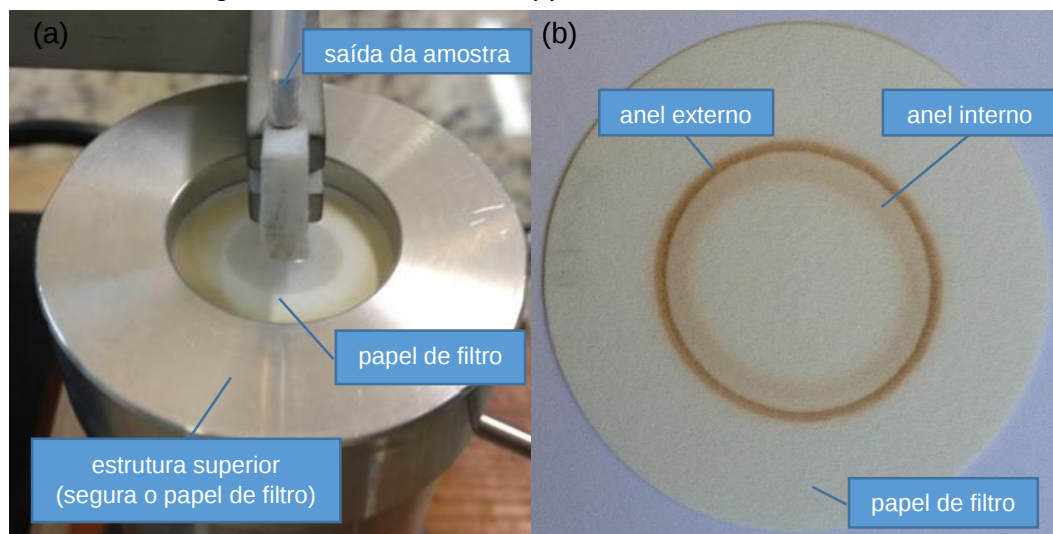
Aditivo	Temperatura (°C)	
	Início da ebulição	Início da carbonização
G	420 - 450	520
T	320 - 380	450
W	220 - 250	330

Quanto à sua solubilidade, os aditivos foram adicionados a diferentes solventes de forma a escolher aquele que melhor os dissolvesse, baseando-se no aspecto visual das misturas (transparência ou turbidez). Dentre os solventes testados, os aditivos foram solúveis em tolueno, tetracloreto de carbono, tetrahidrofurano e éter etílico. Com base nesses resultados e por se tratar de um componente normalmente encontrado nas gasolinas, optou-se por utilizar o tolueno no preparo das soluções estoque.

No que diz respeito ao preparo dos anéis, foi observada a necessidade de se adicionar, após o completo gotejamento da amostra no papel de filtro, um solvente capaz de arrastar possíveis resquícios de aditivos do centro do papel para o anel, de forma a garantir que todo o aditivo presente na gasolina esteja concentrado nele. Devido à elevada solubilidade dos aditivos em tolueno, esse solvente também foi escolhido para a lavagem. Da mesma forma que foi feito com as amostras, a adição do tolueno foi feita em gotas e com um intervalo de tempo entre as adições, de forma a garantir sua completa evaporação. Após a adição do tolueno, os anéis foram deixados por cerca de 5 minutos no sistema, sob aquecimento, de modo a garantir a evaporação de espécies voláteis ainda presentes no anel. O preparo de um anel está mostrado na Figura 15a.

Como pode ser visto na Figura 15b, há um anel externo mais escuro, cuja espessura varia entre 1 e 2 mm, e um anel interno mais claro, cuja espessura é de aproximadamente 2 mm. No que diz respeito ao diâmetro do anel externo, ele varia entre 2,5 e 2,7 cm. Estudos prévios permitiram concluir que os aditivos estão concentrados no anel mais externo, contudo não se conhece a identidade dos outros compostos também concentrados nesse anel ou no anel interno.

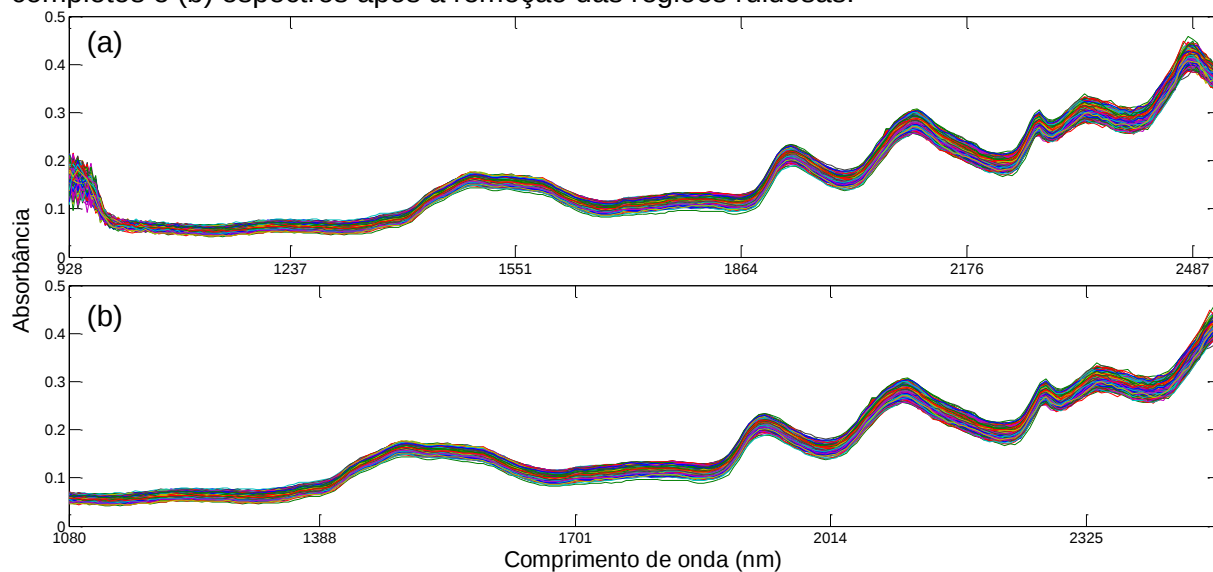
Figura 15 – (a) Produção do anel e (b) papel de filtro com anel produzido a partir de um *blend* de gasolinas contendo 100 ppm do aditivo G.



Fonte: da autora

Parte dos espectros NIR (1.000 espectros) de um anel (produzido a partir da adição de 1.800 μL de amostra) de uma amostra de gasolina contendo 50 ppm do aditivo G estão apresentados na Figura 16a. Como pode ser observado nessa Figura, tanto a região inicial (de 928 a 1.073 nm) quanto a final (de 2.493 a 2.524 nm) dos espectros são bastante ruidosas. Por isso, essas regiões foram eliminadas e restou a região compreendida entre 1.080 e 2.487 nm, composta por 226 comprimentos de onda (Figura 16b).

Figura 16 – Espectros de uma amostra contendo 50 ppm do aditivo G: (a) espectros completos e (b) espectros após a remoção das regiões ruidosas.



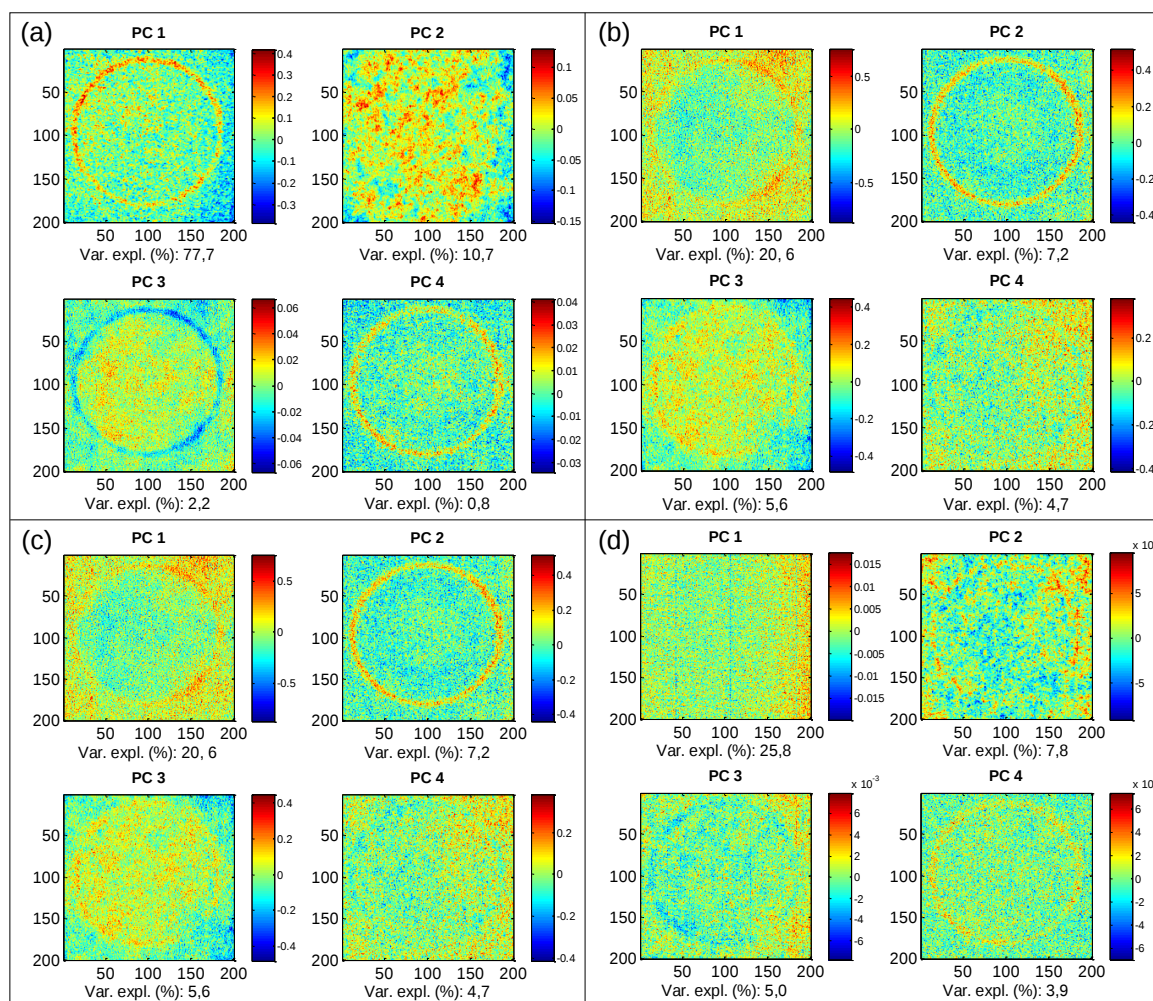
Fonte: da autora

Após a eliminação das regiões mais ruidosas, diferentes técnicas de pré-processamento foram testadas (SNV, MSC e 1° derivada com filtro de Savitzky-Golay com polinômio de segunda ordem e janela de 7 pontos) de modo a identificar aquela que facilitava a visualização dos *pixels* dos anéis nas imagens dos escores da PCA. A análise das imagens dos escores de um anel produzido a partir da adição de 1.800 µL de uma amostra contendo 50 ppm do aditivo G (Figura 17), aditivo cujos anéis se mostraram menos definidos quando comparados aos demais, indicou que o pré-processamento com derivada dificulta a visualização do anel enquanto que com SNV e MSC isso não é observado. É importante observar que as imagens dos anéis foram cortadas em cubos de dimensões espaciais de 200 x 200 *pixels* de forma a eliminar os *pixels* do fundo (placa de teflon utilizada como suporte dos anéis) e igualar o tamanho das imagens.

Optou-se por utilizar a SNV como técnica de pré-processamento pelo fato dessa técnica corrigir os espectros em relação a eles mesmos (autoescalamento nas amostras) e não em relação a um espectro de referência (ex. espectro médio), como é o caso da MSC. Isso porque ao se utilizar um espectro médio da imagem, na qual a maior parte dos *pixels* está relacionada ao papel, ele seria praticamente um espectro médio do próprio papel. Os espectros obtidos após a correção usando a SNV estão mostrados na Figura 18.

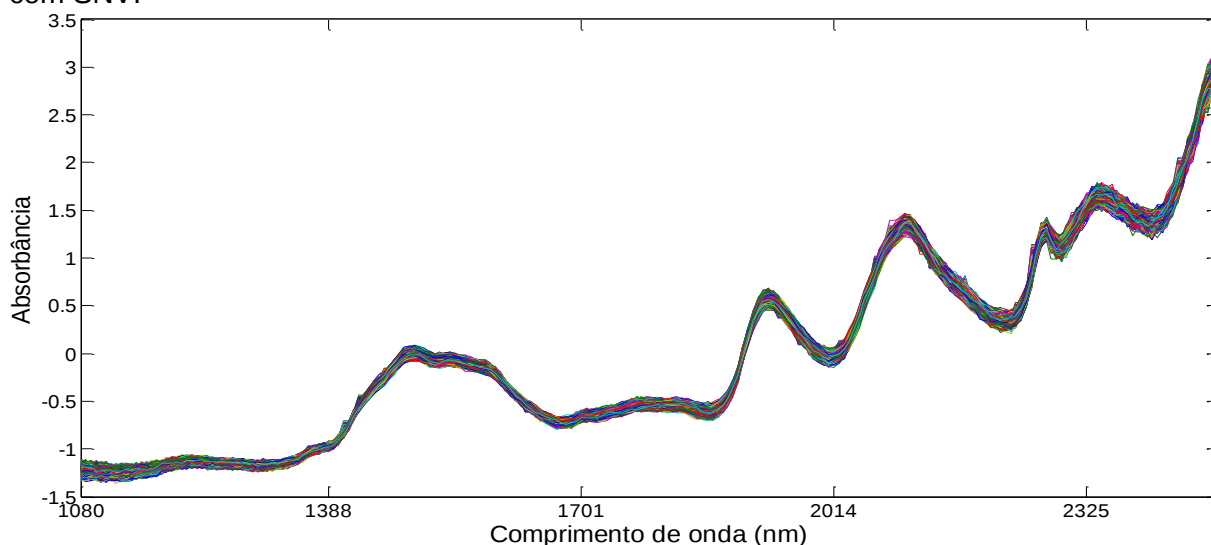
Após a escolha da SNV, diferentes condições de preparo dos anéis foram testadas. Algumas dessas condições testadas foram prontamente descartadas por serem muito demoradas ou produzirem anéis pouco definidos (borrados). Nesses casos, não foram adquiridas imagens dos anéis. Dentre as demais condições, quatro delas (Tabela 5) foram avaliadas quanto à definição dos anéis nas imagens dos escores da PCA. A temperatura de aquecimento escolhida para o preparo dos anéis foi 150 °C, por se tratar de uma temperatura que não causa danos ao papel e na qual o tolueno e os compostos mais voláteis da gasolina evaporam. É importante observar que tanto o intervalo de tempo entre as adições quanto o volume de tolueno adicionado em cada condição (volume total e das gotas) foram avaliados através da observação visual dos anéis preparados.

Figura 17 – Imagens dos escores da PCA de um anel de uma amostra contendo 50 ppm do aditivo G (a) sem pré-processamento, pré-processada com (b) SNV, (c) MSC e (d) 1° derivada com filtro de Savitzky-Golay (polinômio de 2° grau e janela de 7 pontos).



Fonte: da autora

Figura 18 – Espectros de um anel de uma amostra contendo 50 ppm do aditivo G corrigidos com SNV.



Fonte: da autora

Tabela 5 – Condições de preparo dos anéis analisadas por imagem.

Condição	Amostra			Tolueno		
	Volume (µL)		Tempo (s)	Volume (µL)		Tempo (s)
	Total	Gota		Total	Gota	
1	600	10	10	150	25	10
2	1.200	10	10	150	25	10
3	1.800	12	10	200	25	10
4	2.400	10	10	240	40	20

Como pode ser visto na Figura 19, os anéis ficaram mais visíveis a partir condição 3 (adição de 1.800 µL de amostra). Essa condição foi escolhida e, portanto, utilizada em todo o trabalho por ter um tempo de preparo do anel um pouco menor quando comparado à condição 4 (adição de 2.400 µL de amostra). Esse tempo de preparo foi de 35 minutos, incluindo os cinco minutos que o papel foi deixado sob aquecimento para garantir a evaporação dos componentes voláteis.

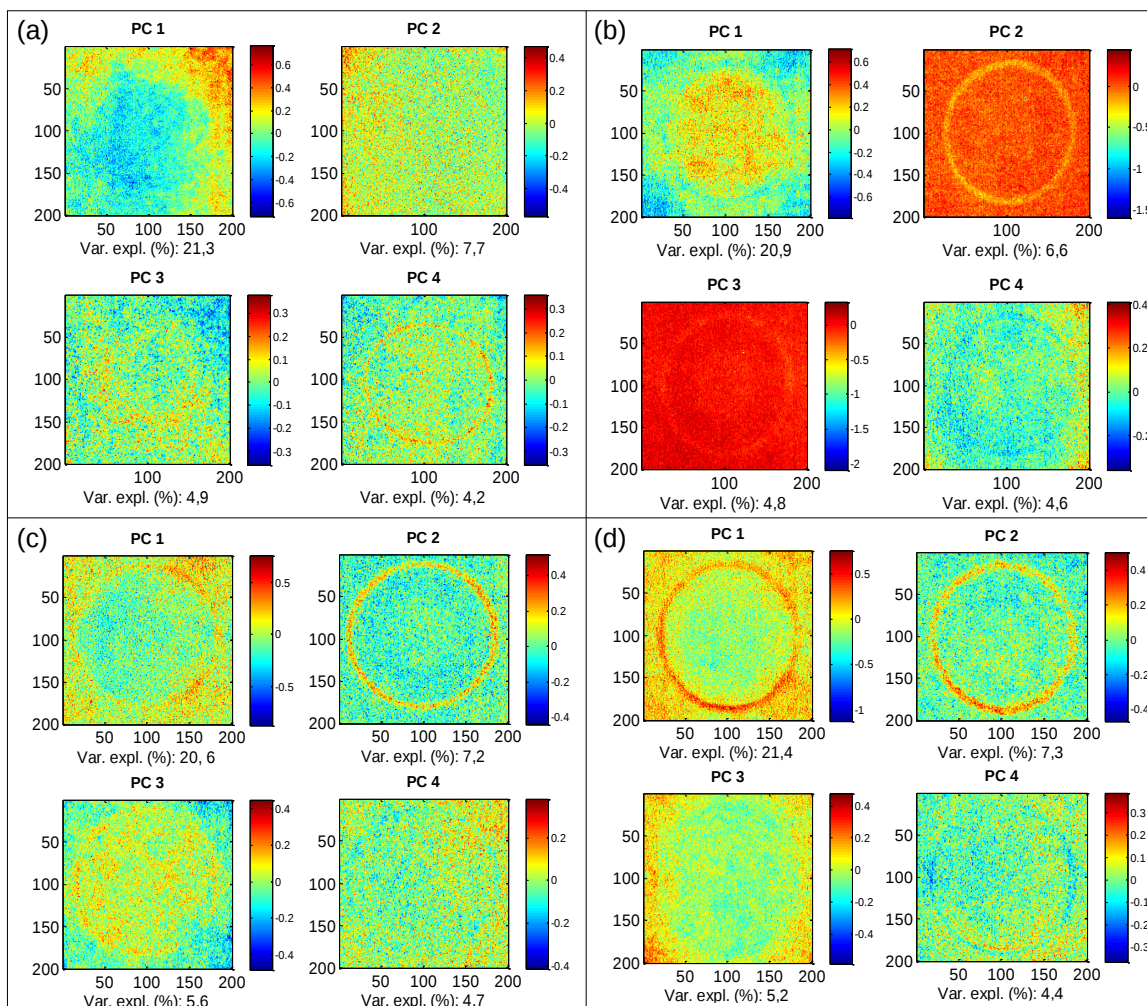
O coeficiente de pré-concentração (K), o qual permite avaliar a eficiência do processo de pre-concentração, pode ser calculado de várias formas (CORTEZ, 2012). Nesse trabalho, ele foi calculado em termos do volume adicionado da amostra ($V_{amostra}$) e do volume definido pelo anel (V_{anel}), segundo a Equação 15.

$$K(v/v) = \frac{V_{amostra}}{V_{anel}} \quad (15)$$

O volume do anel, por sua vez, foi calculado segundo a Equação 16, na qual e_{papel} é a espessura do papel (0,2 mm segundo o fabricante), r_{ext} é o raio do anel externo (1,3 cm aproximadamente) e e_{anel} é a espessura do anel (1,5 mm aproximadamente). É importante ressaltar que os valores de r_{ext} e e_{anel} foram medidos com auxílio de uma régua. Considerando esses valores médios, tem-se que o coeficiente de pré-concentração obtido para a condição 3 ($V_{amostra}=1.800 \mu\text{L}$) é de cerca de 78 vezes.

$$V_{anel} = \pi e_{papel} [r_{ext}^2 - (r_{ext} - e_{anel})^2] \quad (16)$$

Figura 19 – Imagens dos escores da PCA de um anel contendo 50 ppm do aditivo G produzido a partir da condição (a) 1 (600 μ L de amostra), (b) 2 (1200 μ L de amostra), (c) 3 (1800 μ L de amostra) e (d) 4 (2400 μ L de amostra).



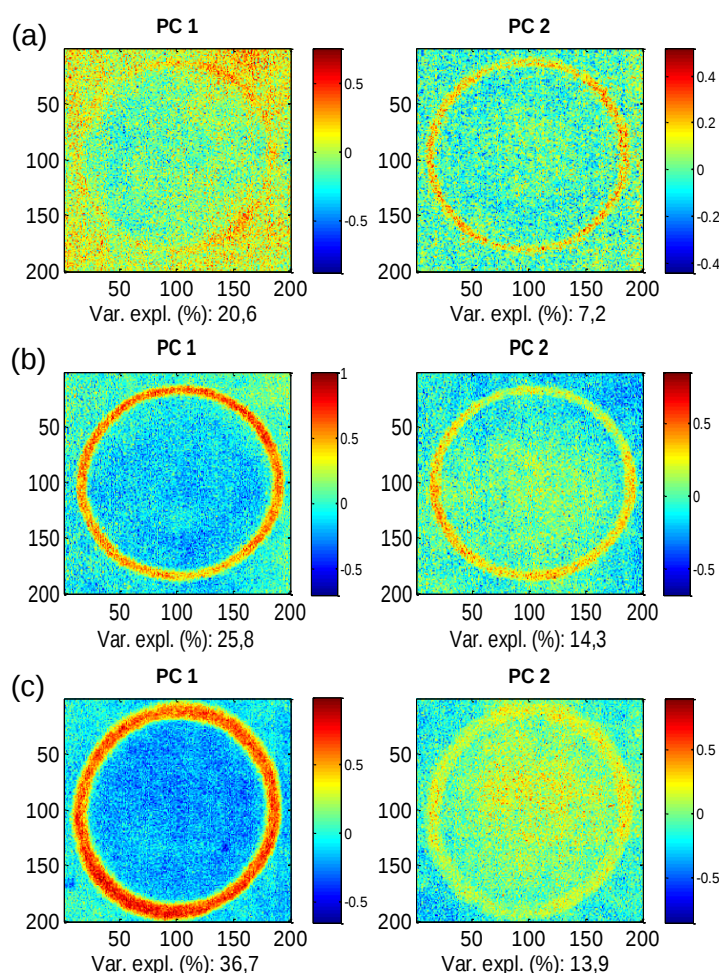
Fonte: da autora

Após a definição da melhor condição de preparo dos anéis, foram preparados 24 anéis de cada aditivo (triplicatas de 8 diferentes concentrações variando entre 50 a 1.000 ppm), totalizando 96 anéis, empregando o *blend* de gasolinas. Nessa etapa do trabalho, o volume de gasolina que se dispunha, de um único fabricante, era insuficiente para preparar todos os anéis. Assim, para evitar problemas de matriz advindos do uso de gasolinas provenientes de diferentes distribuidores, resolveu-se preparar o *blend* misturando 12 litros de gasolinas comuns diferentes.

A observação dos anéis produzidos mostrou a existência de uma tendência de alargamento com o aumento da concentração (Figura 20). Os anéis cuja concentração era de cerca de 50 ppm apresentaram uma espessura de cerca de 1 mm enquanto que alguns anéis cuja concentração era de cerca de 1.000 ppm chegaram a apresentar o dobro da espessura (2 mm). Uma possível causa desse alargamento é

a saturação do papel devido à sua pequena espessura. A quantidade de aditivo era tão elevada que o papel ficava saturado e, como consequência, o aditivo e a gasolina que gotejavam em seguida ficavam retidos antes do anel ou deslocavam os componentes já concentrados para as bordas do papel, aumentando o diâmetro do anel. É importante ressaltar que esse aumento do diâmetro também foi observado, o que corrobora essa hipótese.

Figura 20 – Imagens dos escores da PCA de anéis contendo (a) 50 ppm, (b) 500 ppm e (c) 1.000 ppm do aditivo G.



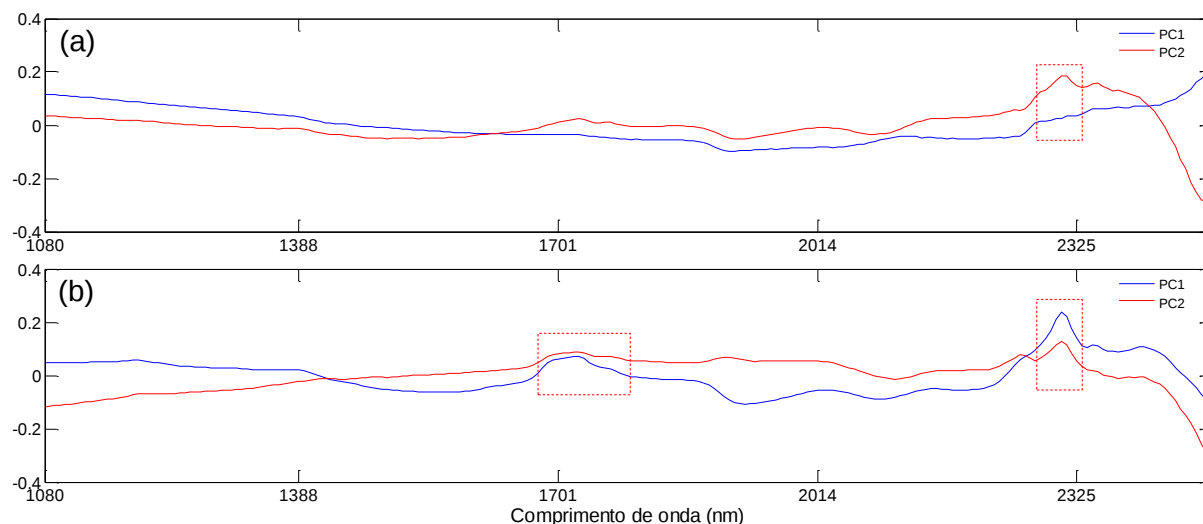
Fonte: da autora

Também é possível observar na Figura 20 que, à medida que a concentração do aditivo aumenta, o anel passa a ficar mais definido na PC1 e menos definido na PC2, enquanto que na concentração mais baixa, o anel está mais definido na PC2.

A análise dos gráficos dos pesos das duas primeiras PCs dos anéis de 50 ppm (Figura 21a) e de 1.000 ppm (Figura 21b) permitiu concluir que uma das regiões

espectrais responsável pela identificação dos *pixels* dos anéis do aditivo G, presente nas amostras analisadas, está localizada entre 2.275 e 2.338 nm. Trata-se da região das bandas de combinação de vibrações C-H, cujos valores dos pesos são positivos na PC1 e PC2 (WORKMAN JR; WEYER, 2012). Espectros cujas absorbâncias nessa região são mais elevadas, espectros de concentrações maiores, têm valores de escores mais positivos na PC1. Assim, nos anéis produzidos a partir de amostras de aditivos mais concentrados, os *pixels* do anel se diferenciam melhor dos *pixels* do papel, como pode ser visto na Figura 20c.

Figura 21 – Gráficos dos pesos da PCA de um anel contendo (a) 50 ppm e (c) 1.000 ppm do aditivo G.



Fonte: da autora

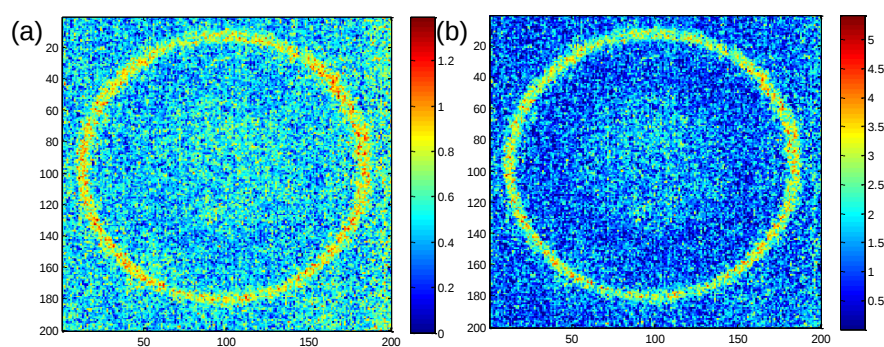
Analisando-se apenas o gráfico dos pesos da PCA do anel contendo 1.000 ppm (Figura 21b), pode-se observar que a região entre 1.670 e 1.789 nm também contribuiu para a diferenciação dos *pixels* do anel, com valores de escores positivos na PC1 e PC2. Essa região corresponde aos primeiros sobretons das vibrações de estiramento C-H (WORKMAN JR; WEYER, 2012).

À medida que as cadeias carbônicas dos compostos se tornam mais longas, há um aumento da quantidade de ligações C-H de grupos metileno (CH_2) em relação à quantidade de ligações de grupos metilas (CH_3), a qual tende a se manter constante por se tratarem de grupos terminais. Assim, compostos de longas cadeias carbônicas lineares ou ramificadas, como é o caso dos aditivos, apresentam uma quantidade maior de ligações C-H e, portanto, contribuem fortemente para o aumento da intensidade de absorção nas regiões destacadas na Figura 21. Isso permite a

identificação dos aditivos através das ligações C-H de suas cadeias carbônicas e não dos seus grupos funcionais específicos.

Foi avaliado se o agrupamento das imagens (imagens de vários anéis juntos) melhorava a definição dos *pixels* dos anéis nas imagens dos escores da PCA e nos mapas de concentração da MCR-ALS. Observou-se que, principalmente para os mapas de concentração, o agrupamento favorecia a definição dos anéis de menores concentrações, como pode ser visto na Figura 22. Isso já era esperado, pois ao se usar os mesmos espectros puros e diferentes concentrações dos componentes na MCR-ALS (forma aumentada), os resultados são melhores.

Figura 22 – Mapas de concentração de um anel contendo 50 ppm do aditivo G para a MCR-ALS (a) individual e (b) aumentada (todas as imagens de diferentes concentrações dos aditivos juntas).



Fonte: da autora

5.1.1. Extração dos espectros do anel

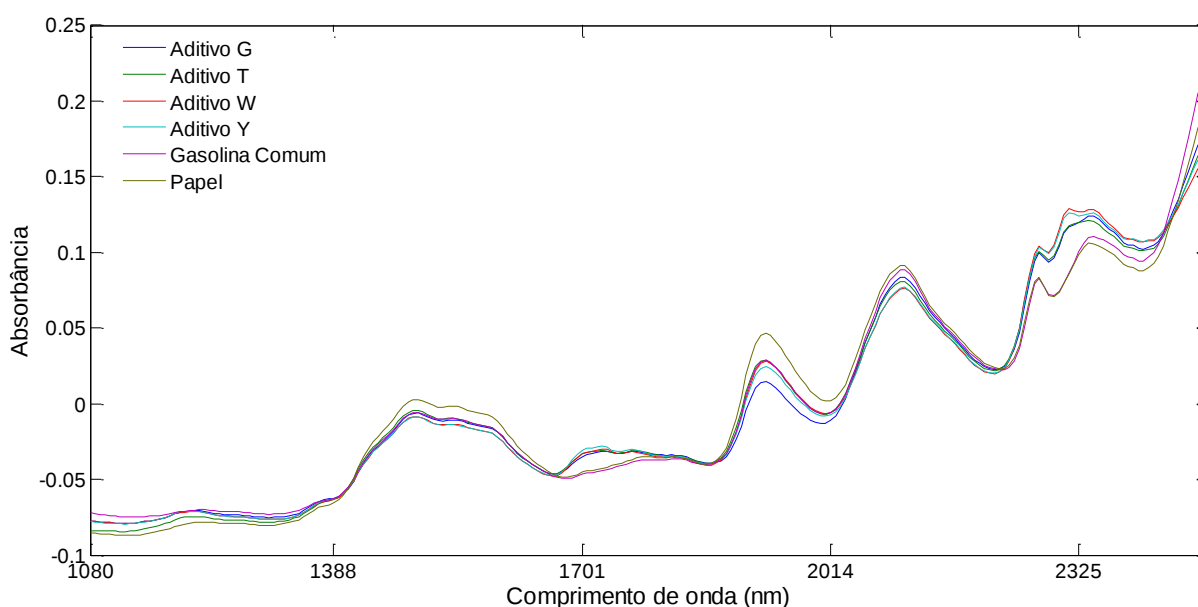
Três abordagens de extração dos espectros dos *pixels* do anel (região de interesse da imagem) foram testadas a fim de escolher aquela que fosse mais rápida e menos subjetiva, de forma que pudesse ser adaptada facilmente à rotina dos laboratórios. A primeira delas baseou-se nos mapas de concentração dos aditivos, obtidos através da MCR-ALS, enquanto que as outras duas se basearam nas imagens dos escores e histogramas da primeira componente da PCA. Nessas duas últimas abordagens, os modelos da PCA foram construídos apenas com as regiões nas quais os aditivos têm perfis espectrais diferentes daqueles observados para o papel e para a gasolina, regiões essas que foram identificadas através da análise dos espectros otimizados obtidos através da MCR-ALS.

5.1.1.1. Mapas de concentração da resolução multivariada de curvas

Tendo-se observado melhores resultados para a MCR-ALS aumentada, tanto na definição das imagens dos anéis quanto na qualidade dos espectros otimizados, optou-se por utilizar essa abordagem. Assim, as imagens dos anéis para as 8 diferentes concentrações de cada aditivo (uma replicata por vez) foram agrupadas e a resolução realizada. Uma vez que se tinha triplicatas de todos os anéis, a MCR-ALS foi aplicada três vezes, resultando em 24 mapas de concentração para cada aditivo.

É importante observar que os espectros dos aditivos e do *blend* de gasolinas comuns utilizados como estimativas iniciais na resolução (espectros médios obtidos a partir daqueles extraídos manualmente) não são realmente espectros puros, pois contêm informações espectrais do papel. A observação dos espectros otimizados pela MCR-ALS (Figura 23) permitiu identificar a existência de duas regiões espectrais nas quais os aditivos têm perfis diferentes do *blend* de gasolinas e do papel. Essas regiões características dos aditivos se localizam entre 1.670 e 1.801 nm (22 comprimentos de onda) e entre 2.238 e 2.425 nm (31 comprimentos de onda) e correspondem às regiões dos primeiros sobretons das vibrações de estiramento C-H e das bandas de combinações de vibrações C-H, respectivamente (WORKMAN JR; WEYER, 2012).

Figura 23 – Espectros otimizados pela MCR-ALS dos aditivos, do *blend* de gasolinas comuns e do papel com as regiões características dos aditivos destacadas em vermelho.



Fonte: da autora

As duas regiões espectrais (entre 1.670 e 1.789 nm e entre 2.275 e 2.338 nm) identificadas nos gráficos dos pesos das PC1 e PC2 (Figura 21) como responsáveis pela diferenciação dos *pixels* do anel fazem parte dessas regiões identificadas nos espectros otimizados. Apesar de se tratarem de regiões relacionadas às vibrações C-H, que ocorrem não só nas moléculas dos aditivos, mas também nas moléculas dos componentes do papel e das gasolinas comuns, é possível que a diferença espectral se deva ao fato das cadeias carbônicas dos aditivos terem uma quantidade maior de ligações C-H de grupos metileno.

Através da observação dos mapas de concentração dos aditivos, foram definidos, visualmente, limites para os valores dos escores de forma a identificar os *pixels* relacionados aos anéis, *pixels* cujos valores estavam acima desses limites. Após a determinação dos limites, a extração dos espectros desses *pixels* foi realizada utilizando uma rotina desenvolvida no MATLAB. Nessa rotina, informam-se os valores limites para cada imagem, os espectros dos *pixels* com valores superiores aos dos limites são extraídos e são calculados os espectros médios ou somados, os quais passam a ser os espectros representativos dos anéis. No caso dos espectros somados, a fim eliminar a informação espectral do papel e contabilizar apenas a dos aditivos e demais compostos de alto peso molecular da gasolina que não foram eliminados por evaporação, o mesmo número de espectros médios representativos do papel deve ser somado para ser, em seguida, subtraído dos espectros somados do anel.

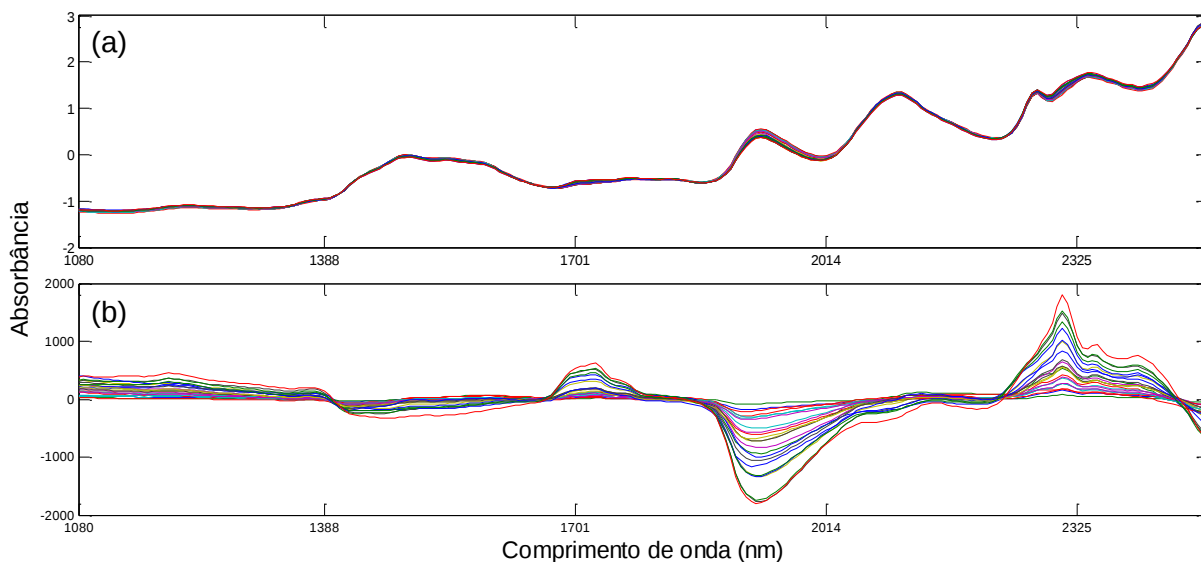
A necessidade de eliminar a informação espectral do papel surgiu da observação do aumento da espessura e diâmetro dos anéis para as maiores concentrações do aditivo. Isso porque com o aumento da concentração um número maior de *pixels* é selecionado e, conseqüentemente, um número maior de espectros é extraído, fazendo com que o somatório das absorbâncias dos espectros seja maior não só pelo aumento da concentração do aditivo, mas, principalmente, pelo aumento da quantidade de espectros do papel sendo somados.

O aumento da espessura e diâmetro dos anéis (aumento da área do anel) resulta em uma diminuição do fator de pré-concentração, o que faz com que a razão do sinal do aditivo em relação ao do papel seja menor do que seria esperado para anéis de dimensões menores, mais concentrados. Assim, ao se calcular a média dos espectros, as absorbâncias são menores do que seriam caso os anéis apresentassem as mesmas dimensões, o que pode resultar em erros na calibração. Pode-se supor,

então, que os espectros somados e subtraídos dos espectros do papel são espectros representativos melhores quando se observa a variação do fator de pré-concentração.

Os espectros médios e somados extraídos dos anéis produzidos a partir do aditivo G estão apresentados na Figura 24. Observa-se, na Figura 24b, a presença de valores de absorbância positivos e negativos. No primeiro caso, as bandas positivas aparecem nas regiões atribuídas aos aditivos, como consequência do aumento da concentração desses. No que diz respeito às bandas negativas, tem-se que elas estão localizadas nas regiões entre 1.413 e 1.657 nm e 1.864 e 2.057 nm. Essas regiões são atribuídas ao primeiro sobretom e às combinações de vibrações das moléculas de água, respectivamente. Assim, os valores negativos observados nos espectros podem significar que nas regiões do papel onde o aditivo está depositado a quantidade de água seja menor quando comparada a quantidade no resto do papel. Como o aditivo é insolúvel em água, é possível que ocorra a formação de uma região hidrofóbica, minimizando a absorção de umidade do ar pelo papel. Assim, ao subtrair os espectros do papel, contendo umidade, dos espectros do anel, cuja umidade é menor, tem-se como resultado valores negativos no espectro.

Figura 24 – Espectros (a) médios e (b) somados e subtraídos dos espectros do papel extraídos através dos mapas de concentração da MCR-ALS dos anéis produzidos a partir das diferentes concentrações do aditivo G na gasolina comum.



Fonte: da autora

Após a extração dos espectros, foram construídos modelos PLS tanto para os espectros médios quanto para os espectros somados e subtraídos do papel, de forma a avaliar os melhores resultados para cada aditivo e escolher o melhor espectro

representativo do anel (ROI). Além disso, os modelos foram construídos utilizando o espectro completo, as variáveis selecionadas pelo *jack-knife* e as regiões identificadas como sendo características dos aditivos. É importante observar que, nessa primeira etapa do trabalho, todos os modelos PLS foram construídos utilizando as triplicatas de cada anel como sendo amostras distintas, de forma a avaliar a dispersão das mesmas. Os resultados de todos os modelos construídos estão apresentados no Apêndice A.

Utilizando-se o teste F (95 % de confiança) para avaliar se os valores de RMSECV dos modelos construídos eram estatisticamente diferentes, observou-se que todos os modelos construídos para os aditivos G e W foram estatisticamente equivalentes, enquanto que os modelos dos aditivos T e Y utilizando os espectros médios foram estatisticamente diferentes daqueles utilizando os espectros somados.

Os melhores modelos foram aqueles que apresentaram os menores erros (RMSEC e RMSECV) e/ou utilizaram um número menor de variáveis latentes. No caso do aditivo G, os modelos utilizando os espectros médios apresentaram erros menores, contudo, como os valores dos erros não foram estatisticamente diferentes, o modelo escolhido foi aquele que utilizou o menor número de variáveis latentes. Quanto ao *jack-knife*, o seu emprego não fez muita diferença para a construção de alguns modelos uma vez que a maioria das variáveis era selecionada, quando isso aconteceu o modelo sem *jack-knife* foi escolhido. Os resultados dos melhores modelos para cada aditivo estão apresentados na Tabela 6.

Tabela 6 – Melhores resultados obtidos para os espectros extraídos pela MCR-ALS com base nos mapas de concentração dos aditivos.

Aditivo	Espectro representativo	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
					Calibração	Validação
G	Somado + JK	1	123	133	0,8778	0,8676
T	Somado	1	58	64	0,9755	0,9719
W	Somado	2	60	70	0,9680	0,9548
Y	Somado (regiões)	1	30	43	0,9914	0,9835

O que se pode observar nos resultados é que os melhores modelos foram aqueles que utilizaram os espectros somados, o que leva a supor que se trata realmente de uma estratégia melhor. Os erros obtidos, no entanto, foram bastante elevados para o aditivo G, provavelmente devido ao fato da seleção da ROI ser feita

de forma visual e esse aditivo não apresentar anéis bem definidos nos mapas de concentração. Essa hipótese é corroborada pelo fato do aditivo Y, o qual apresenta anéis mais definidos, ter sido aquele que apresentou menores erros. Isso permite concluir que a seleção da ROI tem grande influência nos resultados e a sua definição é um fator extremamente importante para a construção dos modelos multivariados.

Essa primeira abordagem de extração dos espectros, por se basear na definição de limites visuais, é bastante subjetiva e demanda um tempo grande para cada análise. Além disso, diferentes analistas podem definir limites diferentes e, dessa forma, extrair mais ou menos *pixels* relacionados ao anel. Como a seleção da ROI influencia diretamente os resultados, a escolha correta dos limites passa a ter uma importância fundamental no resultado analítico.

5.1.1.2. Imagens dos escores das componentes principais

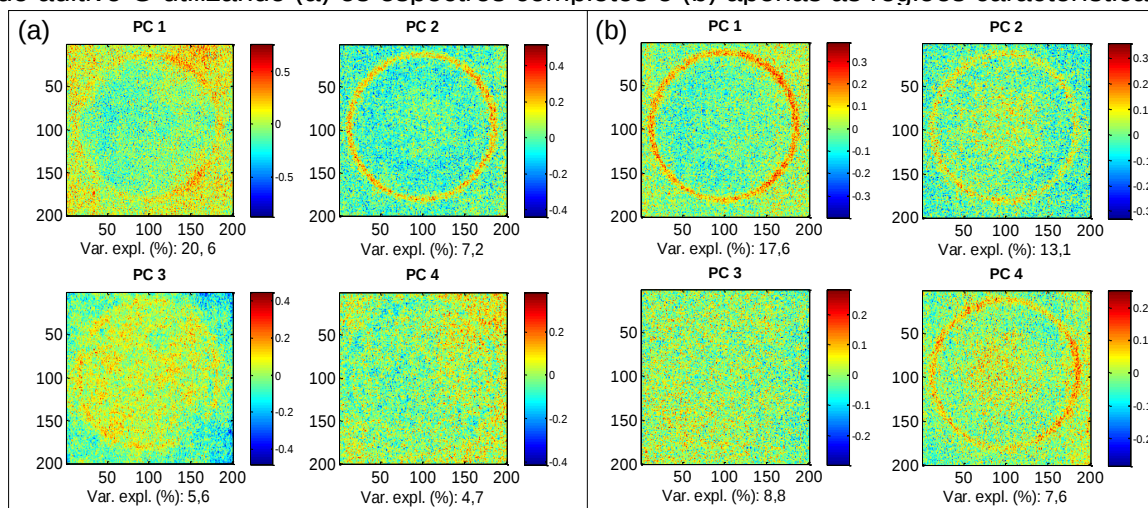
Tendo-se identificado as regiões características dos aditivos (entre 1.670 e 1.801 nm e entre 2.238 e 2.425 nm), a PCA foi construída restringindo os espectros a essas regiões. Essa restrição fez com que a informação do aditivo nos anéis de baixas concentrações passasse a ser explicada pela PC1, da mesma forma que tinha sido observado para os anéis de 500 e 1.000 ppm nas Figuras 20b e 20c. Assim, os anéis deixaram de ser identificados na PC2 (Figura 25a) para serem identificados na PC1 (Figura 25b).

Como pode ser visto na Figura 25b, mesmo restringindo-se o espectro às regiões características dos aditivos, os anéis que apresentam menores concentrações de aditivo (50 ppm) não ficam bem definidos quando comparados àqueles de maiores concentrações. Alguns *pixels* do papel se confundem com os do anel, o que dificulta a definição dos limites visuais e pode levar à incorporação de uma maior quantidade de *pixels* do papel na ROI. Por isso, buscou-se melhorar a definição dos anéis através do agrupamento das imagens, da mesma forma que foi feita na MCR-ALS.

Na construção dos modelos, foram testadas três alternativas de agrupamento (Figura 26): PCA das triplicatas de cada concentração do aditivo, PCA das concentrações divididas em dois grupos (50 a 300 ppm e 300 a 1.000 ppm) e PCA envolvendo todas as concentrações (50 a 1.000 ppm). Foi observado que o agrupamento das concentrações menores com as maiores (Figuras 26b e 26c)

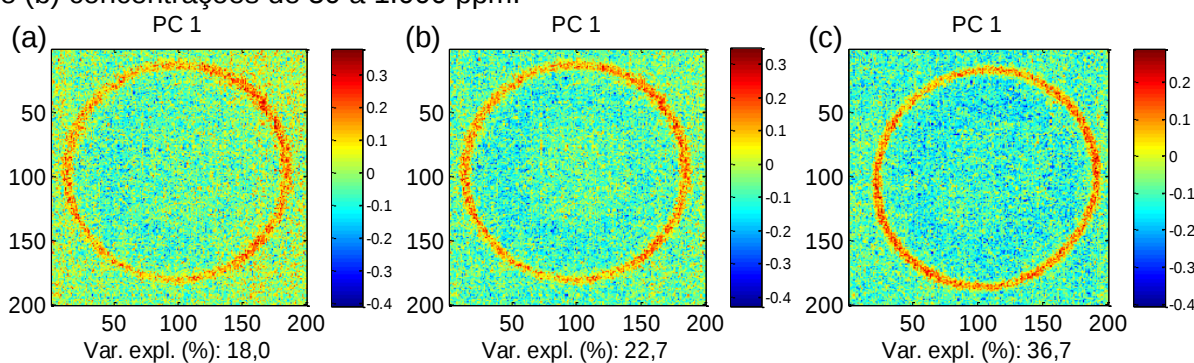
melhorava a definição do anel de menor concentração (50 ppm), uma vez que aumentava a diferença entre os valores de escores dos *pixels* do papel e do anel.

Figura 25 – Imagens dos escores da PCA de um anel de uma amostra contendo 50 ppm do aditivo G utilizando (a) os espectros completos e (b) apenas as regiões características



Fonte: da autora

Figura 26 – Imagens dos escores da primeira componente da PCA de um anel de uma amostra contendo 50 ppm do aditivo G utilizando apenas as regiões características dos aditivos e agrupando-se as imagens pelas (a) triplicatas, (b) concentrações de 50 a 300 ppm e (c) concentrações de 50 a 1.000 ppm.



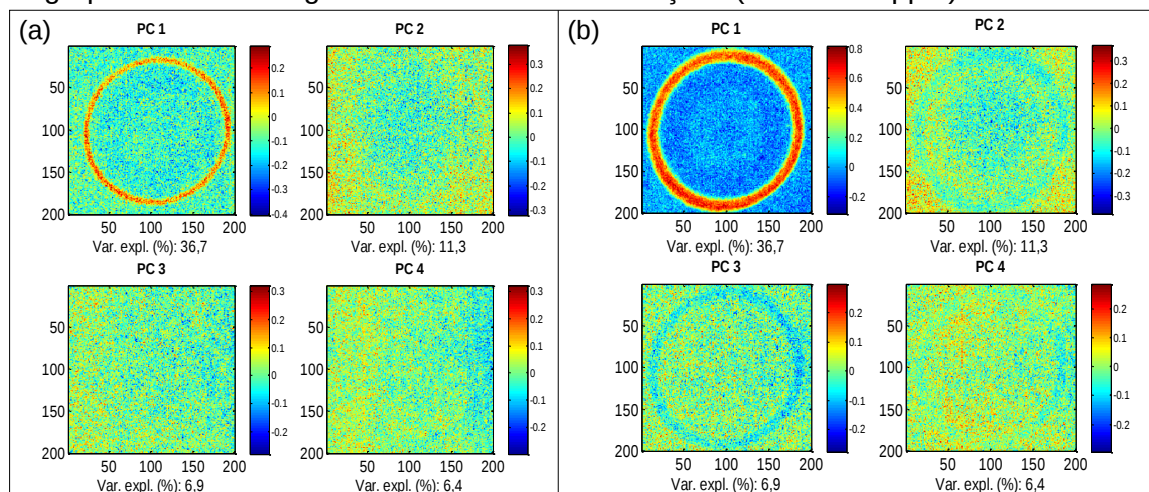
Fonte: da autora

Aparentemente, o agrupamento das imagens de concentrações de 50 a 300 ppm (Figura 26b) e de todas as concentrações juntas (Figura 26c) não apresenta diferenças quanto à definição do anel. Assim, por questão de simplicidade e rapidez, optou-se por construir as PCAs utilizando sempre todas as concentrações agrupadas, da mesma forma que foi feito na MCR-ALS.

É importante ressaltar que quando se seleciona apenas as regiões características dos aditivos e agrupam-se as imagens de todas as concentrações, é possível identificar os anéis analisando-se apenas as imagens dos escores da PC1. Isso porque no caso das concentrações mais baixas, os anéis só podem ser vistos

nas imagens dos escores da PC1 (Figura 27a) e no caso das concentrações mais altas, nas quais pode ser possível visualizá-los nas PC2, PC3 e PC4, eles estão pouco definidos e seus *pixels* se confundem bastante com os do papel, por conta dos valores dos escores serem similares (Figura 27b).

Figura 27 – Imagens dos escores da PCA de um anel de uma amostra contendo (a) 50 ppm e (b) 1.000 ppm do aditivo G utilizando apenas as regiões características dos aditivos e agrupando-se as imagens de todas as concentrações (50 a 1.000 ppm).



Fonte: da autora

Utilizando os valores dos escores da PCA, duas abordagens de extração dos espectros do anel (ROI) foram testadas, a primeira se baseou na determinação visual dos limites com base nas imagens dos escores e a segunda se baseou na determinação dos limites com base nos histogramas dos valores dos escores. É importante observar que, em ambas as abordagens, foram utilizadas apenas as imagens e valores dos escores da PC1, uma vez que essa componente é suficiente para identificar os *pixels* dos anéis.

5.1.1.2.1. Determinação visual dos limites

Essa primeira abordagem foi semelhante àquela utilizada nos mapas de concentração da MCR-ALS, a única diferença é que as imagens dos escores da PCA e não os mapas de concentração foram usados na definição dos limites. Trata-se de uma abordagem mais simples por não necessitar de estimativas iniciais dos espectros puros, os quais não estarão disponíveis na análise de amostras reais de gasolinas aditivadas.

Após a determinação dos limites, uma rotina semelhante àquela desenvolvida para os mapas de concentração foi utilizada para a extração dos espectros representativos (médios e somados). É importante ressaltar que, apesar da PCA ter sido construída utilizando-se apenas as regiões características dos aditivos, os espectros completos foram extraídos. A limitação dos espectros a essas regiões foi usada apenas para melhorar a identificação dos anéis e, conseqüentemente, a seleção dos *pixels*. Após a extração, o mesmo procedimento utilizado para os espectros extraídos pelos mapas de concentração foi adotado para a construção dos modelos empregando PLS. Os resultados de todos os modelos construídos estão apresentados no Apêndice B.

Utilizando-se o teste F (95 % de confiança) para avaliar os valores de RMSECV dos modelos construídos, observou-se, novamente, que todos os modelos construídos para os aditivos G e W foram estatisticamente equivalentes, enquanto que os modelos dos aditivos T e Y utilizando os espectros médios foram estatisticamente diferentes daqueles utilizando os espectros somados. No caso do aditivo Y, o modelo construído com os espectros somados completos foi estatisticamente diferente dos modelos construídos com as variáveis selecionadas pelo *jack-knife* e com as variáveis das regiões características dos aditivos. O modelo construído apenas com as variáveis das regiões características dos aditivos foi selecionado por apresentar menores valores de erros (RMSEC e RMSECV) e por ter usado um número menor de variáveis latentes. Os resultados dos melhores modelos para cada aditivo estão apresentados na Tabela 7.

Tabela 7 – Melhores resultados obtidos para os espectros extraídos pela PCA com base nas imagens dos escores da primeira componente.

Aditivo	Espectro representativo	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
					Calibração	Validação
G	Somado (regiões)	1	91	102	0,9322	0,9220
T	Somado (regiões)	1	46	52	0,9843	0,9819
W	Somado (regiões)	2	49	64	0,9756	0,9627
Y	Somado (regiões)	2	20	29	0,9960	0,9927

Apesar de ter havido uma redução nos valores dos erros obtidos para os aditivos G, T e W, quando comparado com os espectros extraídos pelos mapas de concentração, os valores não foram estatisticamente diferentes e os modelos também

não utilizaram um número menor de variáveis latentes. O único resultado que foi estatisticamente diferente daquele mostrado na Tabela 6 foi o do aditivo Y, cujo modelo apresentou um aumento no número de variáveis latentes (de 1 para 2) e uma redução nos valores dos erros (RMSEC de 30 para 20 e RMSECV de 43 para 29). Como a maioria dos resultados não foram estatisticamente diferentes, tem-se que essa abordagem é melhor por conta da sua simplicidade.

5.1.1.2.2. Determinação dos limites com base nos histogramas

Diante da necessidade de se ter uma abordagem menos subjetiva, que não dependa tanto da habilidade do analista, e que seja simples, buscou-se utilizar os histogramas dos valores dos escores da PCA no lugar das imagens dos escores. Dessa forma, não se depende da habilidade do analista de identificar os valores dos limites visualmente e tem-se uma forma mais objetiva de se extrair os espectros representativos.

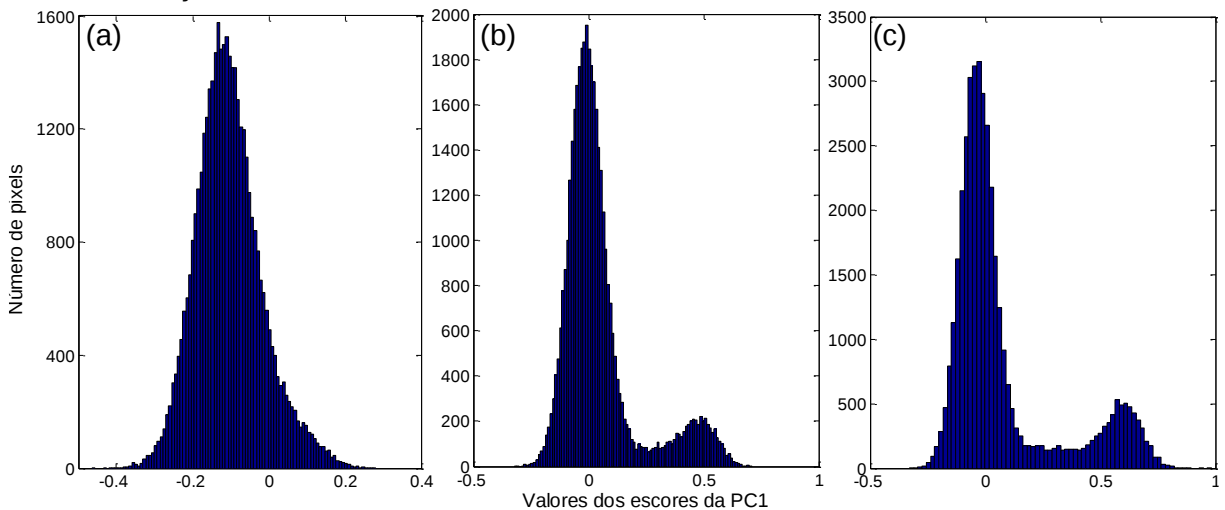
No caso das imagens dos anéis, os histogramas dos valores dos escores são bimodais, sendo a moda maior relacionada ao papel, que corresponde a maior parte dos *pixels* da imagem, e a menor ao anel. Nas imagens de amostras contendo os aditivos em menores concentrações (Figura 28a), os histogramas apresentam uma assimetria e não uma segunda moda distinguível, mas à medida que se aumenta a concentração do aditivo observa-se o aparecimento e a separação da outra moda (Figuras 28b e 28c). Essa diferença de aspecto dos histogramas, com a maior separação da segunda moda à medida que a concentração aumenta, já era esperada pelo fato das maiores concentrações resultarem em anéis com valores de escores maiores e mais diferentes dos do papel.

Agrupando-se as imagens de todas as concentrações, os valores dos escores dos *pixels* dos anéis são sempre os valores positivos da PC1. Assim, a moda da direita está relacionada com o anel e a da esquerda, que compreende valores negativos na PC1, está relacionada ao papel. É importante ressaltar que a rotina desenvolvida no MATLAB para a definição dos limites se baseou nessa disposição das modas.

Nessa rotina, inicia-se construindo os vetores dos histogramas (vetores dos valores dos escores e do número de *pixels* em cada valor) e calcula-se o desvio padrão dos valores dos escores localizados à esquerda do máximo do histograma. Após o cálculo do desvio padrão, o valor limite é determinado como sendo a soma do

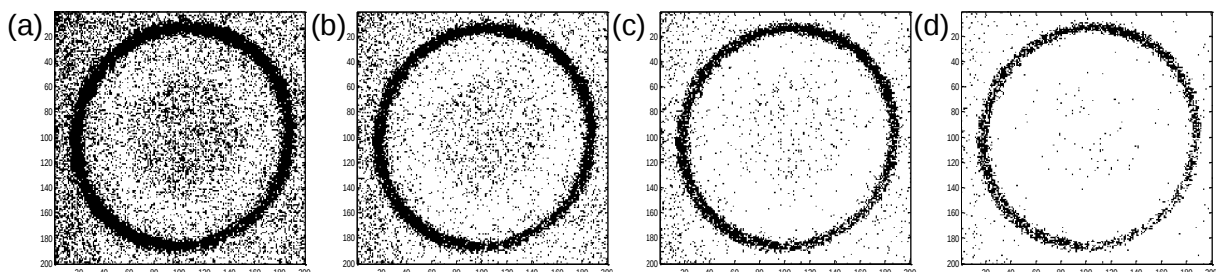
valor do máximo com um determinado número de desvios (ex. 3 desvios padrão). Com esses limites definidos e utilizando uma rotina semelhante à utilizada nas abordagens anteriores para extrair os espectros, os *pixels* com valores de escores superiores ao do limite (ex. valor do máximo acrescido de três desvios padrão) são selecionados e seus espectros extraídos. A fim de definir o melhor limite, variou-se o número de desvios padrão de um a quatro e observou-se os *pixels* que eram selecionados em cada caso (Figura 29).

Figura 28 – Histogramas dos escores da primeira componente principal da PCA de anéis de amostras contendo (a) 100 ppm, (b) 500 ppm e (c) 1.000 ppm do aditivo G agrupando todas as concentrações.



Fonte: da autora

Figura 29 – Imagens de uma amostra de um anel contendo 50 ppm do aditivo G com *pixels* selecionados utilizando (a) um desvio padrão, (b) dois desvios-padrões, (c) três desvios padrões e (d) quatro desvios padrões no cálculo dos limites (os *pixels* em preto são os que estão acima do limite).



Fonte: da autora

O que se pode perceber na Figura 29 é que para quantidades menores de desvios padrão (um e dois desvios) tem-se a seleção de muitos *pixels* do papel, enquanto que para quantidades maiores (três e quatro desvios) deixa-se de selecionar

alguns *pixels* do anel. Assim, deve-se decidir o que é mais importante, selecionar todos os *pixels* do anel, mesmo que isso signifique selecionar *pixels* do papel também, ou garantir que só os *pixels* do anel sejam selecionados, mesmo que alguns deles sejam deixados de fora da seleção. Neste trabalho, optou-se por utilizar três desvios padrão, por ser um valor no qual uma quantidade menor de *pixels* do papel é selecionada sem comprometer muito a quantidade de *pixels* do anel selecionados.

Apesar dessa abordagem selecionar uma maior quantidade de *pixels* do papel do que aquelas baseadas na seleção visual, a simplicidade, objetividade e rapidez justificam sua utilização desde que os resultados encontrados pelos modelos não sejam estatisticamente diferentes daqueles obtidos na abordagem anterior ou, se forem, que não sejam inferiores.

Após a extração dos espectros representativos, o mesmo procedimento adotado para a construção dos modelos de calibração anteriores foi adotado. Os resultados de todos os modelos construídos estão apresentados no Apêndice C.

Os modelos obtidos para o aditivo G utilizando os espectros médios foram estatisticamente diferentes daqueles utilizando os espectros somados e todos os modelos obtidos para o aditivo Y foram estatisticamente equivalentes, ao contrário do que foi observado nas abordagens anteriores. Dos modelos obtidos para o aditivo Y o que apresentou menor valor dos erros foi o construído com os espectros completos somados, o qual utilizou um número maior de variáveis latentes (VL=5). Como os valores dos erros não foram estatisticamente diferentes, o modelo escolhido foi aquele que apresentou menor número de variáveis latentes (VL=1), o qual utilizou as variáveis características dos aditivos nos espectros médios. No caso do aditivo T, o modelo utilizando os espectros médios e *jack-knife* não foi estatisticamente diferente daqueles obtidos com os espectros somados, mas usou uma variável latente a mais. No caso do aditivo W, o modelo com menor valor dos erros, construído com as regiões características dos espectros somados, foi estatisticamente diferente dos demais, contudo utilizou um número maior de variáveis latentes (VL=4) e não foi, portanto, selecionado. Dentre os modelos restantes, apenas o construído com os espectros médios e *jack-knife* foi estatisticamente diferente dos demais. Os resultados dos melhores modelos para cada aditivo estão apresentados na Tabela 8.

Tabela 8 – Melhores resultados obtidos para os espectros extraídos pela PCA com base nos histogramas dos valores dos escores da primeira componente (análise das imagens de todas as concentrações agrupadas).

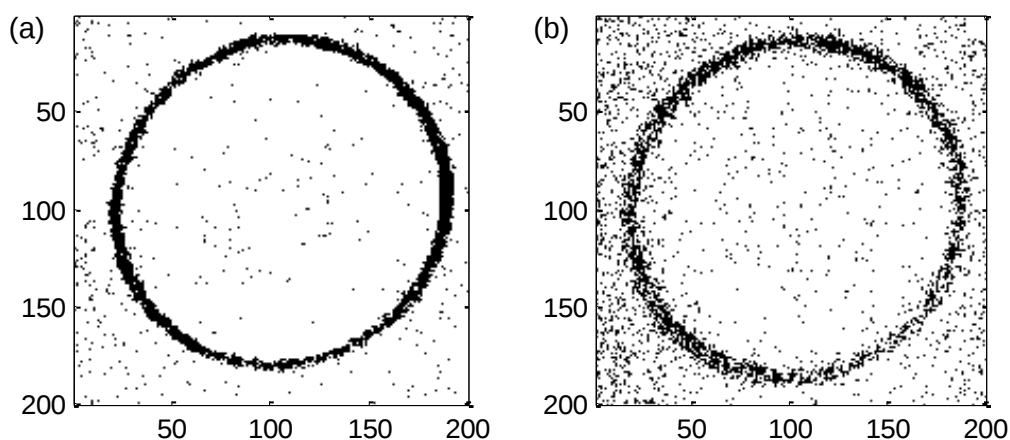
Aditivo	Espectro representativo	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
					Calibração	Validação
G	Somado + JK	1	64	71	0,9668	0,9624
T	Somado + JK	1	53	58	0,9796	0,9772
W	Somado	2	53	63	0,9721	0,9632
Y	Médio (regiões)	1	42	46	0,9829	0,9815

Comparando-se os resultados obtidos por essa abordagem com a baseada na determinação visual dos limites (Tabela 7), tem-se que os resultados para os aditivos T e W foram estatisticamente equivalentes e os resultados para os aditivos G e Y foram estatisticamente diferentes. No que diz respeito ao aditivo G, os erros (RMSEC e RMSECV) foram menores enquanto que para o aditivo Y eles foram maiores. É importante ressaltar, contudo, que o modelo construído para o aditivo Y apresentou uma redução do número de variáveis latentes (de 2 para 1). Apesar do melhor modelo ter sido construído com os espectros médios, tem-se que sua única diferença em relação aos modelos utilizando os espectros somados foi o menor número de variáveis latentes, podendo-se utilizar o modelo construído com as variáveis selecionadas pelo *jack-knife* nos espectros somados sem prejuízo aos resultados.

Uma vez que a construção das PCAs utilizando os anéis agrupados foi realizada a fim de melhorar a identificação visual dos *pixels* dos anéis e essa não ser mais uma condição necessária para a extração dos anéis, realizada a partir dos histogramas, resolveu-se testar a extração dos espectros dos anéis individualmente. Esse procedimento facilita a seleção da ROI por não necessitar de imagens pré-existentes de anéis produzidos a partir de misturas dos aditivos em diferentes concentrações para selecionar a ROI de novas imagens. Isso porque sempre que novas amostras fossem ser analisadas (amostras essas que não teriam sua concentração conhecida) seria necessário agrupá-las com imagens de anéis de concentrações conhecidas (valores elevados de concentração) do mesmo aditivo. Isso tornaria o processo de análise mais demorado e necessitaria do conhecimento prévio da identidade do aditivo contido na gasolina, já que o agrupamento das imagens deveria ser feito com imagens de amostras contendo o mesmo aditivo.

Na Figura 30, é possível observar que a análise individual, com base no histograma, de amostras de baixas concentrações resulta em uma extração de um número menor de *pixels* do anel e, conseqüentemente, de um maior número de *pixels* do papel, uma vez que alguns *pixels* do anel para essas imagens se confundem com os *pixels* do papel. Contudo, desde que não se tenham resultados significativamente inferiores aos dos modelos de calibração, é mais vantajoso realizar a extração individual dos espectros, devido à sua simplicidade e rapidez. A única dificuldade que se tem é que não se pode garantir que os *pixels* com escores positivos da PC1 serão correspondentes ao anel, pois o algoritmo NIPALS pode inverter esses valores. Contudo, no caso de haver a inversão, a qual pode ser percebida na análise rápida das imagens dos escores, basta substituir os valores dos escores pelos seus opostos antes da extração.

Figura 30 – Imagens de um anel contendo 100 ppm do aditivo G com *pixels* selecionados utilizando três desvios padrões com a PCA construída (a) agrupando-se as imagens de todas as concentrações e (b) individualmente para cada imagem.



Fonte: da autora

Após a extração, o mesmo procedimento adotado para a construção dos modelos de calibração anteriores foi repetido. Os resultados de todos os modelos construídos estão apresentados no Apêndice D.

Os modelos utilizando os espectros somados foram estatisticamente diferentes (RMSEC e RMSECV menores), exceto os do aditivo G. No caso do aditivo T, os três modelos construídos utilizando os espectros somados foram equivalentes. No caso do aditivo W, os modelos com as variáveis selecionadas pelo *jack-knife* e com as regiões características foram equivalentes, tendo-se selecionado o último por ter um

número menor de variáveis latentes (VL=2). No caso do Y, o modelo construído apenas com as regiões características foi estatisticamente diferente dos demais e, portanto, foi selecionado. Os resultados dos melhores modelos para cada aditivo estão apresentados na Tabela 9.

Tabela 9 – Melhores resultados obtidos para os espectros extraídos pela PCA com base nos histogramas dos valores dos escores da primeira componente (análise individual das imagens).

Aditivo	Espectro representativo	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
					Calibração	Validação
G	Somado + JK	1	71	79	0,9594	0,9534
T	Somado (regiões)	1	41	46	0,9879	0,9855
W	Somado (regiões)	2	35	41	0,9879	0,9850
Y	Somado (regiões)	2	31	40	0,9908	0,9857

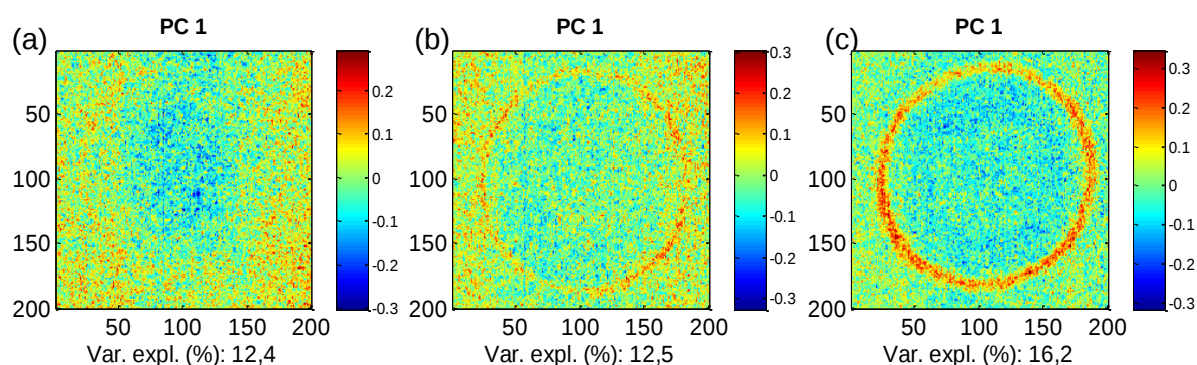
Comparando-se os resultados obtidos por essa abordagem com os mostrados na Tabela 8, tem-se que eles foram estatisticamente equivalentes para os aditivos G, T e Y e estatisticamente diferentes para o aditivo W, o qual apresentou uma redução nos valores dos erros. Considerando os aditivos G, T e Y, apenas o G apresentou um aumento nos valores dos erros, os demais apresentaram um decréscimo. Uma possível causa para esse resultado é o fato dos anéis desse aditivo serem menos definidos e a melhor definição causada pelo agrupamento das imagens, utilizada na abordagem anterior, resultar na seleção de uma maior quantidade de *pixels* do anel em comparação com a quantidade de *pixels* do papel.

Com esses resultados pode-se concluir, de maneira geral, que a seleção da ROI a partir dos histogramas dos valores dos escores da primeira componente da PCA construída individualmente para cada anel é a melhor abordagem entre aquelas que foram testadas. Trata-se de uma forma de extração de espectros mais objetiva, rápida e cujos resultados são melhores ou equivalentes aos obtidos nas abordagens anteriores, considerando-se não só os valores dos erros, mas também o número de variáveis latentes utilizado. Além disso, uma vez que os melhores resultados foram obtidos para os espectros somados e subtraídos dos espectros do papel, pode-se concluir que esses são os espectros que melhor representam os anéis (ROI) e, portanto, serão usados nos próximos modelos.

5.2. Resultados consolidados

Como já foi visto, as regiões consideradas características dos aditivos (entre 1.670 e 1.801 nm entre 2.238 e 2.425 nm) são regiões do espectro relacionadas aos primeiros sobretons das vibrações de estiramento C-H e às bandas de combinações de vibrações C-H. Não se tem, então, informações espectrais específicas dos grupos funcionais característicos dos aditivos detergentes dispersantes, mas sim das suas cadeias carbônicas, as quais podem ser semelhantes às dos componentes químicos presentes no papel e na gasolina comum. Assim, dependendo da composição da gasolina comum, é possível que exista interferência dos seus componentes no momento da seleção dos *pixels* das imagens dos anéis com base nos valores dos escores. Nesta segunda etapa do trabalho, ao se empregar no estudo 10 gasolinas comuns de diferentes distribuidores, foram observadas diferenças não só quanto à coloração desse combustível (algumas praticamente incolores e outras bastante amareladas), mas também quanto às imagens dos escores da PC1. Em algumas imagens dos escores não foi possível identificar a presença do anel, mas em outros casos a sua presença ficou bastante evidente, como pode ser visto na Figura 31. Assim, é possível que a composição da gasolina comum possa afetar a seleção dos *pixels* do anel. Além disso, essas diferenças entre as imagens indicam que a composição dos resíduos das gasolinas é diferente devido à grande variabilidade de composição desse tipo de combustível.

Figura 31 – Imagens dos escores da primeira componente principal da PCA das regiões características dos aditivos de anéis preparados a partir de gasolinas comuns de diferentes distribuidores. Anel (a) não identificado, (b) pouco definido e (c) bem definido.

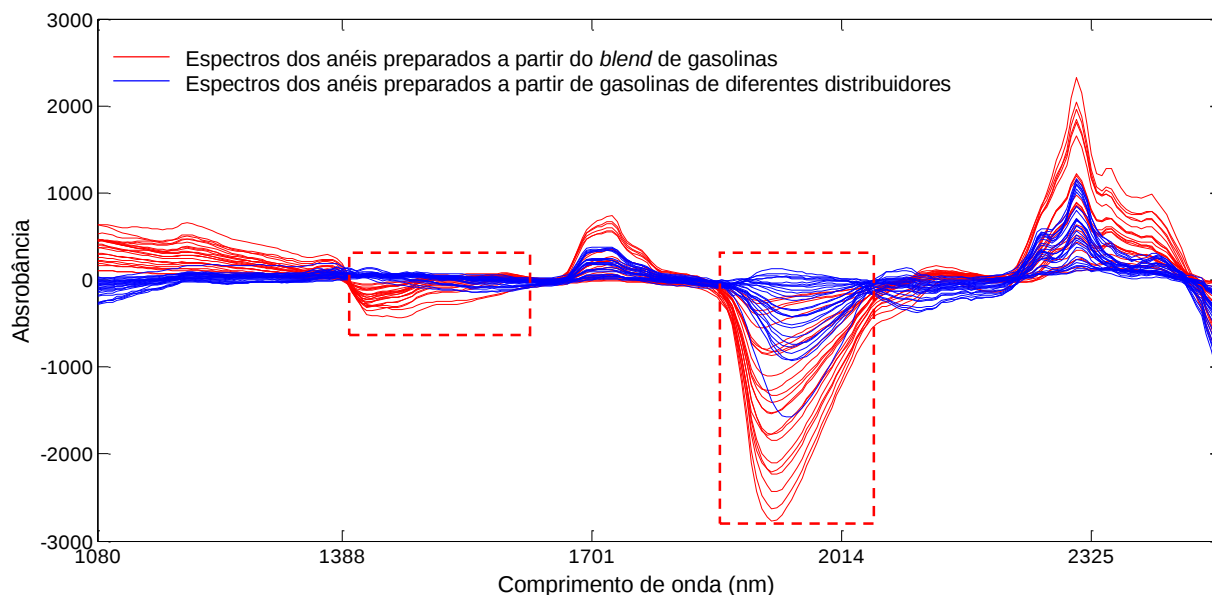


Fonte: da autora

O fato da composição dos resíduos das gasolinas comuns ser diferente torna necessário incluir gasolinas diferentes nos modelos, de forma a introduzir essa variabilidade e torná-los mais robustos. Isso foi feito preparando as misturas dos aditivos em diferentes amostras de gasolina.

Os espectros representativos dos anéis preparados a partir das misturas de gasolina comum e aditivo utilizando como gasolina base o *blend* de gasolinas (primeira etapa do trabalho) e as gasolinas comuns de diferentes distribuidores (segunda etapa do trabalho) estão apresentados na Figura 32. Observa-se grande diferença entre os dois grupos de espectros, diferenças essas que não estão relacionadas à composição das gasolinas comuns utilizadas como base, mas à preparação do anel. Isso porque entre as regiões nas quais foram observadas as maiores diferenças, estão as relacionadas às combinações de vibrações das moléculas de água (entre 1.413 e 1.657 nm e entre 1.864 e 2.057 nm), as quais estão presente no papel, na forma de umidade, e não na gasolina.

Figura 32 – Espectros de todos os anéis do aditivo G somados e subtraídos dos espectros do papel extraídos através dos histogramas da primeira componente da PCA.



Fonte: da autora

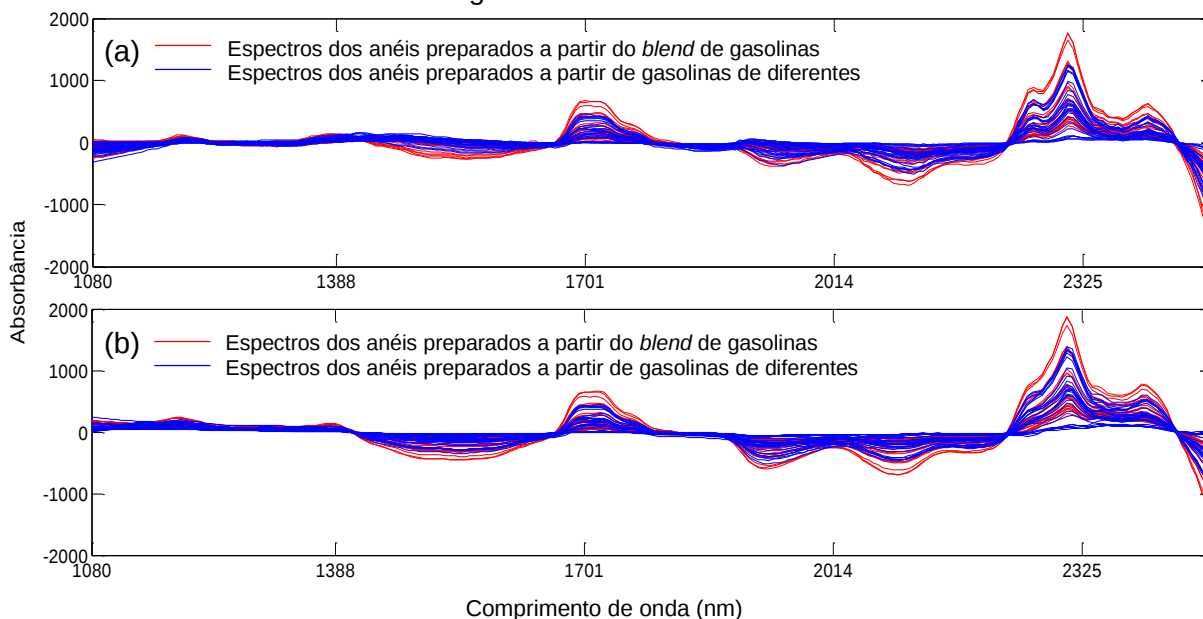
Tendo-se observado essas diferenças e sabendo-se que a compensação dos espectros do papel foi feita de forma diferente, na primeira etapa utilizou-se a média dos espectros de uma imagem de papel (não submetido ao aquecimento) e na segunda a média dos espectros de parte de um dos papéis utilizado no preparo dos anéis (submetido ao aquecimento), e que as imagens foram adquiridas em períodos

diferentes do ano (intervalo de cerca de 6 meses), concluiu-se que essas eram as causas da diferença. Uma explicação para o que foi observado é que ao ser submeter o papel ao aquecimento, parte da sua umidade é perdida e não é recuperada. Assim, seu nível se torna inferior ao do papel não submetido ao aquecimento, principalmente quando o nível de umidade do ar for mais baixo, como foi observado no período de obtenção das imagens utilizadas na segunda etapa do trabalho.

Sabendo-se que anéis serão preparados em diferentes momentos, avaliou-se que a melhor forma de compensar a informação do papel é fazendo a subtração dos espectros do próprio papel, o qual deve ter água retida em suas fibras de forma homogênea, garantido, assim, uma melhor correção da sua contribuição na região do anel. Para fazer essa compensação, duas abordagens foram testadas. No primeiro caso, optou-se por calcular a média dos espectros dos *pixels* das bordas das imagens do papel e, no segundo, dos *pixels* cujos valores dos escores estão abaixo do máximo do histograma, *pixels* necessariamente do papel. Essa segunda abordagem foi feita através de uma modificação na rotina utilizada na extração dos espectros sendo, portanto, mais simples e rápida. A diferença dessa segunda abordagem é que os espectros são selecionados considerando toda a região da imagem que não contém o anel, ou seja, os espectros do papel podem ser obtidos na região interna e externa ao anel. A região interna do anel compreende a região onde a amostra e o tolueno foram gotejados durante o preparo do anel. Como não se sabe ao certo que componentes são levados pelo tolueno para o anel e que outros componentes restam no centro do papel, é possível que ao se considerar os espectros dessa região central alguma informação relacionada a esses componentes também seja eliminada.

Os espectros resultantes dessas duas abordagens aparentemente não são diferentes (Figura 33), mas, a fim de compará-los, foram construídos modelos de regressão para os aditivos T e W. Para cada abordagem, o melhor modelo foi aquele com menores valores de erros percentuais médios de previsão, calculados como a média dos erros percentuais das concentrações previstas das amostras. É importante ressaltar que quando mais de um modelo apresentou valores próximos desses erros, o modelo escolhido foi o com menor valor de RMSECV. Além disso, os modelos construídos com seleção de variáveis pelo *jack-knife* que utilizaram quase todas as variáveis e que, portanto, não apresentaram melhorias significativas nos valores de RMSECV, não foram selecionados.

Figura 33 – Espectros de todos os anéis do aditivo G somados e subtraídos dos espectros do papel extraídos através dos histogramas da primeira componente da PCA. Subtração dos espectros do papel (a) dos *pixels* das bordas da imagem e (b) dos *pixels* com valores de escores abaixo do máximo do histograma.



Fonte: da autora

Escolheu-se incluir amostras produzidas a partir de gasolinas diferentes tanto no conjunto de calibração quanto de previsão porque testes preliminares mostraram que utilizar apenas as amostras produzidas a partir do *blend* de gasolinas comuns na construção dos modelos resultava em erros muito elevados de previsão das amostras de gasolinas diferentes. Isso porque sem incluir amostras de composições diferentes, representativas das amostras que podem vir a ter o parâmetro de interesse previsto, o modelo não é robusto o suficiente.

Os resultados de todos os modelos construídos para os aditivos T e W, incluindo os erros de previsão de cada amostra, estão apresentados nos Apêndices E a L e os resultados dos melhores modelos estão apresentados na Tabela 10. Comparando-se os resultados apresentados nessa tabela, observa-se que os menores valores de erros percentuais médios de previsão foram obtidos para os modelos construídos utilizando-se os espectros dos *pixels* com valores de escores abaixo do máximo do histograma. Uma vez que a diferença entre os resultados não é grande (5% no máximo), pode-se utilizar qualquer uma das abordagens, contudo, pelo fato da segunda abordagem ser simples e rápida ela foi escolhida.

Tabela 10 – Resultados obtidos para os espectros somados e subtraídos dos espectros (1) dos *pixels* das bordas da imagem e (2) dos *pixels* com valores de escores abaixo do máximo do histograma.

Aditivo	Abordagem	Pré-processamento	VL	RMSECV (ppm)	Erro médio de previsão (%)
T	1	2° derivada (janela de 11 ptos) (regiões características dos aditivos)	3	24	14
	2	1° derivada (janela de 7 ptos) + JK	2	28	9
W	1	2° derivada (janela de 11 ptos)	1	52	23
	2	1° derivada (janela de 11 ptos)	2	43	19

Tendo-se definido que a melhor abordagem é a segunda, os modelos dos aditivos G e Y foram construídos, os resultados de todos os modelos desses dois aditivos estão apresentados nos Apêndices M a P e os resultados dos melhores modelos para os quatro aditivos estão apresentados na Tabela 11. A escolha dos melhores modelos se baseou nos valores dos erros percentuais médios de previsão, dos erros da validação cruzada (RMSECV) e nos números de variáveis latentes.

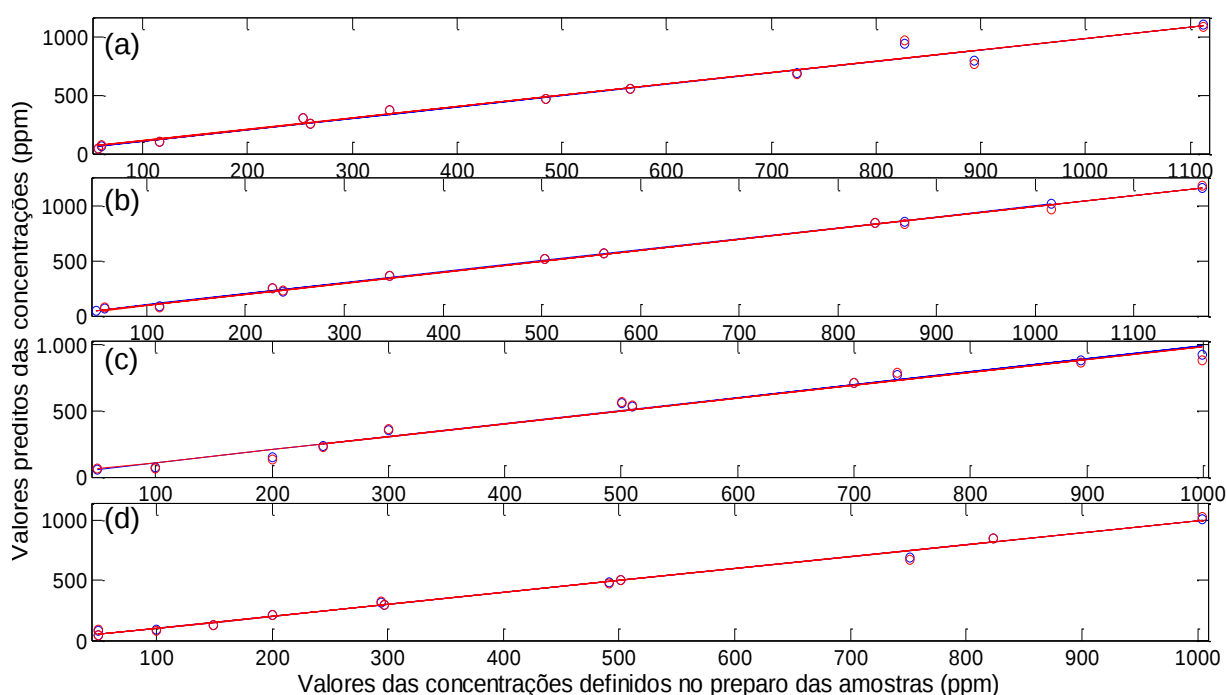
Tabela 11 – Resultados obtidos para os espectros somados e subtraídos dos espectros dos *pixels* com valores de escores abaixo do máximo do histograma.

Aditivo	Pré-processamento	VL	RMSECV (ppm)	Erro médio de previsão (%)
G	2° derivada (janela de 11 pontos) + JK	1	60	14
T	1° derivada (janela de 7 ptos) + JK	2	28	9
W	1° derivada (janela de 11 ptos)	2	43	19
Y	2° derivada (janela de 11 pontos) + JK	2	30	2

Os gráficos dos valores preditos em função das concentrações definidas no preparo das amostras para os melhores modelos de cada aditivo estão apresentados na Figura 34. Através desses gráficos é possível observar ajustes de boa qualidade, principalmente para os casos dos aditivos T e Y.

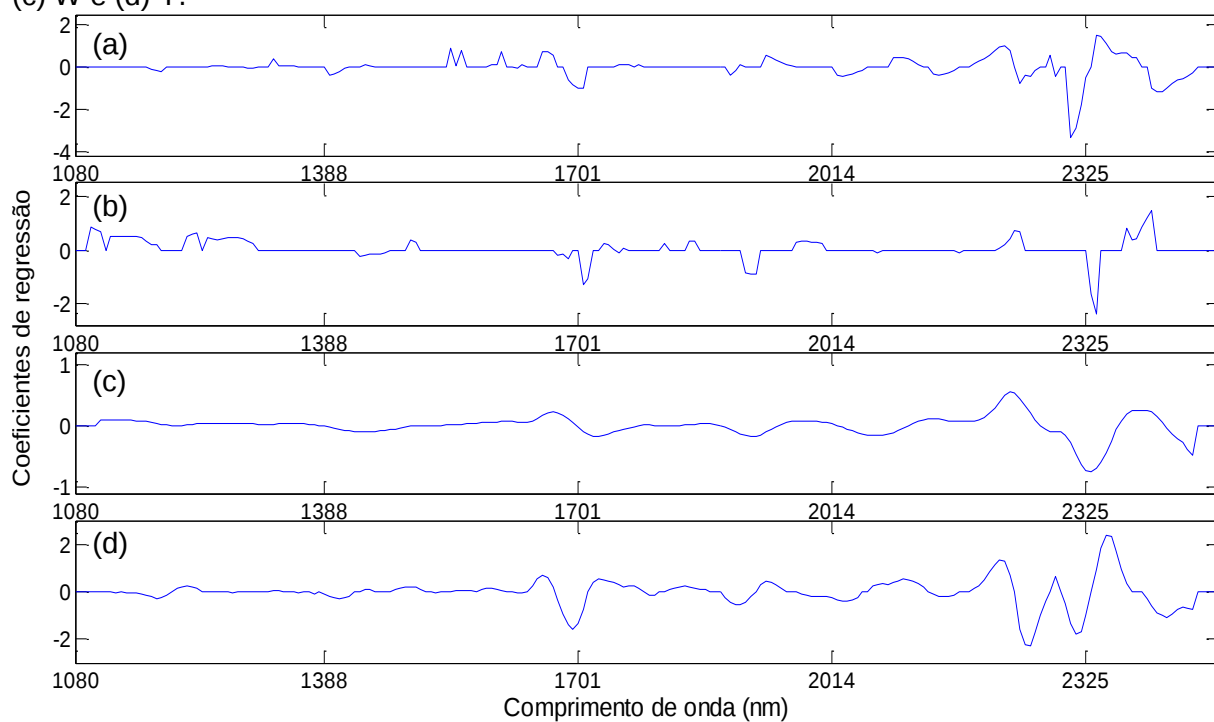
Os gráficos dos coeficientes de regressão (Figura 35) permitem observar que a região cujas variáveis apresentam maiores coeficientes de regressão está entre 2.275 e 2.338 nm, uma das regiões consideradas características dos aditivos e que está relacionada às bandas de combinações de vibrações C-H. Assim, essa região é importante não só na identificação do anel, mas também na construção dos modelos de regressão.

Figura 34 – Gráficos dos valores preditos das concentrações em função das concentrações definidas no preparo das amostras para os aditivos (a) G, (b) T, (c) W e (d) Y.



Fonte: da autora

Figura 35 – Gráficos dos coeficientes de regressão para os modelos dos aditivos (a) G, (b) T, (c) W e (d) Y.



Fonte: da autora

Por fim, tem-se que os modelos construídos com os espectros extraídos a partir dos histogramas dos valores dos escores da PCA e que utilizaram os espectros somados e subtraídos dos espectros médios do papel como representativos da ROI permitem quantificar os aditivos nas gasolinas. Contudo, como diferentes estratégias de pré-processamento foram escolhidas para cada aditivo e, conseqüentemente, modelos diferentes devem ser usados na determinação da concentração de cada tipo de aditivo, é necessário conhecer previamente o aditivo contido na amostra antes de iniciar o processo analítico.

5.2.1. Modelos de classificação

Uma vez que a composição das amostras reais, no que diz respeito ao tipo de aditivo que elas contêm, pode não ser informada pela ANP, torna-se necessário realizar uma etapa prévia de classificação das mesmas visando à identificação do aditivo. Para isso, foram construídos modelos LDA utilizando os algoritmos de seleção de variáveis SPA e GA (selecionando-se variáveis no espectro completo e apenas nas regiões características dos aditivos) e utilizando apenas as variáveis das regiões características.

Assim como na regressão, decidiu-se utilizar os espectros somados e subtraídos dos espectros do papel. No que diz respeito aos pré-processamentos, utilizou-se a 1ª e 2ª derivada com filtro de Savitzky-Golay com janelas de 7 e 11 pontos seguida de normalização pela faixa e a suavização com filtro de Savitzky-Golay com janelas de 3 e 7 pontos também seguida de normalização pela faixa. A normalização foi utilizada para corrigir variações devido às diferenças de concentração, já que as intensidades dos espectros eram muito maiores para as concentrações de 1.000 ppm.

Os resultados dos modelos utilizando a seleção de variáveis pelo SPA estão apresentados nos Apêndices Q e R e, resumidamente, na Tabela 12.

O melhor modelo, considerando-se a porcentagem de acertos no conjunto de previsão, foi aquele construído com os espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa, cuja porcentagem de acertos foi de 90 %.

Tabela 12 – Resultados de classificação (SPA-LDA).

Pré-processamento	Quantidade de variáveis	Acertos na previsão (%)
Suavização (janela de 3 pontos) + normalização pela faixa	56	79
Suavização (janela de 7 pontos) + normalização pela faixa	60	90
1° derivada (janela de 7 pontos) + normalização pela faixa	50	85
1° derivada (janela de 11 pontos) + normalização pela faixa	35	82
2° derivada (janela de 7 pontos) + normalização pela faixa	27	74
2° derivada (janela de 11 pontos) + normalização pela faixa	64	85
Suavização (janela de 3 pontos) + normalização pela faixa (*)	41	87
Suavização (janela de 7 pontos) + normalização pela faixa (*)	27	79
1° derivada (janela de 7 pontos) + normalização pela faixa (*)	26	77
1° derivada (janela de 11 pontos) + normalização pela faixa (*)	28	69
2° derivada (janela de 7 pontos) + normalização pela faixa (*)	25	77
2° derivada (janela de 11 pontos) + normalização pela faixa (*)	64	85

(*) Seleção nas regiões características dos aditivos.

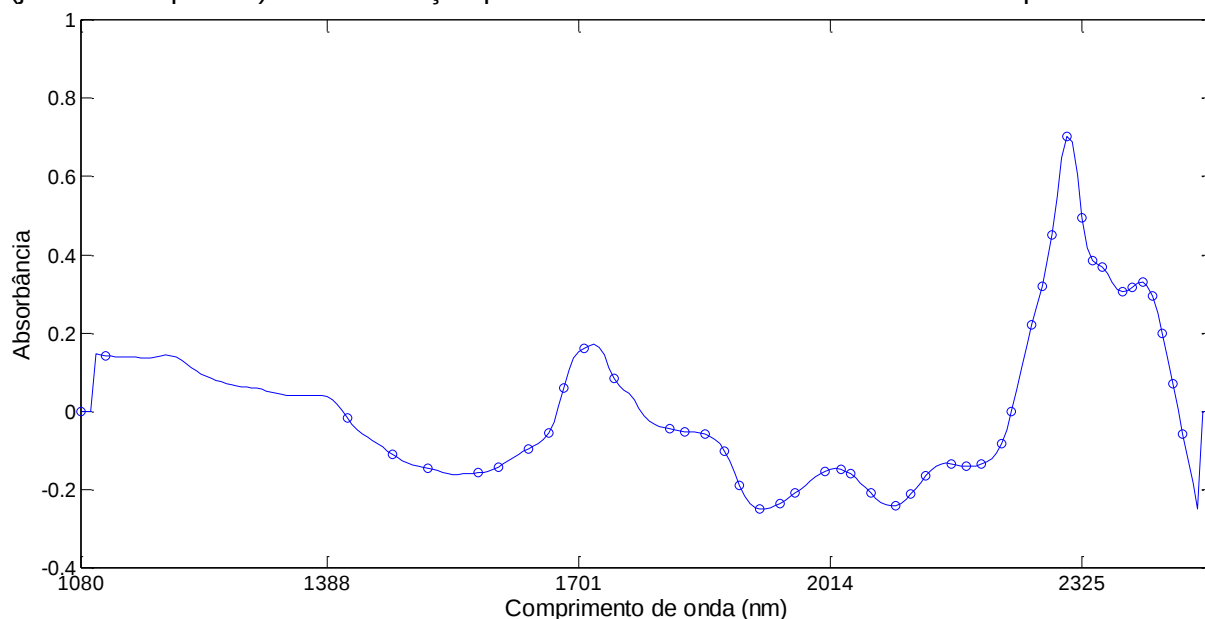
As variáveis selecionadas pelo SPA estão marcadas no espectro da Figura 36, pré-processado conforme no modelo. O que se pode observar nesse espectro é que apesar de algumas das variáveis selecionadas fazerem parte das regiões consideradas características dos aditivos (entre 1.670 e 1.801 nm e entre 2.238 e 2425 nm), outras variáveis foram selecionadas no restante do espectro, inclusive nas regiões relacionadas às combinações de vibrações das moléculas de água (entre 1.413 e 1.657 nm e entre 1.864 e 2.057 nm). Assim, o espectro como um todo influencia na separação das classes dos aditivos.

Os gráficos das funções discriminantes (FDs), apresentados na Figura 37, permitem observar que as amostras dos aditivos G e W apresentam certa sobreposição e que as amostras dos aditivos T e Y encontram-se separadas na FD2. Apesar da sobreposição parcial das amostras dos aditivos G e W observada nos gráficos de duas funções discriminantes, o uso de três funções permite a separação das quatro classes no espaço com poucos erros de classificação. É importante observar que as duas primeiras funções discriminantes (FD1 e FD2) permitem separar os aditivos T e Y dos aditivos G e W.

O modelo selecionado apresenta um total de 12 erros de classificação, considerando-se os conjuntos de treinamento, validação e previsão, desses erros 4 estão relacionados ao aditivo G (amostras classificadas como W) e 7 estão relacionadas ao aditivo W (amostras classificadas como G), como pode ser visto na Tabela 13. Assim, o que se pode concluir é que a maior parte dos erros é resultado

da sobreposição das amostras dos aditivos G e W. O fato dessas amostras estarem próximas no gráfico das FDs permite supor que suas composições químicas sejam semelhantes.

Figura 36 – Espectro médio pré-processado com suavização com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa com as variáveis selecionadas pelo SPA.



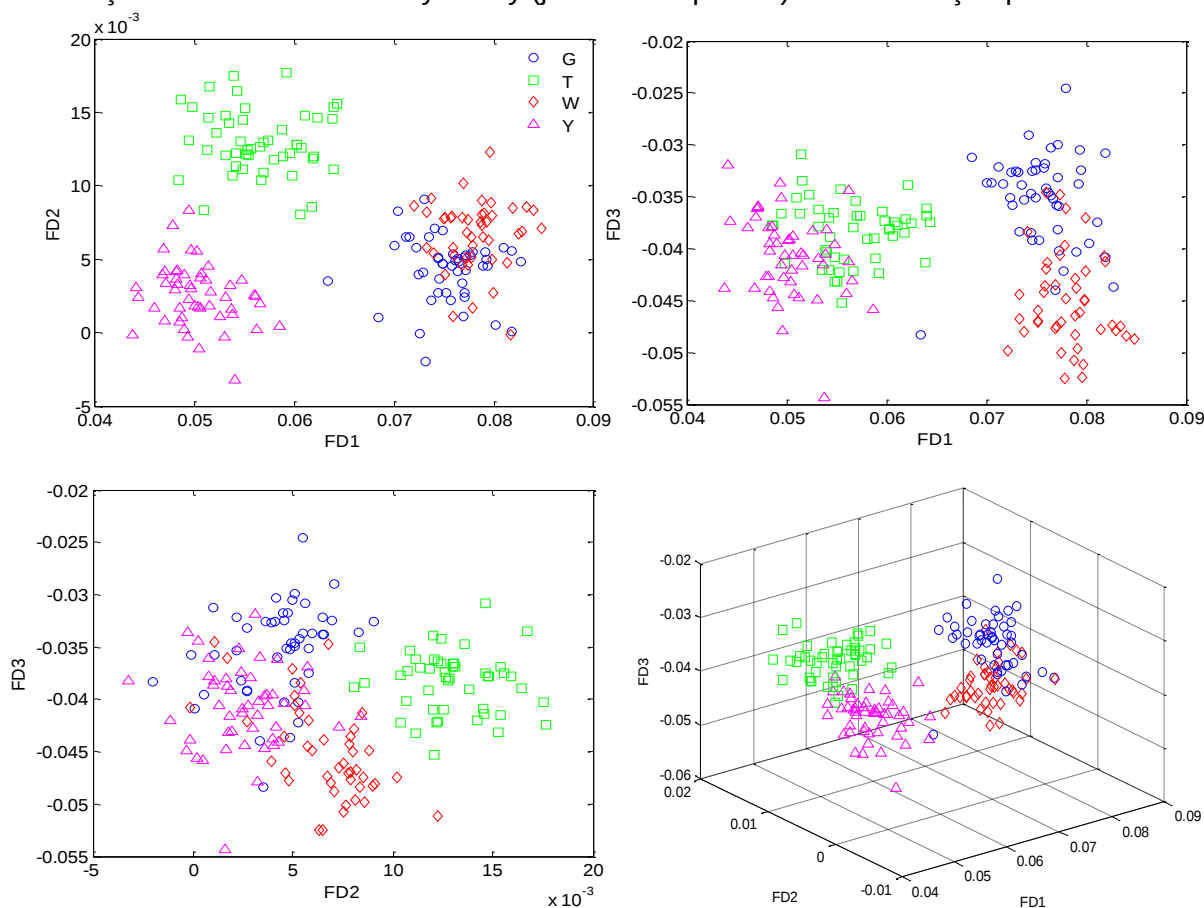
Fonte: da autora

No caso do GA, repetido 5 vezes por ser um algoritmo estocástico, o melhor modelo construído a partir de cada pré-processamento foi escolhido com base na porcentagem de acertos no conjunto de previsão. No caso de modelos com a mesma porcentagem de acertos, a escolha se baseou na quantidade de erros nos conjuntos de validação e treinamento. Os resultados dos modelos utilizando a seleção de variáveis pelo GA estão apresentados nos Apêndices S e T e, resumidamente, na Tabela 14.

Os melhores modelos, considerando-se a porcentagem de acertos no conjunto de previsão e o número de variáveis selecionadas foram aqueles construídos com os espectros pré-processados com a 2ª derivada com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa e com a suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa, tendo-se selecionado as variáveis apenas nas regiões características dos aditivos. Os modelos apresentaram 92 % de acertos e foram construídos com 20 e 28 variáveis, respectivamente. Outros modelos

apresentaram a mesma porcentagem de acertos, mas foram construídos com praticamente o dobro de variáveis.

Figura 37 – Funções discriminantes do modelo SPA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa.



Fonte: da autora

Tabela 13 – Resultados do modelo SPA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa, onde N é a quantidade de amostras de cada classe.

Classe verdadeira	N	Classe predita			
		G	T	W	Y
G	48	43	-	4	1
T	49	-	49	-	-
W	45	7	-	38	-
Y	50	-	-	-	50

Tabela 14 – Resultados de classificação (GA-LDA).

Pré-processamento	Quantidade de variáveis	Acertos na previsão (%)
Suavização (janela de 3 pontos) + normalização pela faixa	53	92
Suavização (janela de 7 pontos) + normalização pela faixa	48	92
1º derivada (janela de 7 pontos) + normalização pela faixa	67	92
1º derivada (janela de 11 pontos) + normalização pela faixa	59	90
2º derivada (janela de 7 pontos) + normalização pela faixa	38	90
2º derivada (janela de 11 pontos) + normalização pela faixa	53	87
Suavização (janela de 3 pontos) + normalização pela faixa (*)	28	92
Suavização (janela de 7 pontos) + normalização pela faixa (*)	15	90
1º derivada (janela de 7 pontos) + normalização pela faixa (*)	16	90
1º derivada (janela de 11 pontos) + normalização pela faixa (*)	22	79
2º derivada (janela de 7 pontos) + normalização pela faixa (*)	20	92
2º derivada (janela de 11 pontos) + normalização pela faixa (*)	23	87

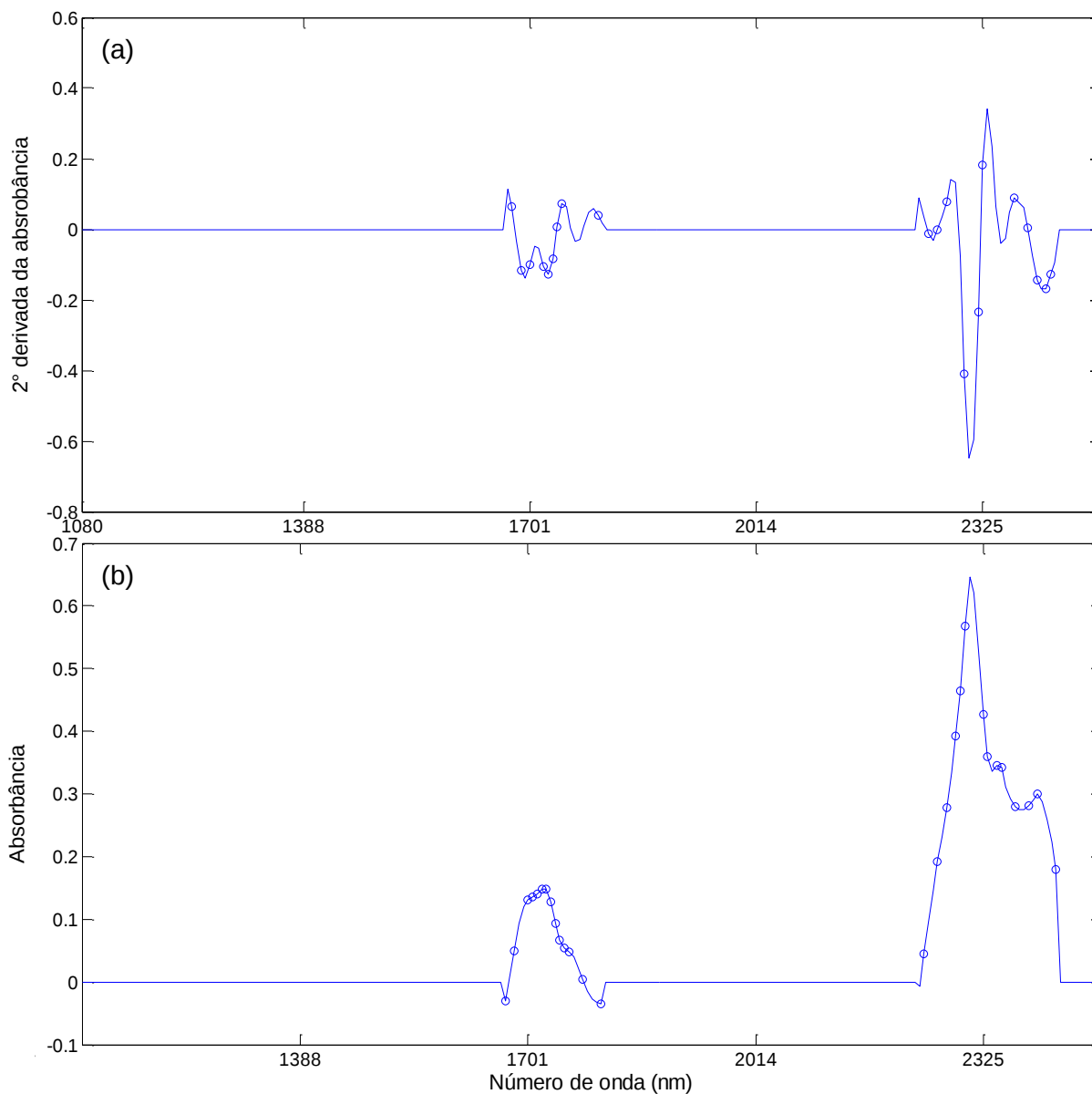
(*) Seleção nas regiões características dos aditivos.

As variáveis selecionadas pelo GA estão marcadas nos espectros da Figura 38, pré-processados conforme nos modelos. Nesses dois modelos, uma vez que as variáveis foram selecionadas apenas nas regiões características dos aditivos, tem-se que apenas elas, e não o espectro completo, foram responsáveis pela classificação, ao contrário do modelo construído com as variáveis selecionadas pelo SPA.

Através da análise dos gráficos das FDs, apresentados nas Figura 39 e 40, observa-se uma melhor separação das classes no segundo modelo, estando as amostras dos aditivos T e Y mais separadas na FD2 e as quatro classes mais separadas no gráfico das três funções discriminantes. Associado a essa maior separação, tem-se uma quantidade de erros totais menor para o segundo modelo, o qual apresenta 11 erros (94 % de acertos) enquanto que o primeiro apresenta 17 erros (91 % de acertos). Assim, o modelo GA-LDA construído utilizando os espectros pré-processados com a suavização foi escolhido como o melhor.

Dos 11 erros totais de classificação, 4 estão relacionados ao aditivo G (amostras classificadas como W) e 6 estão relacionados ao aditivo W (amostras classificadas como G), como pode ser visto na Tabela 15. Esse resultado é bastante semelhante àquele encontrado para o modelo SPA-LDA, corroborando a hipótese de que os aditivos devem ter composições químicas semelhantes.

Figura 38 – Espectros médios pré-processado com (a) 2º derivada com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa e (b) suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa com as variáveis selecionadas pelo GA apenas nas regiões características do espectro.

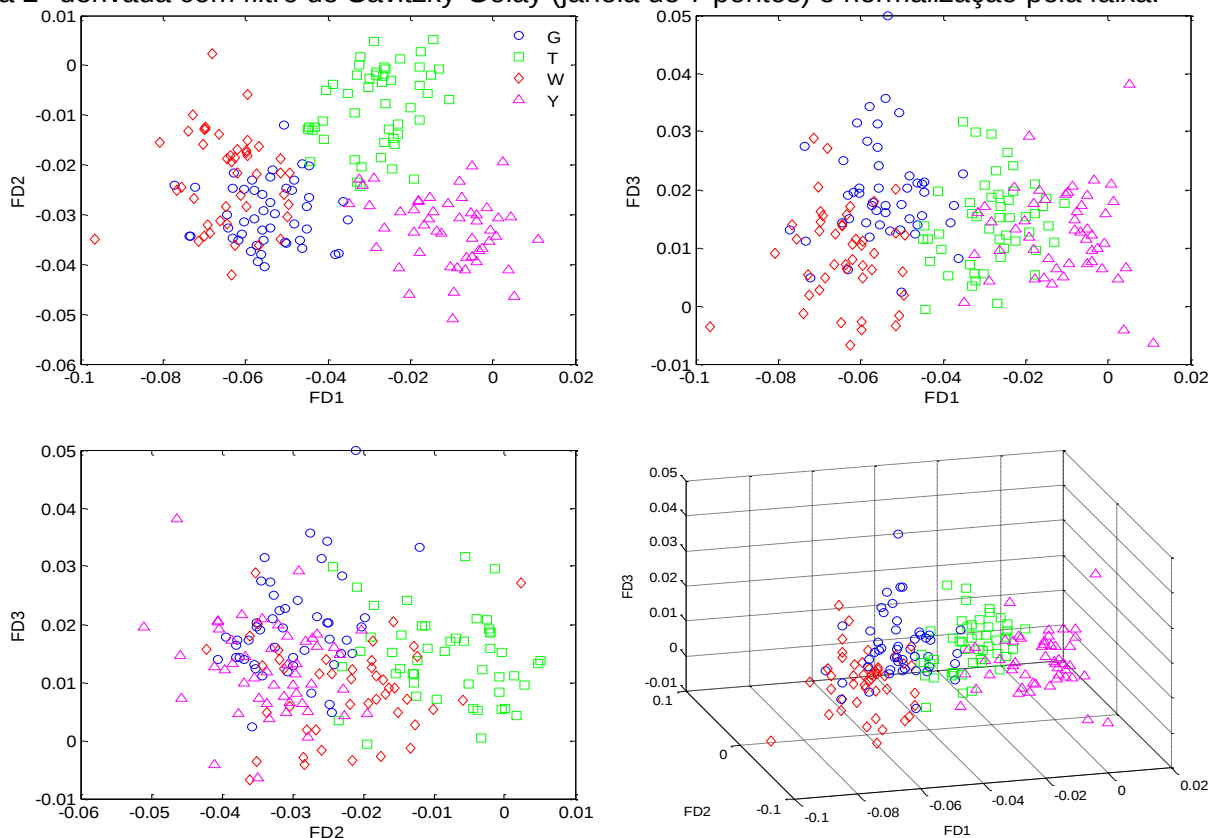


Fonte: da autora

Como as regiões características dos aditivos são formadas por 53 variáveis e o conjunto de treinamento do LDA é composto por 114 amostras, é possível utilizar apenas essas variáveis na construção dos modelos. Os resultados dos modelos utilizando essas variáveis estão apresentados no Apêndice U e, resumidamente, na Tabela 16. Os melhores modelos, considerando-se a porcentagem de acertos no conjunto de previsão, foram aqueles construídos com os espectros pré-processados com a suavização com filtro de Savitzky-Golay (janela de 3 pontos) e 2º derivada com

filtro de Savitzky-Golay (janela de 7 pontos), ambos seguidos de normalização pela faixa. Dentre esses dois modelos, o primeiro foi considerado melhor por ter uma menor quantidade total de erros.

Figura 39 – Funções discriminantes do modelo GA-LDA dos espectros pré-processados com a 2ª derivada com filtro de Savitzky-Golay (janela de 7 pontos) e normalização pela faixa.

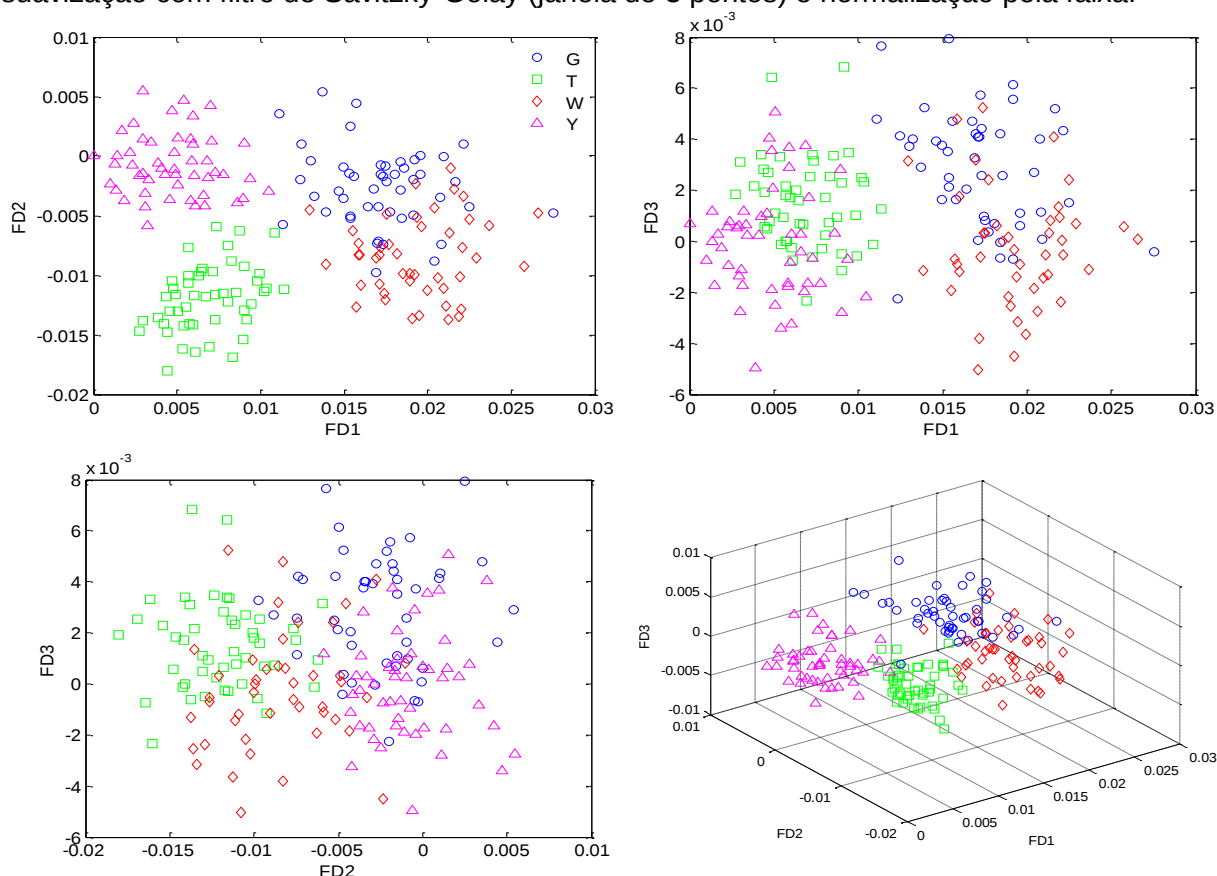


Fonte: da autora

Os gráficos das funções discriminantes, apresentados na Figura 41, permitem observar a separação das amostras dos aditivos T e Y na FD2 e a separação das quatro classes no gráfico das três funções discriminantes, da mesma forma que foi observado nos modelos anteriores.

O modelo apresentou um total de 12 erros de classificação, dos quais 5 estão relacionados ao aditivo G (4 amostras classificadas como W e 1 como Y) e 4 estão relacionados ao aditivo W (amostras classificadas como G), como pode ser visto na Tabela 17. A maior parte dos erros ainda se deve a classificações incorretas das amostras G e W, contudo, nesse modelo 3 amostras T e Y também se confundem.

Figura 40 – Funções discriminantes do modelo GA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa.



Fonte: da autora

Tabela 15 – Resultados do modelo GA-LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa, onde N é a quantidade de amostras de cada classe.

Classe verdadeira	N	Classe predita			
		G	T	W	Y
G	48	43	-	4	1
T	49	-	49	-	-
W	45	6	-	39	-
Y	50	-	-	-	50

Considerando os três melhores modelos, construídos com as variáveis selecionadas pelo SPA, GA e com todas as variáveis das regiões características dos aditivos, tem-se que o melhor foi aquele construído com as variáveis selecionadas pelo GA. Esse modelo apresentou a menor quantidade de erros de previsão e de erros totais, além de ter utilizado uma quantidade menor de variáveis em comparação aos demais.

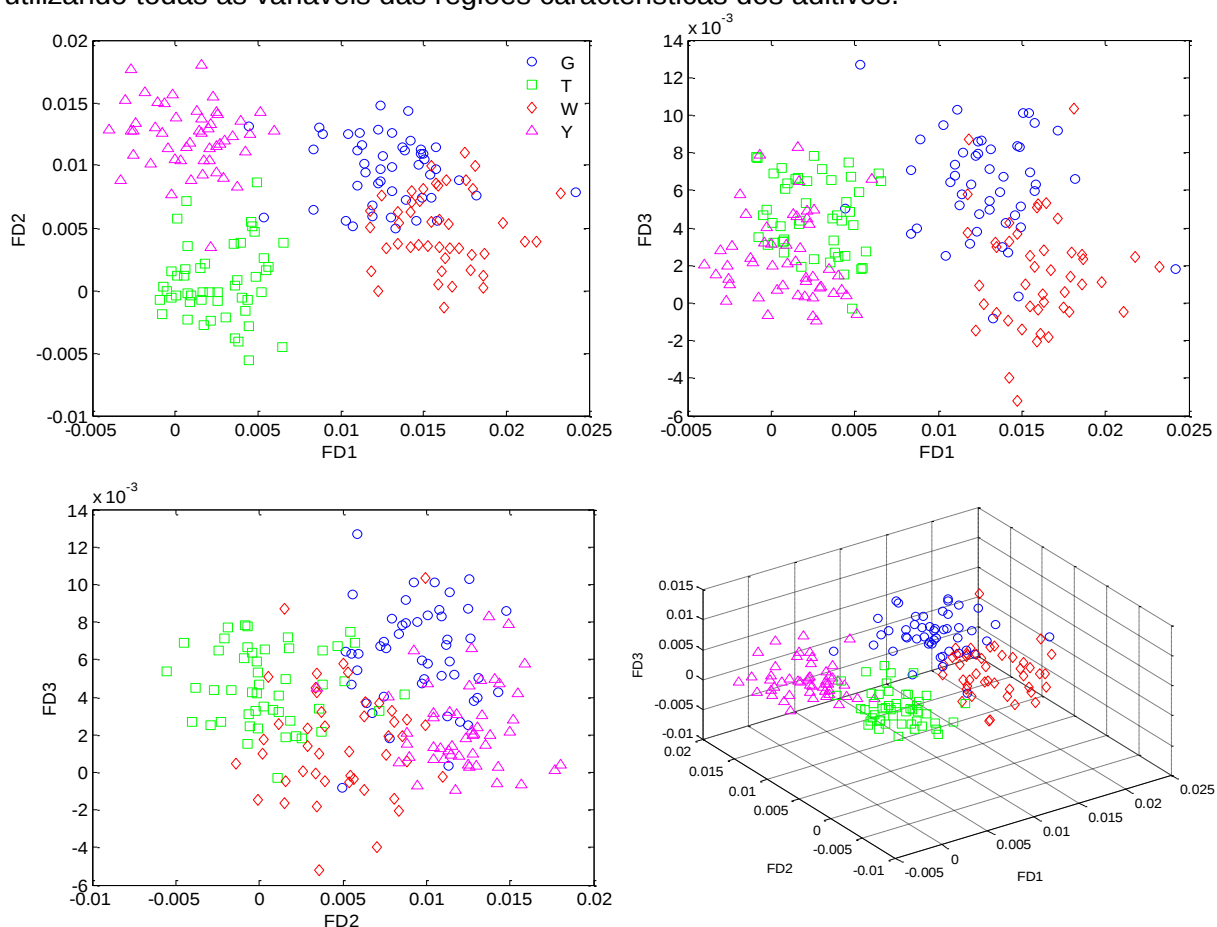
Diante desses resultados, nos quais a maioria das classificações incorretas envolveram amostras dos aditivos G e W, buscou-se calcular os erros percentuais

médios de previsão de concentração resultantes das classificações incorretas (amostras de um dado aditivo sendo quantificadas com o modelo de outro aditivo) e a possibilidade de se construir um modelo único para esses dois aditivos. Os resultados das predições de concentração estão apresentados no Apêndice V e, resumidamente, na Tabela 18.

Tabela 16 – Resultados de classificação (LDA) obtidos para as variáveis das regiões características dos aditivos.

Pré-processamento	Acertos na previsão (%)
Suavização (janela de 3 pontos) + normalização pela faixa	90
Suavização (janela de 7 pontos) + normalização pela faixa	72
1° derivada (janela de 7 pontos) + normalização pela faixa	72
1° derivada (janela de 11 pontos) + normalização pela faixa	72
2° derivada (janela de 7 pontos) + normalização pela faixa	90
2° derivada (janela de 11 pontos) + normalização pela faixa	77

Figura 41 – Funções discriminantes do modelo LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa utilizando todas as variáveis das regiões características dos aditivos.



Fonte: da autora

Tabela 17 – Resultados do modelo LDA dos espectros pré-processados com suavização com filtro de Savitzky-Golay (janela de 3 pontos) e normalização pela faixa utilizando todas as variáveis das regiões características dos aditivos.

Classe verdadeira	N	Classe predita			
		G	T	W	Y
G	48	43	-	4	1
T	49	-	47	-	2
W	45	4	-	41	-
Y	50	-	1	-	49

Tabela 18 – Resultados de previsão obtidos utilizando os quatro modelos de regressão dos aditivos na previsão das amostras dos quatro aditivos.

Amostras do aditivo	Modelo do aditivo	Desvio médio (ppm)	Erro médio de previsão (%)
G	G	66	14
	T	27	31
	W	38	20
	Y	34	37
T	G	62	24
	T	23	9
	W	40	28
	Y	29	24
W	G	63	15
	T	24	40
	W	35	19
	Y	34	42
Y	G	81	36
	T	30	26
	W	51	42
	Y	24	2

Como pode ser visto na tabela, os erros médios de previsão das amostras G e W, quando se usa o modelo incorreto, são no máximo 6 % maiores do que os erros utilizando os modelos corretos. Assim, classificações incorretas entre amostras desses dois aditivos não resultam em erros muito elevados, provavelmente devido a alguma semelhança de composição entre eles. Para os outros aditivos, contudo, a utilização do modelo incorreto resulta em erros percentuais até 20 vezes maiores (amostras do aditivo Y previstas com o modelo do aditivo W).

No que diz respeito ao desvio médio, o qual tende a ser maior quando amostras de composições diferentes das utilizadas na calibração são analisadas, observou-se que, independente da composição das amostras, quando se utilizou os modelos dos aditivos G e W os valores foram sempre os maiores. Isso significa que o modelo

identificou as amostras, mesmo dos seus próprios aditivos, como diferentes das utilizadas na sua construção.

Supondo-se que os aditivos G e W são quimicamente semelhantes, foram construídos modelos de regressão únicos para esses dois aditivos, cujos resultados estão apresentados nos Apêndices W e X. O resultado do melhor modelo, construído com os espectros pré-processados com a 2ª derivada com filtro de Savitzky-Golay (janela de 7 pontos), está apresentado na Tabela 19 juntamente com os resultados dos modelos individuais dos dois aditivos. Aplicando-se o teste F (95 % de confiança), tem-se que o valor de RMSECV do modelo é estatisticamente equivalente aos dos modelos individuais, além disso, o valor do erro médio de previsão é intermediário. Assim, é possível utilizar um único modelo para quantificar esses dois aditivos sem prejuízos nos resultados.

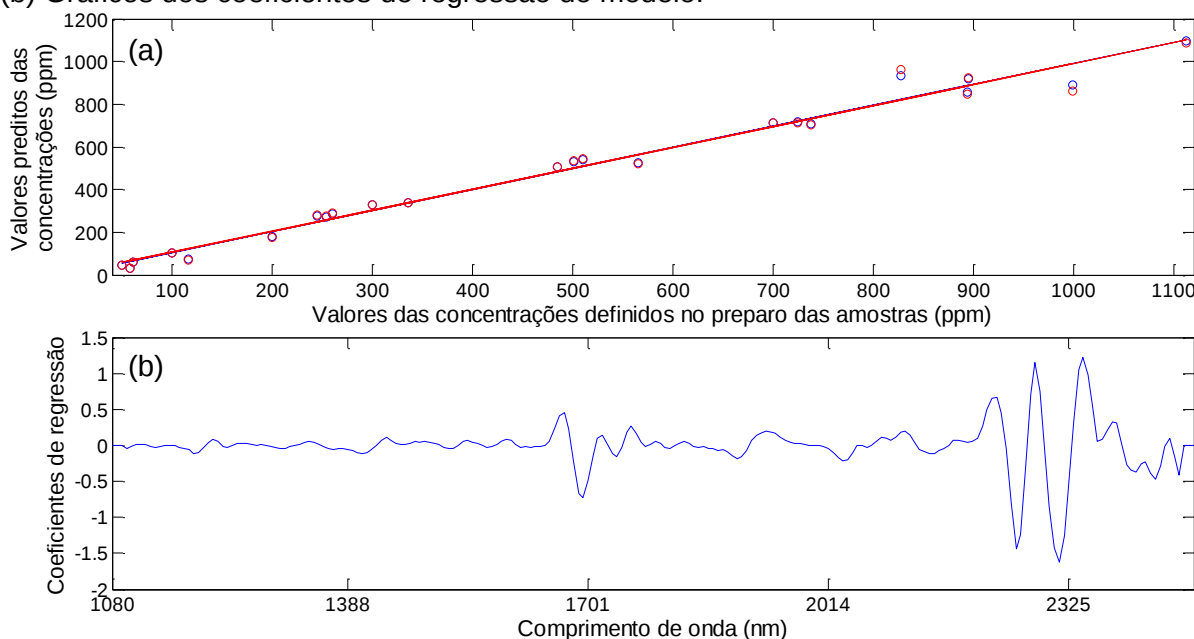
Tabela 19 – Resultados obtidos para os espectros somados subtraídos dos espectros dos *pixels* com valores de escores abaixo do máximo do histograma.

Aditivo	Pré-processamento	VL	RMSECV (ppm)	Erro médio de previsão (%)
G	2º derivada (janela de 11 pontos) + JK	1	60	14
W	1º derivada (janela de 11 ptos)	2	43	19
G e W	2º derivada (janela de 7 pontos)	2	48	15

Os gráficos dos valores previstos em função dos valores das concentrações definidos no preparo das amostras e dos coeficientes de regressão para o melhor modelo estão apresentados na Figura 42.

Através da análise do gráfico dos valores previstos em função dos valores definidos no preparo das amostras (Figura 42a) é possível observar que a curva se ajustou relativamente bem aos pontos, principalmente, para as concentrações mais baixas. No que diz respeito ao gráfico dos coeficientes de regressão do modelo (Figura 42b), observa-se que a região cujas variáveis apresentam maiores coeficientes de regressão está entre 2.232 e 2.344 nm, a qual engloba uma das regiões consideradas características dos aditivos (relacionada às bandas de combinações de vibrações C-H). Esse comportamento dos coeficientes de regressão foi semelhante ao observado nos modelos anteriores, comprovando a importância dessa região.

Figura 42 – (a) Gráficos dos valores preditos das concentrações em função das concentrações definidas no preparo das amostras do modelo conjunto dos aditivos G e W e (b) Gráficos dos coeficientes de regressão do modelo.



Fonte: da autora

Após a definição dos melhores modelos para quantificar os aditivos, o cálculo dos limites de detecção e quantificação foi realizado utilizando o *toolbox* MVC1. Os resultados obtidos estão apresentados na Tabela 20.

Tabela 20 – Limites de detecção e quantificação dos melhores modelos de regressão.

Aditivo	Pré-processamento	LD (ppm)	LQ (ppm)
G	2° derivada (janela de 11 pontos) + JK	22	74
T	1° derivada (janela de 7 ptos) + JK	23	76
W	1° derivada (janela de 11 ptos)	7	25
Y	2° derivada (janela de 11 pontos) + JK	9	28
G e W	2° derivada (janela de 7 pontos)	10	33

O que se pode observar nos resultados é que todos os modelos apresentam valores de limites de detecção inferiores a 25 ppm e valores de limites de quantificação inferiores a 80 ppm. Apesar da concentração mínima dos aditivos não ser conhecida, acredita-se que ela deva ser da ordem de 200 ppm. Assim, uma vez que os limites de quantificação dos modelos estão abaixo desse valor, conclui-se que a metodologia

desenvolvida nesse trabalho pode ser empregada nas análises de rotina dos laboratórios.

6. CONCLUSÃO

A técnica do *ring-oven* permitiu pré-concentrar os aditivos detergentes dispersantes cerca de 78 vezes. Com essa pré-concentração, as concentrações das amostras mais diluídas (50 ppm ou 0,005 %) ficam acima de 0,1 % (m/m), que é, em geral, o limite de detecção admitido para a espectroscopia na região do infravermelho próximo.

Das diferentes estratégias que foram testadas para a extração dos espectros dos anéis, escolheu-se utilizar aquela baseada nos histogramas dos escores da primeira componente pela sua rapidez e não subjetividade. No que diz respeito ao espectro representativo, foi utilizado o espectro somado e, posteriormente, subtraído do mesmo número de espectros médios do papel de forma a evidenciar a informação do aditivo e minimizar a contribuição do sinal referente ao papel. Como foram construídos modelos de calibração PLS individuais para cada aditivo, foi necessário incluir na metodologia uma etapa de classificação a fim de identificar o tipo de aditivo previamente à sua quantificação.

O melhor modelo de classificação, cujo percentual de classificações corretas para o conjunto de previsão foi de 92 %, foi construído com o algoritmo genético associado à análise discriminante linear. Esse modelo utilizou os espectros pré-processados com a suavização com filtro de Savitzky-Golay (polinômio de 2º grau e janela de 7 pontos) seguida da normalização pela faixa. Nesse modelo, das 11 classificações incorretas totais, 10 corresponderam aos aditivos G e W sendo classificados incorretamente (aditivo G sendo classificado como W e vice-versa). Esse resultado pode indicar que esses aditivos são quimicamente semelhantes, o que foi corroborado pelos baixos erros médios de previsão obtidos quando as amostras de um aditivo são preditas utilizando o modelo do outro. Por isso, um modelo de regressão único pôde ser construído para esses dois aditivos.

No que diz respeito à quantificação dos aditivos, os modelos de regressão construídos apresentaram erros médios de previsão entre 2 e 19 % e utilizaram no máximo 2 variáveis latentes.

Um fator limitante desse método é o tempo do preparo do anel (cerca de 35 minutos), contudo, tratando-se de um sistema de baixo custo (entre 1.500 e 2.000 dólares incluindo o computador) e automatizado, é possível que o laboratório de análise possua vários sistemas operando em paralelo. Outro fator limitante,

atualmente, é o custo elevado de uma câmera hiperespectral. Contudo, os anéis uma vez produzidos são estáveis e podem ser facilmente armazenados, sem ocupar muito espaço, e transportados. A logística empregada poderia se constituir das seguintes etapas: os laboratórios localizados em várias regiões do país produziriam os anéis e enviariam os mesmos por correio, por exemplo, para o laboratório no qual a câmera de imagem estivesse instalada. Após aquisição das imagens, as mesmas seriam enviadas de volta aos respectivos laboratórios de origem para tratamento dos dados. É importante observar, contudo, que com o aumento do uso das câmeras hiperespectrais é possível que seu preço diminua. Além disso, o fato das imagens hiperespectrais apresentarem potencial para serem empregadas em diversas análises (alimentos, medicamentos, documentos, etc) pode justificar a sua aquisição pelos laboratórios.

Tem-se, então, uma metodologia promissora para a determinação dos aditivos detergentes dispersantes em gasolina, o qual pode ser facilmente adaptada à rotina dos laboratórios de análise.

É importante observar que essa metodologia não só pode ser aplicada na resolução de outros problemas, sempre que os analitos estejam presentes em baixas concentrações e a análise pelo NIR seja possível, como também outras técnicas analíticas podem ser exploradas em associação com a técnica do *ring-oven*, como foi feito no trabalho de Cortez e Pasquini (2012). No caso específico da determinação dos aditivos, tem-se como perspectiva avaliar a associação da pré-concentração pelo *ring-oven* com a análise por infravermelho médio (refletância total atenuada).

REFERÊNCIAS BIBLIOGRÁFICAS

- AMIGO, J. M. et al. Study of pharmaceutical samples by NIR chemical-image and multivariate analysis. **TrAC Trends in Analytical Chemistry**, v. 27, n. 8, p. 696-713, 2008.
- BRASIL. Portaria MAPA N°105, de 28 de fevereiro de 2013. **Diário Oficial da União**, Brasília, DF, 1 mar. 2013. Seção 1, p. 3, 2013a.
- BRASIL. Resolução ANP N°40, de 25 de outubro de 2013. **Diário Oficial da União**, Brasília, DF, 30 out. 2013. Seção 1, p. 55-56, 2013b.
- BRASIL. Resolução ANP N°1, de 6 de janeiro de 2014. **Diário Oficial da União**, Brasília, DF, 7 jan. 2014. Seção 1, p. 31-33, 2014.
- BRAUER, S. J.; MIIN, T.-C. T. **Method for quantitatively determining detergent fuel additives in fuel samples**. CA 2132806 A1, 23 set. 1994, 25 mar. 1995.
- BRAZ, A.; LÓPEZ-LÓPEZ, M.; GARCÍA-RUIZ, C. Raman spectroscopy for forensic analysis of inks in questioned documents. **Forensic Science International**, v. 232, n. 1-3, p. 206-212, 2013.
- BRERETON, R. G. Introduction to multivariate calibration in analytical chemistry. **Analyst**, v. 125, n. 11, p. 2125-2154, 2000.
- BRERETON, R. G. **Chemometrics for pattern recognition**. 1. ed. Reino Unido: Wiley, 2009. 522 p.
- CHIANG, L. H.; PELL, R. J. Genetic algorithms combined with discriminant analysis for key variable identification. **Journal of Process Control**, v. 14, n. 2, p. 143-155, 2004.
- COLAIOCCO, S.; LATTANZIO, M. Determination of additives in gasoline by SEC coupled to a light-scattering detector—chemometrics application. **Journal of High Resolution Chromatography**, v. 18, n. 6, p. 387-388, 1995.
- CORTEZ, J. **Pré-concentração baseada na técnica de ring-oven para microanálise: determinação simultânea de sódio, ferro e cobre em etanol hidratado combustível por espectroscopia de emissão óptica em plasma induzido por laser (LIBS)**. 2012. 117 p. Tese de doutorado. Instituto de Química, Universidade Estadual de Campinas, Campinas. 2012.
- CORTEZ, J.; PASQUINI, C. Ring-oven based preconcentration technique for microanalysis: Simultaneous determination of Na, Fe, and Cu in fuel ethanol by laser induced breakdown spectroscopy. **Analytical Chemistry**, v. 85, n. 3, p. 1547-1554, 2013.

DA SILVA, M. P. F. et al. Classification of gasoline as with or without dispersant and detergent additives using infrared spectroscopy and multivariate classification. **Fuel**, v. 116, n. 0, p. 151-157, 2014.

DE JUAN, A. et al. Chemometric Tools for Image Analysis. In: SALZER, R. e SIESLER, H. W. **Infrared and Raman Spectroscopic Imaging** 1. ed. Alemanha: Wiley-VCH, 2009. p. 65-109.

DE MEDEIROS, A. R. B. **Uso de ATR/FTIR e FTNIR associado a técnicas quimiométricas para quantificação de aditivos em gasolina automotiva**. 2009. 76 p. Dissertação de Mestrado. Instituto de Química, Universidade de Brasília, Brasília. 2009.

DEROSA, T. F. **Next generation of international chemical additives: A critical review of current US patents**. 1. ed. Oxford: Elsevier, 2012. 578 p.

DOS SANTOS, A. P. F. **Desenvolvimento de métodos para separação e identificação do biodiesel e/ou óleo vegetal no óleo diesel e de aditivos em gasolina**. 2012. 84 p. Dissertação de Mestrado. Escola de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2012.

FEARN, T. et al. On the geometry of SNV and MSC. **Chemometrics and Intelligent Laboratory Systems**, v. 96, n. 1, p. 22-26, 2009.

FEDEROVA, G. V. **Analytical method for the detection and quantification of fuel additives**. US 20040214341 A1, 25 abr. 2003, 28 out. 2004.

FRANCH-LAGE, F. et al. Fast assessment of the surface distribution of API and excipients in tablets using NIR-hyperspectral imaging. **International Journal of Pharmaceutics**, v. 411, n. 1–2, p. 27-35, 2011.

GELADI, P. L. M.; GRAHN, H. F.; BURGUER, J. E. Multivariate Images, Hyperspectral Imaging: Background and Equipment. In: GRAHN, H. F. e GELADI, P. **Techniques and Applications of Hyperspectral Image Analysis** ed. Inglaterra: Wiley, 2007. p. 390.

GRAUPNER, O. et al. **Gasoline additives**. US 20050172545 A1, 14 mar. 2003, 11 ago. 2005.

GUOQING, L.; LILI, W.; MINGXING, F. **Method for separating and measuring cleaning agent in gasoline**. CN 1673737 A, 22 mar. 2004, 28 set. 2005.

HERBSTMAN, S. et al. **Gasoline detergent additive**. US 5030249 A, 1 out. 1990, 9 jul. 1991.

HOLLER, F. J.; SKOOG, D. A.; CROUCH, S. R. **Princípios de Análise Instrumental**. Tradução: PASQUINI, C. et al 6. ed. Porto Alegre: Bookman, 2009. 1056 p.

HONORATO, F. A. et al. Using principal component analysis to find the best calibration settings for simultaneous spectroscopic determination of several gasoline properties. **Fuel**, v. 87, n. 17-18, p. 3706-3709, 2008.

JAUMOT, J. et al. A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB. **Chemometrics and Intelligent Laboratory Systems**, v. 76, n. 1, p. 101-110, 2005.

KARAMAN, İ. et al. Comparison of Sparse and Jack-knife partial least squares regression methods for variable selection. **Chemometrics and Intelligent Laboratory Systems**, v. 122, n. 0, p. 65-77, 2013.

KARPOV, S. A. Improving the environmental and performance properties of automotive gasolines. Detergent additives. **Chemistry and Technology of Fuels and Oils**, v. 43, n. 3, p. 173-178, 2007.

KELEMEN, S. R. et al. **Gasoline additives for reducing the amount of internal combustion engine intake valve deposits and combustion chamber deposits**. US 20030167680 A1, 19 nov. 2002, 11 set. 2003.

LIN, J.-J. et al. Thermal Stability and Combustion Behaviors of Poly(oxybutylene)amides. **Polymer Journal**, v. 34, n. 2, p. 72-80, 2002.

MALPAS, R. E. et al. **Detection system**. US 20080190354 A1, 30 out. 2007, 14 ago. 2008.

MARTENS, H.; MARTENS, M. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). **Food Quality and Preference**, v. 11, n. 1-2, p. 5-16, 2000.

MOREIRA, E. D. T. et al. Near infrared reflectance spectrometry classification of cigarettes using the successive projections algorithm for variable selection. **Talanta**, v. 79, n. 5, p. 1260-1264, 2009.

NIKITINA, E. A. et al. Detergent additives to automotive gasolines. **Chemistry and Technology of Fuels and Oils**, v. 42, n. 1, p. 30-34, 2006.

NOWAK, A. V. **Determining the concentration of additives in petroleum fuels**. US 5244809 A, 29 out. 1992, 14 set. 1993.

OLIVIERI, A. C.; GOICOECHEA, H. C.; IÑÓN, F. A. MVC1: an integrated MatLab toolbox for first-order multivariate calibration. **Chemometrics and Intelligent Laboratory Systems**, v. 73, n. 2, p. 189-197, 2004.

ORONITE, C. Polyisobutylene (PIB)-Based Fuel Additives Disponível em: <https://www.oronite.com/pdfs/polyisobutylene_productstewardshipsummary.pdf>. Acesso em: 15 maio 2014.

PASQUINI, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, v. 14, p. 198-219, 2003.

PIQUERAS, S. et al. Resolution and segmentation of hyperspectral biomedical images by Multivariate Curve Resolution-Alternating Least Squares. **Analytica Chimica Acta**, v. 705, n. 1-2, p. 182-192, 2011.

PONTES, M. J. C. et al. The successive projections algorithm for spectral variable selection in classification problems. **Chemometrics and Intelligent Laboratory Systems**, v. 78, n. 1-2, p. 11-18, 2005.

PRATS-MONTALBÁN, J. M.; DE JUAN, A.; FERRER, A. Multivariate image analysis: A review with applications. **Chemometrics and Intelligent Laboratory Systems**, v. 107, n. 1, p. 1-23, 2011.

READING, K.; MCFARLANE, E. **Test kit and method for detection of additives in fuel compositions**. US 20140038303 A1, 23 jan. 2012, 6 fev. 2014.

RINNAN, Å.; BERG, F. V. D.; ENGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. **TrAC Trends in Analytical Chemistry**, v. 28, n. 10, p. 1201-1222, 2009.

SILIN, M. A. et al. Synthesis and testing of polyalkenyl succinimides as components of detergent additives for motor fuels. **Petroleum Chemistry**, v. 52, n. 4, p. 272-277, 2012.

SILVA, A. C. et al. Detection of adulteration in hydrated ethyl alcohol fuel using infrared spectroscopy and supervised pattern recognition methods. **Talanta**, v. 93, p. 129-134, 2012.

STEVENSON, P. R.; THETFORD, D.; VILARDO, J. S. **Quaternary ammonium salt of a Mannich compound**. US 7906470 B2, 1 set. 2006, 15 mar. 2011.

VIDAL, M.; AMIGO, J. M. Pre-processing of hyperspectral images. Essential steps before image analysis. **Chemometrics and Intelligent Laboratory Systems**, v. 117, n. 0, p. 138-148, 2012.

WEISZ, H. **Microanalysis by the Ring-oven Technique**. 2. ed. Londres: Pergamon Press, 1970. 170 p.

WEISZ, H. Recent applications of the ring-oven technique : A brief review. **Analytica Chimica Acta**, v. 202, n. 0, p. 25-34, 1987.

WORKMAN JR, J.; WEYER, L. **Practical guide and spectral atlas for interpretive near-infrared spectroscopy**. 2. ed. Estados Unidos: CRC Press, 2012. 326 p.

APÊNDICES

APÊNDICE A - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos manualmente dos mapas de concentração da MCR-ALS aumentada.

Aditivo	Espectros representativos	Variáveis utilizadas	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
						Calibração	Validação
G	Espectros médios	Espectro completo	2	100	117	0,9191	0,8972
		Variáveis selecionadas pelo Jack Knife	2	98	114	0,9223	0,9037
		Regiões características dos aditivos	2	98	114	0,9222	0,9028
	Espectros somados	Espectro completo	2	109	135	0,9035	0,8641
		Variáveis selecionadas pelo Jack Knife	1	123	133	0,8778	0,8676
		Regiões características dos aditivos	2	111	133	0,8993	0,8683
T	Espectros médios	Espectro completo	2	81	97	0,9514	0,9362
		Variáveis selecionadas pelo Jack Knife	1	86	97	0,9458	0,9368
		Regiões características dos aditivos	1	91	102	0,9392	0,9299
	Espectros somados	Espectro completo	1	58	64	0,9755	0,9719
		Variáveis selecionadas pelo Jack Knife	1	58	64	0,9755	0,9719
		Regiões características dos aditivos	2	55	63	0,9776	0,9730
W	Espectros médios	Espectro completo	4	59	80	0,9656	0,9415
		Variáveis selecionadas pelo Jack Knife	2	65	77	0,9586	0,9460
		Regiões características dos aditivos	2	69	83	0,9524	0,9371
	Espectros somados	Espectro completo	2	60	70	0,9680	0,9548
		Variáveis selecionadas pelo Jack Knife	2	57	70	0,9682	0,9555
		Regiões características dos aditivos	3	42	60	0,9825	0,9672
Y	Espectros médios	Espectro completo	2	51	63	0,9746	0,9649
		Variáveis selecionadas pelo Jack Knife	1	58	64	0,9676	0,9632
		Regiões características dos aditivos	1	62	69	0,9630	0,9573
	Espectros somados	Espectro completo	5	22	41	0,9954	0,9850
		Variáveis selecionadas pelo Jack Knife	2	28	40	0,9922	0,9857
		Regiões características dos aditivos	1	30	43	0,9914	0,9835

APÊNDICE B - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos manualmente das imagens dos escores da PCA aumentada.

Aditivo	Espectros representativos	Variáveis utilizadas	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
						Calibração	Validação
G	Espectros médios	Espectro completo	2	93	109	0,9292	0,9121
		Variáveis selecionadas pelo Jack Knife	1	99	108	0,9203	0,9124
		Regiões características dos aditivos	1	104	114	0,9115	0,9027
	Espectros somados	Espectro completo	1	99	112	0,9198	0,9069
		Variáveis selecionadas pelo Jack Knife	1	99	112	0,9198	0,9069
		Regiões características dos aditivos	1	91	102	0,9322	0,9220
T	Espectros médios	Espectro completo	2	80	96	0,9528	0,9379
		Variáveis selecionadas pelo Jack Knife	1	81	92	0,9512	0,9422
		Regiões características dos aditivos	1	88	99	0,9426	0,9341
	Espectros somados	Espectro completo	1	48	53	0,9830	0,9810
		Variáveis selecionadas pelo Jack Knife	1	48	53	0,9831	0,9810
		Regiões características dos aditivos	1	46	52	0,9843	0,9819
W	Espectros médios	Espectro completo	2	72	85	0,9488	0,9341
		Variáveis selecionadas pelo Jack Knife	2	71	84	0,9502	0,9359
		Regiões características dos aditivos	2	63	76	0,9601	0,9468
	Espectros somados	Espectro completo	3	49	70	0,9755	0,9542
		Variáveis selecionadas pelo Jack Knife	2	49	65	0,9760	0,9614
		Regiões características dos aditivos	2	49	64	0,9756	0,9627
Y	Espectros médios	Espectro completo	1	61	68	0,9638	0,9591
		Variáveis selecionadas pelo Jack Knife	1	61	68	0,9634	0,9587
		Regiões características dos aditivos	1	57	65	0,9680	0,9626
	Espectros somados	Espectro completo	3	28	45	0,9922	0,9822
		Variáveis selecionadas pelo Jack Knife	4	21	33	0,9955	0,9901
		Regiões características dos aditivos	2	20	29	0,9960	0,9927

APÊNDICE C - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos através dos histogramas dos escores da PCA aumentada.

Aditivo	Espectros representativos	Variáveis utilizadas	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
						Calibração	Validação
G	Espectros médios	Espectro completo	2	88	101	0,9379	0,9233
		Variáveis selecionadas pelo Jack Knife	1	94	103	0,9283	0,9204
		Regiões características dos aditivos	1	98	107	0,9221	0,9144
	Espectros somados	Espectro completo	2	57	69	0,9733	0,9643
		Variáveis selecionadas pelo Jack Knife	1	64	71	0,9668	0,9624
		Regiões características dos aditivos	1	68	75	0,9623	0,9575
T	Espectros médios	Espectro completo	2	70	83	0,9641	0,9536
		Variáveis selecionadas pelo Jack Knife	2	69	82	0,9653	0,9550
		Regiões características dos aditivos	1	77	86	0,9560	0,9495
	Espectros somados	Espectro completo	3	37	59	0,9898	0,9766
		Variáveis selecionadas pelo Jack Knife	1	53	58	0,9796	0,9772
		Regiões características dos aditivos	1	61	67	0,9726	0,9693
W	Espectros médios	Espectro completo	3	60	86	0,9645	0,9327
		Variáveis selecionadas pelo Jack Knife	2	76	98	0,9433	0,9118
		Regiões características dos aditivos	4	59	79	0,9656	0,9432
	Espectros somados	Espectro completo	2	53	63	0,9721	0,9632
		Variáveis selecionadas pelo Jack Knife	2	53	63	0,9721	0,9633
		Regiões características dos aditivos	4	20	28	0,9958	0,9929
Y	Espectros médios	Espectro completo	2	38	44	0,9864	0,9824
		Variáveis selecionadas pelo Jack Knife	2	37	44	0,9866	0,9827
		Regiões características dos aditivos	1	42	46	0,9829	0,9815
	Espectros somados	Espectro completo	5	22	41	0,9952	0,9849
		Variáveis selecionadas pelo Jack Knife	2	32	44	0,9898	0,9826
		Regiões características dos aditivos	2	33	47	0,9892	0,9802

APÊNDICE D - Modelos de regressão construídos com os espectros representativos da região de interesse extraídos através dos histogramas dos escores da PCA individual de cada imagem.

Aditivo	Espectros representativos	Variáveis utilizadas	VL	RMSEC (ppm)	RMSECV (ppm)	R ²	
						Calibração	Validação
G	Espectros médios	Espectro completo	2	85	99	0,9416	0,9267
		Variáveis selecionadas pelo Jack Knife	2	80	97	0,9479	0,9295
		Regiões características dos aditivos	5	55	95	0,9753	0,9324
	Espectros somados	Espectro completo	2	62	77	0,9693	0,9561
		Variáveis selecionadas pelo Jack Knife	1	71	79	0,9594	0,9534
		Regiões características dos aditivos	1	72	81	0,9575	0,9516
T	Espectros médios	Espectro completo	3	66	87	0,9680	0,9484
		Variáveis selecionadas pelo Jack Knife	1	75	84	0,9583	0,9517
		Regiões características dos aditivos	1	86	96	0,9453	0,9372
	Espectros somados	Espectro completo	1	49	54	0,9821	0,9802
		Variáveis selecionadas pelo Jack Knife	1	49	54	0,9821	0,9803
		Regiões características dos aditivos	1	41	46	0,9879	0,9855
W	Espectros médios	Espectro completo	4	56	78	0,9691	0,9444
		Variáveis selecionadas pelo Jack Knife	1	83	92	0,9323	0,9230
		Regiões características dos aditivos	2	68	80	0,9542	0,9418
	Espectros somados	Espectro completo	2	51	60	0,9743	0,9671
		Variáveis selecionadas pelo Jack Knife	4	25	37	0,9940	0,9875
		Regiões características dos aditivos	2	35	41	0,9879	0,9850
Y	Espectros médios	Espectro completo	2	48	58	0,9774	0,9703
		Variáveis selecionadas pelo Jack Knife	2	48	57	0,9780	0,9709
		Regiões características dos aditivos	1	56	62	0,9700	0,9656
	Espectros somados	Espectro completo	3	42	59	0,9831	0,9684
		Variáveis selecionadas pelo Jack Knife	2	55	65	0,9701	0,9625
		Regiões características dos aditivos	2	31	40	0,9908	0,9857

APÊNDICE E - Modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	VL	RMSEC (ppm)	RMSECV (ppm)	R ²		Erro médio de previsão (%)
				Calibração	Validação	
Suavização (janela de 3 pontos)	3	28	62	0,9943	0,9770	22
Suavização (janela de 3 pontos) com JK	1	58	73	0,9758	0,9678	59
Suavização (janela de 3 pontos) (regiões dos aditivos)	3	20	32	0,9973	0,9937	22
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	3	22	36	0,9967	0,9924	48
Suavização (janela de 7 pontos)	3	28	61	0,9945	0,9775	21
Suavização (janela de 7 pontos) com JK	3	20	33	0,9973	0,9934	13
Suavização (janela de 7 pontos) (regiões dos aditivos)	3	22	34	0,9967	0,9929	24
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	2	30	55	0,9937	0,9819	32
1º derivada (janela de 7 pontos)	3	23	39	0,9962	0,9909	17
1º derivada (janela de 7 pontos) com JK	3	24	36	0,9960	0,9922	16
1º derivada (janela de 7 pontos) (regiões dos aditivos)	3	15	23	0,9985	0,9969	15
1º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	16	41	0,9981	0,9897	19
1º derivada (janela de 11 pontos)	3	24	44	0,9959	0,9886	19
1º derivada (janela de 11 pontos) com JK	3	28	44	0,9946	0,9885	20
1º derivada (janela de 11 pontos) (regiões dos aditivos)	3	19	30	0,9975	0,9947	19
1º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	3	19	29	0,9975	0,9949	19
2º derivada (janela de 7 pontos)	3	14	27	0,9986	0,9956	17
2º derivada (janela de 7 pontos) com JK	2	12	36	0,9989	0,9924	16
2º derivada (janela de 7 pontos) (regiões dos aditivos)	3	11	22	0,9991	0,9971	16
2º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	12	42	0,9990	0,9892	16
2º derivada (janela de 11 pontos)	3	20	32	0,9971	0,9937	15
2º derivada (janela de 11 pontos) com JK	3	19	28	0,9974	0,9952	15
2º derivada (janela de 11 pontos) (regiões dos aditivos)	3	15	24	0,9983	0,9966	14
2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	3	15	23	0,9983	0,9969	13

APÊNDICE F - Resultados de previsão utilizando os modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	Valor			(continua) Erro de previsão (%)
	Predito	Desvio	Referência	
Suavização (janela de 3 pontos)	217,0	36,9	165,6	31
	60,4	48,1	83,2	27
	221,2	78,4	274,8	20
	430,6	64,1	357,6	20
	551,2	92,6	606,8	9
Suavização (janela de 3 pontos) com JK	174,4	54,8	165,6	5
	-72,5	123,6	83,2	187
	149,4	68,6	274,8	46
	281,1	70,5	357,6	21
	400,0	92,3	606,8	34
Suavização (janela de 3 pontos) (regiões dos aditivos)	201,8	25,4	165,6	22
	62,4	22,8	83,2	25
	322,1	38,8	274,8	17
	470,1	41,4	357,6	31
	699,6	57,1	606,8	15
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	171,3	43,6	165,6	3
	-82,6	87,9	83,2	199
	211,6	103,9	274,8	23
	364,4	85,1	357,6	2
	537,7	157,3	606,8	11
Suavização (janela de 7 pontos)	219,1	37,5	165,6	32
	61,0	48,8	83,2	27
	229,2	87,2	274,8	17
	444,1	68,2	357,6	24
	562,5	102,6	606,8	7
Suavização (janela de 7 pontos) com JK	192,6	16,7	165,6	16
	80,7	34,3	83,2	3
	304,5	63,8	274,8	11
	448,1	57,7	357,6	25
	669,3	89,7	606,8	10
Suavização (janela de 7 pontos) (regiões dos aditivos)	203,7	28,4	165,6	23
	62,0	24,0	83,2	25
	330,3	42,6	274,8	20
	474,1	47,0	357,6	33
	709,3	62,5	606,8	17
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	208,3	28,2	165,6	26
	27,6	56,5	83,2	67
	328,6	53,7	274,8	20
	475,5	53,4	357,6	33
	693,8	89,6	606,8	14
1° derivada (janela de 7 pontos)	195,4	26,9	165,6	18
	55,4	29,9	83,2	33
	246,1	44,5	274,8	10
	424,4	35,1	357,6	19
	580,7	67,2	606,8	4
1° derivada (janela de 7 pontos) com JK	201,9	26,9	165,6	22
	60,2	25,5	83,2	28
	246,6	49,9	274,8	10
	420,1	41,5	357,6	17
	580,9	82,4	606,8	4
	180,3	12,9	165,6	9

1° derivada (janela de 7 pontos) (regiões dos aditivos)	62,1	14,8	83,2	25
	293,2	16,9	274,8	7
	440,8	18,6	357,6	23
	678,3	27,0	606,8	12
1° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	186,6	19,3	165,6	13
	55,6	29,6	83,2	33
	299,6	27,5	274,8	9
	448,2	33,2	357,6	25
	698,1	44,4	606,8	15
1° derivada (janela de 11 pontos)	210,5	26,6	165,6	27
	55,3	33,9	83,2	34
	250,6	55,7	274,8	9
	438,9	40,6	357,6	23
	593,2	81,7	606,8	2
1° derivada (janela de 11 pontos) com JK	214,4	32,2	165,6	29
	54,6	32,2	83,2	34
	243,4	63,4	274,8	11
	423,7	54,9	357,6	18
	574,2	102,2	606,8	5
1° derivada (janela de 11 pontos) (regiões dos aditivos)	192,4	22,4	165,6	16
	62,2	19,9	83,2	25
	312,2	34,5	274,8	14
	450,3	38,8	357,6	26
	703,3	51,7	606,8	16
1° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	193,1	18,5	165,6	17
	61,8	19,4	83,2	26
	311,8	35,1	274,8	13
	448,6	39,8	357,6	25
	703,4	50,8	606,8	16
2° derivada (janela de 7 pontos)	188,2	23,1	165,6	14
	53,5	19,9	83,2	36
	244,5	20,2	274,8	11
	437,1	19,8	357,6	22
	599,0	31,1	606,8	1
2° derivada (janela de 7 pontos) com JK	184,6	21,2	165,6	11
	54,3	24,6	83,2	35
	252,1	22,1	274,8	8
	438,0	28,6	357,6	23
	610,8	36,7	606,8	1
2° derivada (janela de 7 pontos) (regiões dos aditivos)	182,2	16,4	165,6	10
	55,2	16,0	83,2	34
	267,3	15,2	274,8	3
	455,4	17,2	357,6	27
	631,3	26,3	606,8	4
2° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	182,5	21,4	165,6	10
	54,2	28,7	83,2	35
	267,2	24,9	274,8	3
	452,3	33,6	357,6	27
	628,6	43,4	606,8	4
2° derivada (janela de 11 pontos)	187,3	25,2	165,6	13
	57,2	24,4	83,2	31
	245,0	30,3	274,8	11
	418,9	27,3	357,6	17
	582,1	51,5	606,8	4
2° derivada (janela de 11 pontos) com JK	191,6	23,9	165,6	16
	57,9	20,0	83,2	30
	250,4	33,1	274,8	9
	420,4	27,8	357,6	18
	591,5	59,7	606,8	3

2° derivada (janela de 11 pontos) (regiões dos aditivos)	171,9	19,2	165,6	4
	59,6	16,5	83,2	28
	288,9	20,8	274,8	5
	434,9	24,0	357,6	22
	659,4	35,8	606,8	9
2° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	171,9	18,4	165,6	4
	60,0	16,0	83,2	28
	288,7	23,2	274,8	5
	433,6	26,7	357,6	21
	663,6	38,3	606,8	9

APÊNDICE G - Modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	VL	RMSEC (ppm)	RMSECV (ppm)	R ²		Erro médio de previsão (%)
				Calibração	Validação	
Suavização (janela de 3 pontos)	3	18	45	0,9977	0,9879	12
Suavização (janela de 3 pontos) com JK	2	32	49	0,9927	0,9857	14
Suavização (janela de 3 pontos) (regiões dos aditivos)	3	26	53	0,9952	0,9831	20
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	3	26	39	0,9952	0,9911	10
Suavização (janela de 7 pontos)	3	19	46	0,9974	0,9871	13
Suavização (janela de 7 pontos) com JK	2	30	45	0,9935	0,9877	15
Suavização (janela de 7 pontos) (regiões dos aditivos)	3	28	57	0,9944	0,9807	21
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	3	26	39	0,9950	0,9909	10
1º derivada (janela de 7 pontos)	2	17	48	0,9978	0,9863	10
1º derivada (janela de 7 pontos) com JK	2	12	28	0,9989	0,9953	9
1º derivada (janela de 7 pontos) (regiões dos aditivos)	2	24	46	0,9958	0,9870	15
1º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	19	35	0,9974	0,9927	16
1º derivada (janela de 11 pontos)	3	19	44	0,9975	0,9883	14
1º derivada (janela de 11 pontos) com JK	2	17	43	0,9980	0,9889	11
1º derivada (janela de 11 pontos) (regiões dos aditivos)	2	37	65	0,9901	0,9748	15
1º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	3	27	50	0,9949	0,9850	16
2º derivada (janela de 7 pontos)	2	19	39	0,9975	0,9909	13
2º derivada (janela de 7 pontos) com JK	1	28	35	0,9945	0,9926	18
2º derivada (janela de 7 pontos) (regiões dos aditivos)	2	22	43	0,9964	0,9891	13
2º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	1	26	33	0,9952	0,9933	13
2º derivada (janela de 11 pontos)	2	19	40	0,9974	0,9904	14
2º derivada (janela de 11 pontos) com JK	1	37	47	0,9902	0,9869	24
2º derivada (janela de 11 pontos) (regiões dos aditivos)	2	25	42	0,9957	0,9894	16
2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	31	38	0,9932	0,9912	19

APÊNDICE H - Resultados de previsão utilizando os modelos de regressão do aditivo T construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	Valor			(continua) Erro de previsão (%)
	Predito	Desvio	Referência	
Suavização (janela de 3 pontos)	202,1	25,1	165,6	22
	75,0	37,2	83,2	10
	283,9	56,3	274,8	3
	436,3	34,8	357,6	22
	599,0	59,7	606,8	1
Suavização (janela de 3 pontos) com JK	197,5	23,7	165,6	19
	97,4	42,8	83,2	17
	273,1	69,9	274,8	1
	425,2	48,9	357,6	19
	527,7	51,8	606,8	13
Suavização (janela de 3 pontos) (regiões dos aditivos)	208,7	28,4	165,6	26
	59,2	41,9	83,2	29
	315,5	46,9	274,8	15
	433,5	40,4	357,6	21
	664,9	67,7	606,8	10
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	174,6	25,1	165,6	5
	61,8	30,1	83,2	26
	284,2	37,7	274,8	3
	410,3	30,5	357,6	15
	599,1	42,2	606,8	1
Suavização (janela de 7 pontos)	204,9	25,3	165,6	24
	72,2	38,3	83,2	13
	286,3	57,8	274,8	4
	436,3	36,1	357,6	22
	605,2	62,2	606,8	0
Suavização (janela de 7 pontos) com JK	198,3	23,5	165,6	20
	100,2	40,1	83,2	20
	281,2	69,9	274,8	2
	433,2	44,5	357,6	21
	532,7	51,2	606,8	12
Suavização (janela de 7 pontos) (regiões dos aditivos)	211,4	30,4	165,6	28
	57,5	44,8	83,2	31
	318,4	49,9	274,8	16
	433,1	43,9	357,6	21
	667,0	72,1	606,8	10
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	180,9	22,3	165,6	9
	66,9	26,8	83,2	20
	293,3	42,2	274,8	7
	414,3	35,1	357,6	16
	608,2	41,0	606,8	0
1° derivada (janela de 7 pontos)	189,7	22,0	165,6	15
	69,1	37,3	83,2	17
	270,5	35,3	274,8	2
	407,9	29,5	357,6	14
	611,2	48,5	606,8	1
1° derivada (janela de 7 pontos) com JK	188,1	12,4	165,6	14
	73,3	23,2	83,2	12
	277,5	25,2	274,8	1
	429,3	19,7	357,6	20
	611,5	35,8	606,8	1

1° derivada (janela de 7 pontos) (regiões dos aditivos)	190,1	20,9	165,6	15
	57,2	33,0	83,2	31
	293,8	29,9	274,8	7
	410,8	29,4	357,6	15
	650,0	43,6	606,8	7
1° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	185,8	15,8	165,6	12
	56,6	26,4	83,2	32
	291,6	23,8	274,8	6
	424,8	25,3	357,6	19
	660,9	32,8	606,8	9
1° derivada (janela de 11 pontos)	193,5	20,5	165,6	17
	55,6	30,8	83,2	33
	269,6	41,4	274,8	2
	418,6	26,3	357,6	17
	609,8	54,5	606,8	1
1° derivada (janela de 11 pontos) com JK	183,6	19,1	165,6	11
	65,7	34,0	83,2	21
	267,8	38,7	274,8	3
	421,4	32,1	357,6	18
	597,3	47,8	606,8	2
1° derivada (janela de 11 pontos) (regiões dos aditivos)	199,1	30,6	165,6	20
	54,5	44,2	83,2	35
	286,4	53,2	274,8	4
	396,0	51,4	357,6	11
	631,0	76,1	606,8	4
1° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	180,8	35,2	165,6	9
	52,3	35,3	83,2	37
	301,6	55,4	274,8	10
	415,9	42,5	357,6	16
	645,2	67,7	606,8	6
2° derivada (janela de 7 pontos)	179,2	18,3	165,6	8
	53,6	30,5	83,2	36
	269,6	21,8	274,8	2
	411,9	22,3	357,6	15
	618,8	34,8	606,8	2
2° derivada (janela de 7 pontos) com JK	176,9	17,2	165,6	7
	31,6	31,7	83,2	62
	250,6	26,1	274,8	9
	386,9	30,9	357,6	8
	573,1	51,4	606,8	6
2° derivada (janela de 7 pontos) (regiões dos aditivos)	177,7	19,0	165,6	7
	53,3	32,1	83,2	36
	281,2	22,1	274,8	2
	416,2	24,1	357,6	16
	631,4	36,4	606,8	4
2° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	174,9	16,0	165,6	6
	47,8	34,3	83,2	43
	256,8	27,8	274,8	7
	386,2	31,1	357,6	8
	585,8	59,8	606,8	3
2° derivada (janela de 11 pontos)	183,1	18,3	165,6	11
	47,8	32,5	83,2	43
	271,4	25,5	274,8	1
	409,8	22,0	357,6	15
	609,7	38,6	606,8	0
2° derivada (janela de 11 pontos) com JK	186,8	21,9	165,6	13
	14,9	36,8	83,2	82
	242,0	41,7	274,8	12
	369,7	40,3	357,6	3

	541,9	80,1	606,8	11
2° derivada (janela de 11 pontos) (regiões dos aditivos)	180,8	18,1	165,6	9
	46,8	32,3	83,2	44
	289,2	23,6	274,8	5
	416,8	22,3	357,6	17
	628,1	38,3	606,8	4
2° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	184,8	17,2	165,6	12
	29,8	38,1	83,2	64
	265,4	39,3	274,8	3
	388,5	31,8	357,6	9
	570,8	79,6	606,8	6

APÊNDICE I - Modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	VL	RMSEC (ppm)	RMSECV (ppm)	R ²		Erro médio de previsão (%)
				Calibração	Validação	
Suavização (janela de 3 pontos)	1	47	55	0,9769	0,9736	30
Suavização (janela de 3 pontos) com JK	1	47	55	0,9769	0,9736	30
Suavização (janela de 3 pontos) (regiões dos aditivos)	1	47	55	0,9769	0,9742	32
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	1	47	55	0,9769	0,9742	32
Suavização (janela de 7 pontos)	1	47	55	0,9773	0,9739	31
Suavização (janela de 7 pontos) com JK	1	47	55	0,9772	0,9739	31
Suavização (janela de 7 pontos) (regiões dos aditivos)	1	47	55	0,9770	0,9743	32
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	1	47	55	0,9770	0,9743	32
1º derivada (janela de 7 pontos)	1	46	54	0,9782	0,9745	26
1º derivada (janela de 7 pontos) com JK	1	46	54	0,9781	0,9745	26
1º derivada (janela de 7 pontos) (regiões dos aditivos)	1	45	54	0,9789	0,9747	27
1º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	1	45	54	0,9789	0,9748	27
1º derivada (janela de 11 pontos)	1	45	53	0,9788	0,9758	27
1º derivada (janela de 11 pontos) com JK	1	45	53	0,9787	0,9758	27
1º derivada (janela de 11 pontos) (regiões dos aditivos)	1	46	54	0,9782	0,9748	29
1º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	46	54	0,9781	0,9747	29
2º derivada (janela de 7 pontos)	1	52	65	0,9719	0,9633	26
2º derivada (janela de 7 pontos) com JK	1	52	64	0,9722	0,9646	25
2º derivada (janela de 7 pontos) (regiões dos aditivos)	1	55	68	0,9689	0,9599	27
2º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	1	54	67	0,9695	0,9613	26
2º derivada (janela de 11 pontos)	1	42	52	0,9815	0,9767	23
2º derivada (janela de 11 pontos) com JK	1	42	52	0,9814	0,9770	23
2º derivada (janela de 11 pontos) (regiões dos aditivos)	1	44	54	0,9802	0,9752	23
2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	44	53	0,9801	0,9754	23

APÊNDICE J - Resultados de previsão utilizando os modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel das bordas da imagem (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	Valor			(continua) Erro de previsão (%)
	Predito	Desvio	Referência	
Suavização (janela de 3 pontos)	234,0	39,6	150,1	56
	60,7	36,8	99,6	39
	137,2	41,1	167,0	18
	344,2	61,4	413,9	17
	418,7	47,6	539,5	22
Suavização (janela de 3 pontos) com JK	234,2	40,2	150,1	56
	60,6	36,7	99,6	39
	137,3	40,9	167,0	18
	344,3	62,2	413,9	17
	418,6	49,1	539,5	22
Suavização (janela de 3 pontos) (regiões dos aditivos)	231,9	49,6	150,1	54
	52,1	33,4	99,6	48
	131,6	32,1	167,0	21
	346,0	76,5	413,9	16
	422,1	70,7	539,5	22
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	231,9	49,6	150,1	54
	52,1	33,4	99,6	48
	131,6	32,1	167,0	21
	346,0	76,5	413,9	16
	422,1	70,7	539,5	22
Suavização (janela de 7 pontos)	234,3	40,0	150,1	56
	59,3	36,3	99,6	40
	135,9	40,3	167,0	19
	341,1	61,4	413,9	18
	418,8	48,0	539,5	22
Suavização (janela de 7 pontos) com JK	234,6	40,8	150,1	56
	59,2	36,1	99,6	41
	136,0	39,9	167,0	19
	341,2	62,4	413,9	18
	418,6	49,7	539,5	22
Suavização (janela de 7 pontos) (regiões dos aditivos)	231,3	49,5	150,1	54
	52,0	33,4	99,6	48
	131,7	32,1	167,0	21
	346,4	75,8	413,9	16
	422,6	70,4	539,5	22
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	231,3	49,5	150,1	54
	52,0	33,4	99,6	48
	131,7	32,1	167,0	21
	346,4	75,8	413,9	16
	422,6	70,4	539,5	22
1° derivada (janela de 7 pontos)	208,8	43,2	150,1	39
	54,4	35,4	99,6	45
	138,5	36,2	167,0	17
	365,5	53,0	413,9	12
	440,3	55,1	539,5	18
1° derivada (janela de 7 pontos) com JK	208,8	44,1	150,1	39
	54,3	35,5	99,6	45
	138,2	35,6	167,0	17
	365,4	54,9	413,9	12
	440,3	57,1	539,5	18
	201,3	46,6	150,1	34

1° derivada (janela de 7 pontos) (regiões dos aditivos)	43,5	33,0	99,6	56
	140,7	33,8	167,0	16
	370,5	60,9	413,9	10
	451,3	63,9	539,5	16
1° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	201,2	47,5	150,1	34
	43,5	32,8	99,6	56
	140,6	33,5	167,0	16
	370,6	63,3	413,9	10
	451,4	66,2	539,5	16
1° derivada (janela de 11 pontos)	213,0	44,5	150,1	42
	54,4	35,1	99,6	45
	136,6	34,2	167,0	18
	361,2	55,6	413,9	13
	438,4	54,0	539,5	19
1° derivada (janela de 11 pontos) com JK	213,2	45,2	150,1	42
	54,4	35,2	99,6	45
	136,6	33,8	167,0	18
	361,0	56,5	413,9	13
	438,2	55,1	539,5	19
1° derivada (janela de 11 pontos) (regiões dos aditivos)	206,4	52,6	150,1	38
	40,9	31,7	99,6	59
	138,1	30,6	167,0	17
	364,3	70,0	413,9	12
	448,2	69,0	539,5	17
1° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	206,7	52,8	150,1	38
	41,0	31,8	99,6	59
	138,1	30,6	167,0	17
	364,0	71,2	413,9	12
	447,9	70,2	539,5	17
2° derivada (janela de 7 pontos)	204,8	48,8	150,1	36
	47,5	44,1	99,6	52
	148,6	48,7	167,0	11
	366,0	57,4	413,9	12
	434,1	67,0	539,5	20
2° derivada (janela de 7 pontos) com JK	202,8	50,2	150,1	35
	48,7	43,2	99,6	51
	150,4	47,9	167,0	10
	368,9	56,8	413,9	11
	437,1	69,5	539,5	19
2° derivada (janela de 7 pontos) (regiões dos aditivos)	208,9	52,2	150,1	39
	48,2	45,3	99,6	52
	150,6	50,6	167,0	10
	363,4	65,9	413,9	12
	427,9	77,3	539,5	21
2° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	207,1	54,5	150,1	38
	49,5	44,5	99,6	50
	152,3	50,0	167,0	9
	366,6	65,4	413,9	11
	431,0	80,7	539,5	20
2° derivada (janela de 11 pontos)	191,9	37,8	150,1	28
	48,8	32,8	99,6	51
	140,8	35,3	167,0	16
	382,5	45,7	413,9	8
	459,8	50,6	539,5	15
2° derivada (janela de 11 pontos) com JK	191,9	38,5	150,1	28
	48,7	32,4	99,6	51
	140,6	34,9	167,0	16
	382,4	46,9	413,9	8
	459,8	52,2	539,5	15

2° derivada (janela de 11 pontos) (regiões dos aditivos)	193,7	39,9	150,1	29
	50,1	32,5	99,6	50
	143,5	34,5	167,0	14
	382,8	52,0	413,9	8
	457,6	58,6	539,5	15
2° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	193,8	40,8	150,1	29
	49,9	32,4	99,6	50
	143,2	34,0	167,0	14
	382,5	52,8	413,9	8
	457,5	59,8	539,5	15

APÊNDICE K - Modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	VL	RMSEC (ppm)	RMSECV (ppm)	R ²		Erro médio de previsão (%)
				Calibração	Validação	
Suavização (janela de 3 pontos)	2	39	54	0,9838	0,9745	19
Suavização (janela de 3 pontos) com JK	1	50	60	0,9744	0,9689	32
Suavização (janela de 3 pontos) (regiões dos aditivos)	2	34	43	0,9880	0,9823	22
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	2	33	45	0,9883	0,9828	23
Suavização (janela de 7 pontos)	2	40	54	0,9837	0,9744	19
Suavização (janela de 7 pontos) com JK	1	51	62	0,9728	0,9671	32
Suavização (janela de 7 pontos) (regiões dos aditivos)	2	35	46	0,9871	0,9814	23
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	2	35	46	0,9875	0,9820	23
1º derivada (janela de 7 pontos)	1	45	55	0,9790	0,9742	29
1º derivada (janela de 7 pontos) com JK	1	45	55	0,9789	0,9742	29
1º derivada (janela de 7 pontos) (regiões dos aditivos)	1	45	55	0,9785	0,9741	33
1º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	1	45	55	0,9784	0,9741	33
1º derivada (janela de 11 pontos)	2	30	43	0,9905	0,9844	19
1º derivada (janela de 11 pontos) com JK	1	36	45	0,9862	0,9828	26
1º derivada (janela de 11 pontos) (regiões dos aditivos)	2	32	44	0,9895	0,9830	24
1º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	37	45	0,9856	0,9825	30
2º derivada (janela de 7 pontos)	1	41	52	0,9821	0,9770	26
2º derivada (janela de 7 pontos) com JK	1	42	52	0,9820	0,9770	26
2º derivada (janela de 7 pontos) (regiões dos aditivos)	1	44	54	0,9802	0,9749	28
2º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	1	44	54	0,9802	0,9749	28
2º derivada (janela de 11 pontos)	1	37	46	0,9859	0,9817	23
2º derivada (janela de 11 pontos) com JK	1	36	46	0,9861	0,9820	23
2º derivada (janela de 11 pontos) (regiões dos aditivos)	1	38	48	0,9849	0,9805	25
2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	38	47	0,9851	0,9808	24

APÊNDICE L - Resultados de previsão utilizando os modelos de regressão do aditivo W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	Valor			(continua) Erro de previsão (%)
	Predito	Desvio	Referência	
Suavização (janela de 3 pontos)	215,2	50,2	150,1	43
	68,7	59,5	99,6	31
	148,3	34,2	167,0	11
	415,2	41,8	413,9	0
	496,2	43,6	539,5	8
Suavização (janela de 3 pontos) com JK	273,9	78,0	150,1	82
	78,8	45,9	99,6	21
	129,6	43,8	167,0	22
	356,8	80,8	413,9	14
	436,3	81,6	539,5	19
Suavização (janela de 3 pontos) (regiões dos aditivos)	194,2	40,5	150,1	29
	44,2	36,4	99,6	56
	138,9	31,3	167,0	17
	426,5	37,4	413,9	3
	508,7	42,0	539,5	6
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	193,7	42,1	150,1	29
	43,0	37,3	99,6	57
	136,9	29,5	167,0	18
	426,2	35,7	413,9	3
	508,0	44,8	539,5	6
Suavização (janela de 7 pontos)	216,1	50,3	150,1	44
	70,4	59,9	99,6	29
	148,0	34,1	167,0	11
	414,9	42,3	413,9	0
	496,8	43,2	539,5	8
Suavização (janela de 7 pontos) com JK	278,4	80,7	150,1	85
	82,3	47,7	99,6	17
	128,1	44,7	167,0	23
	351,7	83,9	413,9	15
	432,4	84,3	539,5	20
Suavização (janela de 7 pontos) (regiões dos aditivos)	196,9	42,1	150,1	31
	44,4	36,8	99,6	55
	138,6	33,1	167,0	17
	426,4	39,6	413,9	3
	507,0	44,6	539,5	6
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	196,6	43,6	150,1	31
	43,7	39,9	99,6	56
	136,5	31,6	167,0	18
	425,7	38,8	413,9	3
	506,1	48,8	539,5	6
1° derivada (janela de 7 pontos)	252,6	76,6	150,1	68
	68,4	51,2	99,6	31
	134,3	31,7	167,0	20
	367,9	62,1	413,9	11
	450,4	70,8	539,5	17
	252,8	77,3	150,1	68

1° derivada (janela de 7 pontos) com JK	68,4	52,2	99,6	31
	134,3	31,6	167,0	20
	367,8	62,9	413,9	11
	450,4	71,2	539,5	17
1° derivada (janela de 7 pontos) (regiões dos aditivos)	253,5	83,6	150,1	69
	49,4	41,1	99,6	50
	135,0	31,1	167,0	19
	370,0	66,5	413,9	11
	454,8	75,4	539,5	16
1° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	253,7	84,3	150,1	69
	49,5	41,4	99,6	50
	135,1	31,3	167,0	19
	369,9	67,8	413,9	11
	454,7	76,1	539,5	16
1° derivada (janela de 11 pontos)	179,2	37,9	150,1	19
	45,3	50,8	99,6	54
	140,9	24,5	167,0	16
	427,0	28,1	413,9	3
	519,0	34,3	539,5	4
1° derivada (janela de 11 pontos) com JK	225,6	62,0	150,1	50
	60,0	46,0	99,6	40
	133,9	26,4	167,0	20
	389,9	52,0	413,9	6
	474,7	60,2	539,5	12
1° derivada (janela de 11 pontos) (regiões dos aditivos)	177,4	41,0	150,1	18
	20,6	36,3	99,6	79
	139,6	26,0	167,0	16
	432,0	29,3	413,9	4
	525,4	31,7	539,5	3
1° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	221,7	69,0	150,1	48
	33,5	34,5	99,6	66
	132,9	26,4	167,0	20
	397,0	58,0	413,9	4
	484,2	64,6	539,5	10
2° derivada (janela de 7 pontos)	226,0	61,2	150,1	51
	55,5	40,6	99,6	44
	145,4	32,7	167,0	13
	373,6	56,4	413,9	10
	463,1	64,2	539,5	14
2° derivada (janela de 7 pontos) com JK	226,4	61,5	150,1	51
	55,6	39,6	99,6	44
	145,4	32,7	167,0	13
	373,2	56,5	413,9	10
	462,6	64,1	539,5	14
2° derivada (janela de 7 pontos) (regiões dos aditivos)	231,7	66,1	150,1	54
	51,0	37,8	99,6	49
	146,6	34,2	167,0	12
	369,5	60,6	413,9	11
	457,0	69,6	539,5	15
2° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	231,6	66,1	150,1	54
	50,9	37,1	99,6	49
	146,6	34,1	167,0	12
	369,6	60,6	413,9	11
	457,1	69,6	539,5	15

2° derivada (janela de 11 pontos)	219,9	57,1	150,1	46
	61,3	39,4	99,6	38
	143,4	26,5	167,0	14
	389,3	48,5	413,9	6
	481,8	52,8	539,5	11
2° derivada (janela de 11 pontos) com JK	218,8	57,4	150,1	46
	61,1	39,9	99,6	39
	143,7	26,1	167,0	14
	390,2	48,7	413,9	6
	482,7	53,2	539,5	11
2° derivada (janela de 11 pontos) (regiões dos aditivos)	225,7	62,4	150,1	50
	56,4	31,7	99,6	43
	146,3	27,4	167,0	12
	388,1	52,9	413,9	6
	479,3	58,9	539,5	11
2° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	224,4	62,9	150,1	50
	56,3	31,5	99,6	43
	146,7	26,8	167,0	12
	389,2	53,3	413,9	6
	480,4	59,7	539,5	11

APÊNDICE M - Modelos de regressão do aditivo G construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	VL	RMSEC (ppm)	RMSECV (ppm)	R ²		Erro médio de previsão (%)
				Calibração	Validação	
Suavização (janela de 3 pontos)	2	41	60	0,9851	0,9734	17
Suavização (janela de 3 pontos) com JK	2	41	59	0,9853	0,9739	17
Suavização (janela de 3 pontos) (regiões dos aditivos)	2	46	68	0,9812	0,9657	22
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	2	47	69	0,9809	0,9652	22
Suavização (janela de 7 pontos)	2	41	60	0,9851	0,9734	17
Suavização (janela de 7 pontos) com JK	2	41	59	0,9853	0,9739	17
Suavização (janela de 7 pontos) (regiões dos aditivos)	2	46	68	0,9812	0,9657	22
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	2	47	69	0,9809	0,9652	22
1º derivada (janela de 7 pontos)	2	39	59	0,9865	0,9742	19
1º derivada (janela de 7 pontos) com JK	2	39	60	0,9864	0,9738	20
1º derivada (janela de 7 pontos) (regiões dos aditivos)	2	43	66	0,9838	0,9678	21
1º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	1	54	67	0,9746	0,9669	18
1º derivada (janela de 11 pontos)	2	41	62	0,9851	0,9716	21
1º derivada (janela de 11 pontos) com JK	1	58	72	0,9701	0,9621	18
1º derivada (janela de 11 pontos) (regiões dos aditivos)	2	47	72	0,9804	0,9620	23
1º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	56	69	0,9729	0,9644	18
2º derivada (janela de 7 pontos)	2	37	56	0,9881	0,9771	17
2º derivada (janela de 7 pontos) com JK	2	37	55	0,9879	0,9774	17
2º derivada (janela de 7 pontos) (regiões dos aditivos)	2	37	57	0,9879	0,9760	17
2º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	37	57	0,9877	0,9764	17
2º derivada (janela de 11 pontos)	2	38	58	0,9872	0,9756	18
2º derivada (janela de 11 pontos) com JK	1	48	60	0,9795	0,9730	14
2º derivada (janela de 11 pontos) (regiões dos aditivos)	2	39	60	0,9865	0,9731	17
2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	55	68	0,9735	0,9657	17

APÊNDICE N - Resultados de previsão utilizando os modelos de regressão do aditivo G construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	Valor			(continua) Erro de previsão (%)
	Predito	Desvio	Referência	
Suavização (janela de 3 pontos)	136,7	47,0	166,5	18
	106,8	60,1	112,4	5
	325,8	48,9	220,0	48
	481,0	47,6	423,8	13
	547,0	51,2	560,6	2
Suavização (janela de 3 pontos) com JK	136,3	47,4	166,5	18
	108,9	61,0	112,4	3
	327,8	46,1	220,0	49
	482,3	44,7	423,8	14
	548,2	49,0	560,6	2
Suavização (janela de 3 pontos) (regiões dos aditivos)	138,2	64,9	166,5	17
	73,3	68,4	112,4	35
	316,2	58,0	220,0	44
	484,3	51,9	423,8	14
	556,1	60,2	560,6	1
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	139,4	65,5	166,5	16
	71,1	70,0	112,4	37
	315,4	55,0	220,0	43
	483,6	52,0	423,8	14
	555,0	61,5	560,6	1
Suavização (janela de 7 pontos)	136,7	47,0	166,5	18
	106,8	60,1	112,4	5
	325,8	48,9	220,0	48
	481,0	47,6	423,8	13
	547,0	51,2	560,6	2
Suavização (janela de 7 pontos) com JK	136,3	47,4	166,5	18
	108,9	61,0	112,4	3
	327,8	46,1	220,0	49
	482,3	44,7	423,8	14
	548,2	49,0	560,6	2
Suavização (janela de 7 pontos) (regiões dos aditivos)	138,2	64,9	166,5	17
	73,3	68,4	112,4	35
	316,2	58,0	220,0	44
	484,3	51,9	423,8	14
	556,1	60,2	560,6	1
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	139,4	65,5	166,5	16
	71,1	70,0	112,4	37
	315,4	55,0	220,0	43
	483,6	52,0	423,8	14
	555,0	61,5	560,6	1
1° derivada (janela de 7 pontos)	140,1	48,1	166,5	16
	69,3	61,6	112,4	38
	289,1	47,5	220,0	31
	466,4	48,5	423,8	10
	554,8	60,4	560,6	1
1° derivada (janela de 7 pontos) com JK	139,8	44,5	166,5	16
	66,4	54,6	112,4	41
	291,0	51,6	220,0	32
	470,0	48,9	423,8	11
	557,7	60,1	560,6	1

1° derivada (janela de 7 pontos) (regiões dos aditivos)	141,3	59,9	166,5	15
	56,6	63,4	112,4	50
	281,9	67,9	220,0	28
	462,6	62,4	423,8	9
	551,5	77,0	560,6	2
1° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	180,2	61,5	166,5	8
	53,8	42,0	112,4	52
	234,6	80,3	220,0	7
	400,6	95,8	423,8	5
	473,8	110,7	560,6	15
1° derivada (janela de 11 pontos)	135,1	51,4	166,5	19
	71,0	68,2	112,4	37
	302,5	47,0	220,0	37
	477,1	49,5	423,8	13
	565,4	60,9	560,6	1
1° derivada (janela de 11 pontos) com JK	190,2	71,5	166,5	14
	61,2	52,0	112,4	46
	238,1	84,8	220,0	8
	400,9	92,7	423,8	5
	468,6	115,1	560,6	16
1° derivada (janela de 11 pontos) (regiões dos aditivos)	139,6	71,0	166,5	16
	52,0	76,7	112,4	54
	295,8	72,4	220,0	34
	471,0	68,1	423,8	11
	557,5	84,7	560,6	1
1° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	179,9	73,4	166,5	8
	51,5	45,1	112,4	54
	248,3	89,9	220,0	13
	414,7	98,2	423,8	2
	485,8	115,7	560,6	13
2° derivada (janela de 7 pontos)	143,6	40,3	166,5	14
	64,5	49,8	112,4	43
	261,8	48,7	220,0	19
	442,3	53,4	423,8	4
	534,1	69,5	560,6	5
2° derivada (janela de 7 pontos) com JK	142,6	39,6	166,5	14
	65,4	49,2	112,4	42
	264,0	50,2	220,0	20
	444,2	54,3	423,8	5
	537,3	69,8	560,6	4
2° derivada (janela de 7 pontos) (regiões dos aditivos)	146,0	41,2	166,5	12
	60,9	51,5	112,4	46
	257,2	56,4	220,0	17
	438,7	59,7	423,8	4
	530,3	81,3	560,6	5
2° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	145,0	40,5	166,5	13
	62,1	52,1	112,4	45
	259,1	55,7	220,0	18
	440,1	58,1	423,8	4
	532,9	78,4	560,6	5
2° derivada (janela de 11 pontos)	141,2	42,4	166,5	15
	67,7	56,1	112,4	40
	271,0	46,0	220,0	23
	453,3	48,0	423,8	7
	543,8	61,7	560,6	3
2° derivada (janela de 11 pontos) com JK	172,4	53,1	166,5	4
	64,7	39,6	112,4	42
	231,7	64,0	220,0	5
	401,6	77,3	423,8	5

	480,7	94,7	560,6	14
2° derivada (janela de 11 pontos) (regiões dos aditivos)	141,7	46,8	166,5	15
	66,4	59,3	112,4	41
	266,9	55,0	220,0	21
	452,4	55,3	423,8	7
	542,8	73,2	560,6	3
2° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	183,4	62,6	166,5	10
	61,9	41,7	112,4	45
	220,3	69,4	220,0	0
	385,8	89,7	423,8	9
	457,9	112,9	560,6	18

APÊNDICE O - Modelos de regressão do aditivo Y construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	VL	RMSEC (ppm)	RMSECV (ppm)	R ²		Erro médio de previsão (%)
				Calibração	Validação	
Suavização (janela de 3 pontos)	3	20	32	0,9960	0,9911	7
Suavização (janela de 3 pontos) com JK	2	23	30	0,9944	0,9920	5
Suavização (janela de 3 pontos) (regiões dos aditivos)	2	24	33	0,9941	0,9907	4
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	2	24	32	0,9942	0,9909	4
Suavização (janela de 7 pontos)	3	19	31	0,9961	0,9916	7
Suavização (janela de 7 pontos) com JK	2	24	32	0,9939	0,9911	7
Suavização (janela de 7 pontos) (regiões dos aditivos)	2	24	33	0,9940	0,9906	5
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	2	24	32	0,9941	0,9909	4
1º derivada (janela de 7 pontos)	2	24	33	0,9941	0,9903	4
1º derivada (janela de 7 pontos) com JK	2	24	33	0,9941	0,9906	3
1º derivada (janela de 7 pontos) (regiões dos aditivos)	2	19	25	0,9961	0,9945	5
1º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	20	25	0,9958	0,9943	5
1º derivada (janela de 11 pontos)	2	25	37	0,9933	0,9882	5
1º derivada (janela de 11 pontos) com JK	2	25	36	0,9934	0,9883	5
1º derivada (janela de 11 pontos) (regiões dos aditivos)	2	19	26	0,9961	0,9940	5
1º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	2	19	26	0,9961	0,9940	5
2º derivada (janela de 7 pontos)	2	25	32	0,9933	0,9911	2
2º derivada (janela de 7 pontos) com JK	2	25	32	0,9935	0,9911	2
2º derivada (janela de 7 pontos) (regiões dos aditivos)	2	26	32	0,9931	0,9911	3
2º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	25	32	0,9933	0,9912	3
2º derivada (janela de 11 pontos)	2	24	31	0,9941	0,9918	2
2º derivada (janela de 11 pontos) com JK	2	23	30	0,9943	0,9919	2
2º derivada (janela de 11 pontos) (regiões dos aditivos)	2	23	29	0,9945	0,9925	3
2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	2	23	29	0,9946	0,9926	3

APÊNDICE P - Resultados de previsão utilizando os modelos de regressão do aditivo Y construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	Valor			(continua) Erro de previsão (%)
	Predito	Desvio	Referência	
Suavização (janela de 3 pontos)	133,7	32,3	147,2	9
	84,5	23,6	100,0	16
	214,0	21,9	203,4	5
	372,3	32,0	363,8	2
	615,1	39,2	606,4	1
Suavização (janela de 3 pontos) com JK	151,0	22,7	147,2	3
	103,3	21,5	100,0	3
	222,5	20,1	203,4	9
	344,0	18,8	363,8	5
	584,3	31,1	606,4	4
Suavização (janela de 3 pontos) (regiões dos aditivos)	145,4	23,4	147,2	1
	99,0	21,1	100,0	1
	225,3	24,5	203,4	11
	350,0	25,5	363,8	4
	572,8	32,8	606,4	6
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	145,7	23,9	147,2	1
	99,0	22,5	100,0	1
	224,9	24,3	203,4	11
	350,2	26,2	363,8	4
	573,1	32,8	606,4	5
Suavização (janela de 7 pontos)	133,2	31,2	147,2	9
	84,2	23,1	100,0	16
	215,2	21,5	203,4	6
	372,6	30,3	363,8	2
	614,3	38,5	606,4	1
Suavização (janela de 7 pontos) com JK	154,8	26,6	147,2	5
	106,9	25,5	100,0	7
	223,3	19,3	203,4	10
	340,9	20,3	363,8	6
	579,7	37,0	606,4	4
Suavização (janela de 7 pontos) (regiões dos aditivos)	145,3	23,4	147,2	1
	98,4	20,8	100,0	2
	225,4	25,0	203,4	11
	350,4	25,0	363,8	4
	572,6	33,0	606,4	6
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	145,7	23,8	147,2	1
	98,5	22,4	100,0	2
	225,1	24,9	203,4	11
	350,4	26,0	363,8	4
	573,0	32,9	606,4	6
1° derivada (janela de 7 pontos)	144,3	26,0	147,2	2
	106,4	25,8	100,0	6
	208,9	23,1	203,4	3
	361,6	29,6	363,8	1
	569,8	43,9	606,4	6
1° derivada (janela de 7 pontos) com JK	144,3	26,7	147,2	2
	106,0	26,3	100,0	6
	208,5	24,0	203,4	3
	362,5	30,5	363,8	0
	571,3	45,6	606,4	6

1° derivada (janela de 7 pontos) (regiões dos aditivos)	135,6	16,8	147,2	8
	92,5	16,7	100,0	8
	210,3	18,1	203,4	3
	377,0	20,2	363,8	4
	597,7	23,9	606,4	1
1° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	136,1	18,7	147,2	7
	91,7	18,2	100,0	8
	209,4	20,9	203,4	3
	377,7	22,1	363,8	4
	599,2	23,4	606,4	1
1° derivada (janela de 11 pontos)	146,4	30,9	147,2	1
	108,5	30,3	100,0	8
	213,9	26,4	203,4	5
	356,3	35,8	363,8	2
	555,9	55,5	606,4	8
1° derivada (janela de 11 pontos) com JK	146,3	31,1	147,2	1
	108,5	30,3	100,0	8
	213,7	26,1	203,4	5
	356,5	36,1	363,8	2
	556,2	56,0	606,4	8
1° derivada (janela de 11 pontos) (regiões dos aditivos)	136,5	18,4	147,2	7
	92,5	19,1	100,0	8
	214,2	23,9	203,4	5
	373,0	24,4	363,8	3
	587,2	26,1	606,4	3
1° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	136,7	18,2	147,2	7
	92,6	19,1	100,0	7
	214,0	23,2	203,4	5
	373,0	24,5	363,8	3
	586,9	26,3	606,4	3
2° derivada (janela de 7 pontos)	138,1	24,2	147,2	6
	100,1	23,8	100,0	0
	200,6	19,6	203,4	1
	372,2	23,7	363,8	2
	596,3	31,4	606,4	2
2° derivada (janela de 7 pontos) com JK	138,7	24,4	147,2	6
	100,5	23,2	100,0	0
	200,5	18,7	203,4	1
	371,9	23,1	363,8	2
	595,6	30,9	606,4	2
2° derivada (janela de 7 pontos) (regiões dos aditivos)	141,8	21,6	147,2	4
	95,8	22,0	100,0	4
	200,0	17,7	203,4	2
	376,8	22,5	363,8	4
	602,6	29,5	606,4	1
2° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	142,3	21,8	147,2	3
	96,3	21,2	100,0	4
	200,0	17,3	203,4	2
	376,3	21,4	363,8	3
	601,8	28,3	606,4	1
2° derivada (janela de 11 pontos)	139,4	22,1	147,2	5
	98,6	22,8	100,0	1
	201,4	19,5	203,4	1
	373,4	22,8	363,8	3
	593,4	32,6	606,4	2
2° derivada (janela de 11 pontos) com JK	138,8	22,1	147,2	6
	98,8	22,9	100,0	1
	202,0	19,7	203,4	1
	373,8	23,0	363,8	3

	593,8	33,3	606,4	2
2° derivada (janela de 11 pontos) (regiões dos aditivos)	141,6	19,2	147,2	4
	94,7	19,3	100,0	5
	203,7	17,1	203,4	0
	379,9	20,1	363,8	4
	603,1	26,6	606,4	1
2° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	141,1	19,8	147,2	4
	95,0	19,9	100,0	5
	204,6	17,0	203,4	1
	380,1	20,8	363,8	4
	603,6	27,1	606,4	0

APÊNDICE Q - Modelos de classificação SPA-LDA construídos com variáveis selecionadas no espectro completo.

Pré-processamento (após a extração dos espectros da ROI)	Número de variáveis selecionadas	Número de erros		
		Treinamento	Teste	Previsão
Suavização (janela de 3 pontos) com normalização pela faixa	56	3	6	8
Suavização (janela de 7 pontos) com normalização pela faixa	60	4	4	4
1° derivada (janela de 7 pontos) com normalização pela faixa	50	3	4	6
1° derivada (janela de 11 pontos) com normalização pela faixa	35	10	4	7
2° derivada (janela de 7 pontos) com normalização pela faixa	27	19	5	10
2° derivada (janela de 11 pontos) com normalização pela faixa	64	1	3	6

APÊNDICE R - Modelos de classificação SPA-LDA construídos com variáveis selecionadas apenas nas regiões características dos aditivos.

Pré-processamento (após a extração dos espectros da ROI)	Número de variáveis selecionadas	Número de erros		
		Treinamento	Teste	Previsão
Suavização (janela de 3 pontos) com normalização pela faixa	41	7	8	5
Suavização (janela de 7 pontos) com normalização pela faixa	27	15	2	8
1° derivada (janela de 7 pontos) com normalização pela faixa	26	13	5	9
1° derivada (janela de 11 pontos) com normalização pela faixa	28	12	4	12
2° derivada (janela de 7 pontos) com normalização pela faixa	25	13	8	9
2° derivada (janela de 11 pontos) com normalização pela faixa	64	1	3	6

APÊNDICE S - Modelos de classificação GA-LDA construídos com variáveis selecionadas no espectro completo.

(continua)

Pré-processamento (após a extração dos espectros da ROI)	Número de variáveis selecionadas	Número de erros		
		Treinamento	Teste	Previsão
Suavização (janela de 3 pontos) com normalização pela faixa	58	3	3	7
	57	3	5	4
	48	5	3	4
	53	3	3	3
	54	2	3	5
Suavização (janela de 7 pontos) com normalização pela faixa	48	5	2	3
	46	4	3	5
	32	8	2	5
	50	4	2	4
	49	5	4	3
1° derivada (janela de 7 pontos) com normalização pela faixa	53	6	3	4
	49	3	4	4
	52	3	6	3
	67	2	3	3
	56	3	4	5
1° derivada (janela de 11 pontos) com normalização pela faixa	57	3	5	7
	59	3	3	4
	61	3	5	5
	63	2	8	6
	45	2	5	6
2° derivada (janela de 7 pontos) com normalização pela faixa	47	8	0	6
	45	9	4	7
	51	5	4	7
	50	7	5	8
	38	8	2	4

2° derivada (janela de 11 pontos) com normalização pela faixa	58	2	2	6
	53	3	3	5
	48	4	2	7
	55	4	4	7
	70	2	2	6

APÊNDICE T - Modelos de classificação GA-LDA construídos com variáveis seleccionadas apenas nas regiões características dos aditivos.

(continua)

Pré-processamento (após a extração dos espectros da ROI)	Número de variáveis seleccionadas	Número de erros		
		Treinamento	Teste	Previsão
Suavização (janela de 3 pontos) com normalização pela faixa	24	9	3	5
	22	13	3	6
	21	19	3	4
	23	9	2	4
	28	5	3	3
Suavização (janela de 7 pontos) com normalização pela faixa	19	17	1	4
	18	17	0	5
	15	16	1	4
	18	19	0	8
	19	13	0	5
1° derivada (janela de 7 pontos) com normalização pela faixa	15	21	5	11
	19	20	3	8
	16	27	3	4
	22	17	3	8
	17	20	2	10
1° derivada (janela de 11 pontos) com normalização pela faixa	17	19	6	9
	17	24	5	9
	22	18	4	8
	15	24	6	11
	16	23	5	11
2° derivada (janela de 7 pontos) com normalização pela faixa	23	8	3	4
	19	18	3	5
	20	11	3	3
	21	12	3	4
	24	18	1	4

2° derivada (janela de 11 pontos) com normalização pela faixa	23	14	4	5
	25	13	4	6
	16	21	4	5
	23	11	4	9
	21	11	3	9

APÊNDICE U - Modelos de classificação LDA construídos com variáveis as 53 variáveis das regiões características dos aditivos.

Pré-processamento (após a extração dos espectros da ROI)	Número de erros		
	Treinamento	Teste	Previsão
Suavização (janela de 3 pontos) com normalização pela faixa	1	7	4
Suavização (janela de 7 pontos) com normalização pela faixa	4	3	11
1° derivada (janela de 7 pontos) com normalização pela faixa	6	8	11
1° derivada (janela de 11 pontos) com normalização pela faixa	6	8	11
2° derivada (janela de 7 pontos) com normalização pela faixa	2	11	4
2° derivada (janela de 11 pontos) com normalização pela faixa	3	9	9

APÊNDICE V - Resultados de previsão utilizando os melhores modelos de regressão construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Amostras do aditivo	Modelo do aditivo	Valor			(continua) Erro de previsão (%)
		Predito	Desvio	Referência	
G	G	172,4	53,1	166,5	4
		64,7	39,6	112,4	42
		231,7	64,0	220,0	5
		401,6	77,3	423,8	5
		480,7	94,7	560,6	14
	T	32,9	22,0	166,5	80
		61,5	24,0	112,4	45
		266,2	29,1	220,0	21
		452,7	29,1	423,8	7
		556,3	31,0	560,6	1
	W	146,2	37,4	166,5	12
		40,2	56,7	112,4	64
		245,7	33,3	220,0	12
		434,9	28,5	423,8	3
		521,3	34,1	560,6	7
	Y	59,5	25,8	166,5	64
		13,1	31,2	112,4	88
		204,6	29,4	220,0	7
		378,3	40,0	423,8	11
		471,1	44,1	560,6	16
T	G	250,3	32,3	165,6	51
		51,3	43,6	83,2	38
		274,2	55,4	274,8	0
		403,8	58,2	357,6	13
		514,8	119,7	606,8	15
	T	188,1	12,4	165,6	14
		73,3	23,2	83,2	12
		277,5	25,2	274,8	1
		429,3	19,7	357,6	20
		611,5	35,8	606,8	1
	W	260,4	29,0	165,6	57
		33,1	52,5	83,2	60
		281,6	30,9	274,8	2
		428,7	31,5	357,6	20
		593,6	56,5	606,8	2
	Y	178,8	17,4	165,6	8
		12,3	28,1	83,2	85
		237,0	28,2	274,8	14
		362,0	30,8	357,6	1
		535,6	39,3	606,8	12
W	G	213,8	69,9	150,1	42
		85,3	42,9	99,6	14
		169,1	32,0	167,0	1
		396,2	80,2	413,9	4
		478,5	91,2	539,5	11
	T	28,6	26,1	150,1	81
		17,6	20,6	99,6	82
		126,1	15,8	167,0	24
		449,5	28,7	413,9	9
		551,9	30,4	539,5	2

	W	179,2	37,9	150,1	19
		45,3	50,8	99,6	54
		140,9	24,5	167,0	16
		427,0	28,1	413,9	3
		519,0	34,3	539,5	4
	Y	75,2	30,2	150,1	50
		26,0	31,8	99,6	100
		101,2	23,4	167,0	39
		375,7	39,9	413,9	9
		466,1	42,7	539,5	14
Y	G	307,6	91,2	147,2	109
		148,9	38,8	100,0	49
		231,7	53,3	203,4	14
		374,1	89,9	363,8	3
		562,9	133,3	606,4	7
	T	100,0	42,2	147,2	32
		132,3	18,6	100,0	32
		258,9	25,4	203,4	27
		459,4	29,5	363,8	26
		693,1	32,6	606,4	14
	W	291,2	64,2	147,2	98
		157,1	45,4	100,0	57
		266,2	43,2	203,4	31
		418,1	44,3	363,8	15
		648,3	57,0	606,4	7
	Y	138,8	22,1	147,2	6
		98,8	22,9	100,0	1
		202,0	19,7	203,4	1
		373,8	23,0	363,8	3
		593,8	33,3	606,4	2

APÊNDICE W - Modelos de regressão dos aditivos G e W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	VL	RMSEC (ppm)	RMSECV (ppm)	R ²		Erro médio de previsão (%)
				Calibração	Validação	
Suavização (janela de 3 pontos)	2	45	55	0,9806	0,9739	16
Suavização (janela de 3 pontos) com JK	2	45	54	0,9807	0,9742	16
Suavização (janela de 3 pontos) (regiões dos aditivos)	2	44	53	0,9813	0,9752	19
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	2	44	53	0,9816	0,9755	19
Suavização (janela de 7 pontos)	2	44	54	0,9812	0,9746	16
Suavização (janela de 7 pontos) com JK	2	44	54	0,9812	0,9747	16
Suavização (janela de 7 pontos) (regiões dos aditivos)	2	45	54	0,9811	0,9749	20
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	2	44	53	0,9813	0,9751	20
1º derivada (janela de 7 pontos)	2	40	49	0,9850	0,9792	17
1º derivada (janela de 7 pontos) com JK	2	40	48	0,9851	0,9796	17
1º derivada (janela de 7 pontos) (regiões dos aditivos)	2	41	50	0,9843	0,9781	19
1º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	41	50	0,9842	0,9781	19
1º derivada (janela de 11 pontos)	2	41	50	0,9843	0,9784	18
1º derivada (janela de 11 pontos) com JK	2	41	50	0,9842	0,9781	19
1º derivada (janela de 11 pontos) (regiões dos aditivos)	2	44	54	0,9818	0,9750	22
1º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	1	52	59	0,9742	0,9699	21
2º derivada (janela de 7 pontos)	2	39	48	0,9852	0,9797	15
2º derivada (janela de 7 pontos) com JK	2	40	48	0,9851	0,9798	15
2º derivada (janela de 7 pontos) (regiões dos aditivos)	2	40	48	0,9850	0,9796	15
2º derivada (janela de 7 pontos) (regiões dos aditivos) com JK	2	40	48	0,9849	0,9797	15
2º derivada (janela de 11 pontos)	2	40	49	0,9849	0,9792	15
2º derivada (janela de 11 pontos) com JK	2	40	48	0,9851	0,9798	16
2º derivada (janela de 11 pontos) (regiões dos aditivos)	2	40	49	0,9846	0,9789	15
2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	2	40	49	0,9846	0,9790	15

APÊNDICE X - Resultados de previsão utilizando os modelos de regressão dos aditivos G e W construídos com os espectros somados e subtraídos do espectro médio do papel dos *pixels* cujos valores dos escores estavam abaixo do máximo do histograma (extração dos espectros da ROI através dos histogramas dos escores da PCA individual de cada imagem).

Pré-processamento (após a extração dos espectros da ROI)	Valor			(continua) Erro de previsão (%)
	Predito	Desvio	Referência	
Suavização (janela de 3 pontos)	144,6	47,6	166,5	13
	90,2	68,2	112,4	20
	306,5	42,5	220,0	39
	468,5	36,6	423,8	11
	534,3	41,7	560,6	5
	187,8	48,4	150,1	25
	73,2	51,4	99,6	26
	175,4	31,2	167,0	5
	456,8	35,1	413,9	10
	533,3	41,7	539,5	1
Suavização (janela de 3 pontos) com JK	144,4	48,9	166,5	13
	90,0	70,8	112,4	20
	307,8	39,4	220,0	40
	469,5	33,1	423,8	11
	535,6	40,5	560,6	4
	187,3	47,5	150,1	25
	72,5	52,1	99,6	27
	175,5	32,2	167,0	5
	457,8	33,3	413,9	11
	534,6	40,5	539,5	1
Suavização (janela de 3 pontos) (regiões dos aditivos)	142,9	62,5	166,5	14
	54,9	63,2	112,4	51
	295,2	50,9	220,0	34
	468,8	39,3	423,8	11
	540,1	48,5	560,6	4
	178,3	53,3	150,1	19
	52,4	33,9	99,6	47
	161,2	37,1	167,0	3
	458,3	41,5	413,9	11
	538,5	49,1	539,5	0
Suavização (janela de 3 pontos) (regiões dos aditivos) com JK	143,6	61,3	166,5	14
	54,4	66,0	112,4	52
	293,8	45,2	220,0	34
	468,6	37,4	423,8	11
	540,7	49,1	560,6	4
	177,4	51,8	150,1	18
	52,5	35,8	99,6	47
	160,2	36,2	167,0	4
	458,2	40,2	413,9	11
	539,0	50,0	539,5	0
Suavização (janela de 7 pontos)	141,5	47,9	166,5	15
	87,6	68,3	112,4	22
	303,9	40,9	220,0	38
	467,7	35,9	423,8	10
	535,6	41,7	560,6	4
	187,1	49,1	150,1	25
	72,3	51,0	99,6	27
	174,1	31,0	167,0	4
	455,9	34,4	413,9	10
	534,7	41,7	539,5	1

Suavização (janela de 7 pontos) com JK	141,2	50,0	166,5	15
	87,3	71,4	112,4	22
	305,1	37,7	220,0	39
	468,5	32,1	423,8	11
	536,7	40,6	560,6	4
	186,7	48,1	150,1	24
	71,4	51,4	99,6	28
	174,3	32,7	167,0	4
	456,8	32,5	413,9	10
	535,8	40,6	539,5	1
	142,8	63,5	166,5	14
Suavização (janela de 7 pontos) (regiões dos aditivos)	54,8	64,3	112,4	51
	295,8	51,0	220,0	34
	469,1	39,3	423,8	11
	540,0	49,0	560,6	4
	178,3	54,1	150,1	19
	51,9	32,8	99,6	48
	161,1	37,1	167,0	3
	458,5	41,7	413,9	11
	538,5	49,6	539,5	0
	143,7	63,3	166,5	14
	53,8	66,1	112,4	52
Suavização (janela de 7 pontos) (regiões dos aditivos) com JK	294,5	46,7	220,0	34
	468,6	37,7	423,8	11
	540,1	50,2	560,6	4
	177,9	53,8	150,1	19
	51,6	34,5	99,6	48
	160,1	36,1	167,0	4
	458,1	40,9	413,9	11
	538,6	51,0	539,5	0
	137,5	38,5	166,5	17
	58,5	61,0	112,4	48
	276,7	35,8	220,0	26
1° derivada (janela de 7 pontos)	458,5	31,6	423,8	8
	547,2	41,2	560,6	2
	164,5	50,7	150,1	10
	54,1	50,4	99,6	46
	164,3	25,3	167,0	2
	453,5	30,4	413,9	10
	542,7	42,5	539,5	1
	136,2	36,7	166,5	18
	58,5	56,0	112,4	48
	277,6	35,1	220,0	26
	459,7	28,6	423,8	8
1° derivada (janela de 7 pontos) com JK	548,9	37,7	560,6	2
	162,9	47,9	150,1	9
	53,4	44,5	99,6	46
	164,0	25,9	167,0	2
	454,7	27,9	413,9	10
	544,5	38,6	539,5	1
	134,5	41,4	166,5	19
	43,8	48,9	112,4	61
	271,9	42,7	220,0	24
	457,1	32,9	423,8	8
	547,1	41,3	560,6	2
1° derivada (janela de 7 pontos) (regiões dos aditivos)	171,2	57,9	150,1	14
	46,3	39,7	99,6	53
	164,4	30,3	167,0	2
	451,7	31,7	413,9	9

	543,6	42,0	539,5	1
	135,1	40,5	166,5	19
	43,4	44,4	112,4	61
	271,6	43,4	220,0	23
	456,9	33,8	423,8	8
1° derivada (janela de 7 pontos) (regiões dos aditivos) com JK	546,7	41,0	560,6	2
	171,7	58,7	150,1	14
	46,3	35,6	99,6	54
	164,4	30,8	167,0	2
	451,5	31,7	413,9	9
	543,2	42,0	539,5	1
	134,7	40,6	166,5	19
	58,9	67,6	112,4	48
	285,9	34,8	220,0	30
	465,0	31,3	423,8	10
1° derivada (janela de 11 pontos)	552,1	41,9	560,6	2
	161,4	53,6	150,1	7
	49,5	52,7	99,6	50
	162,1	24,8	167,0	3
	458,9	30,1	413,9	11
	548,5	42,7	539,5	2
	134,5	38,9	166,5	19
	55,3	63,0	112,4	51
	285,1	36,4	220,0	30
	463,8	30,2	423,8	9
1° derivada (janela de 11 pontos) com JK	551,5	39,2	560,6	2
	164,3	50,7	150,1	9
	47,1	46,9	99,6	53
	163,0	25,1	167,0	2
	457,9	28,9	413,9	11
	547,9	40,0	539,5	2
	133,6	48,6	166,5	20
	37,6	56,5	112,4	67
	283,3	44,8	220,0	29
	463,7	32,6	423,8	9
1° derivada (janela de 11 pontos) (regiões dos aditivos)	550,7	43,1	560,6	2
	169,7	67,2	150,1	13
	35,5	39,6	99,6	64
	161,3	32,0	167,0	3
	457,5	31,2	413,9	11
	548,1	42,5	539,5	2
	173,2	68,0	166,5	4
	36,3	36,6	112,4	68
	240,0	74,8	220,0	9
	414,3	79,7	423,8	2
1° derivada (janela de 11 pontos) (regiões dos aditivos) com JK	488,5	97,7	560,6	13
	223,4	89,3	150,1	49
	56,5	41,8	99,6	43
	154,0	27,6	167,0	8
	405,8	82,6	413,9	2
	487,3	95,6	539,5	10
	136,1	33,4	166,5	18
	57,1	46,5	112,4	49
	256,8	34,5	220,0	17
	441,2	37,2	423,8	4
2° derivada (janela de 7 pontos)	537,3	46,6	560,6	4
	173,1	45,9	150,1	15
	68,0	47,8	99,6	32
	172,2	32,6	167,0	3

2º derivada (janela de 11 pontos) (regiões dos aditivos) com JK	170,8	30,1	167,0	2
	445,5	29,9	413,9	8
	539,2	42,0	539,5	0
	135,9	32,5	166,5	18
	58,1	50,7	112,4	48
	261,0	35,3	220,0	19
	450,8	30,4	423,8	6
	543,8	41,1	560,6	3
	170,4	53,3	150,1	14
	63,3	39,4	99,6	36
	170,7	30,3	167,0	2
	446,0	30,0	413,9	8
	539,9	42,5	539,5	0