



**UnB**

Instituto de Química  
Programa de Pós-Graduação em Química

TESE DE DOUTORADO

**Desenvolvimento de métodos quimiométricos para triagem de  
novas substâncias psicoativas em selos utilizando técnicas  
espectroscópicas na região do infravermelho**

LAIZ DE OLIVEIRA MAGALHÃES

ORIENTADOR: JEZ WILLIAN BATISTA BRAGA

BRASÍLIA, DF  
2019



**UnB**

Instituto de Química  
Programa de Pós-Graduação em Química

TESE DE DOUTORADO

**Desenvolvimento de métodos quimiométricos para triagem de  
novas substâncias psicoativas em selos utilizando técnicas  
espectroscópicas na região do infravermelho**

LAIZ DE OLIVEIRA MAGALHÃES

ORIENTADOR: JEZ WILLIAN BATISTA BRAGA

Tese apresentada ao Instituto de  
Química da Universidade de Brasília  
como parte dos requisitos exigidos para  
a obtenção do Título de Doutor em  
Química.

BRASÍLIA, DF  
2019

## Folha de Aprovação

Comunicamos a aprovação da Defesa de Tese do (a) aluno (a) Laiz de Oliveira Magalhães, matrícula nº 15/0183020, intitulada “*Desenvolvimento de Métodos Quimiométricos para triagem de novas substâncias psicoativas em selos utilizando técnicas espectroscópicas na região do infravermelho*”, apresentada no (a) Sala de Reuniões do IQ da Universidade de Brasília (UnB) em 29 de novembro de 2019.

Prof. Dr. Jez Willian Batista Braga  
Presidente de Banca (IQ/UnB)

Prof. Dr. Adriano Otávio Maldaner  
Membro Titular (INC/PF)

Dra. Diana Brito da Justa Neves  
Membro Titular (INC/PF)

Prof.<sup>a</sup> Dra. Ana Cristi Basile Dias  
Membro Titular (IQ/UnB)

Prof. Dr. Fabrício Machado Silva  
Membro Suplente (IQ/UnB)

Em 29 de novembro de 2019.

“O essencial é invisível aos olhos”.

Antoine de Saint-Exupéry

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, pelo dom da vida e pela possibilidade de passar anos na Universidade de Brasília, estudando e desenvolvendo pesquisa, principalmente em tempos tão conturbados como o presente. Deus é inspiração e a força que me leva adiante.

Agradeço à minha família, nas pessoas de Lúcio, meu pai, Leny, minha mãe, Lara, minha irmã e Luiz, meu irmão, pelos anos de convivência e incentivo a realizar os melhores projetos de vida. Agradeço em especial ao meu esposo, Daniel, que sempre me acompanhou em toda a minha jornada acadêmica e, com amor, me incentivou a ir mais longe, investir em estudo, amar o conhecimento e propagá-lo na minha profissão de professora. Por todo esse amor dedicado a mim, nas simples atitudes, muito obrigada! Também não posso deixar de agradecer à família de meu esposo: Carlos e Conceição, meus sogros, Leandro e Mateus, meus cunhados, bem como Deizy, minha concunhada, que compartilharam comigo as alegrias e as dificuldades em se desenvolver pesquisa no país.

No desenvolvimento desta tese de Doutorado, devo imensos agradecimentos àqueles que me acompanharam de perto, dentro da Universidade de Brasília. Agradeço ao meu orientador de Doutorado, Jez, pelo acolhimento, pelos ensinamentos, pela postura e exemplo de humildade durante estes quase cinco anos de trabalho conjunto. O senhor é uma inspiração, como profissional e como ser humano! Agradeço também ao professor Alexandre, que me orientou no Mestrado, pelo carinho e incentivo a continuar na vida acadêmica e pelas boas risadas desde 2010. Agradeço a todos os professores do grupo AQQUA-UnB, que em algum momento desta caminhada me acompanharam: Ana Cristi, Fernando, Valéria e Fernanda.

Agradeço às agências de fomento (CAPES, CNPQ e FAPDF) que, por muito tempo, foram capazes de investir com afinco a pesquisa brasileira, em especial a que eu desenvolvi. Agradeço a professora Ingrid, coordenadora do projeto CAPES Pró-Forense, pela amizade e dedicação a seus alunos vinculados ao projeto, nunca hesitando em enviar nossos trabalhos para congressos e fornecendo materiais essenciais para que eles pudessem ser desenvolvidos.

Agradeço à Polícia Civil do Distrito Federal, na pessoa do perito Luciano, que me acolheu de braços abertos no Instituto de Criminalística da PCDF para desenvolver algo novo, ainda nem pensado. Seu apoio, conhecimento, entusiasmo e dedicação foram essenciais para o desenvolvimento deste trabalho. Também não posso deixar de agradecer à estudante de iniciação científica, Maísa, que me acompanhou nos últimos

anos do Doutorado, me incentivando a estudar mais (claro, muitas perguntas que ela fazia) e produzindo, de fato, junto comigo, os resultados deste trabalho.

Agradeço às minhas amigas, Angélica, Camila e Fabiana, que, desenvolvendo seus próprios projetos de Doutorado, puderam ser apoio dentro e fora dos laboratórios. Vocês são queridas demais, que levarei para toda a vida. Agradeço também à Carla, que, nas contáveis vezes que nos encontramos, sempre foi uma grande incentivadora dos meus projetos profissionais e de vida.

Por fim, agradeço a banca da defesa desta tese, pela disponibilidade e pelas contribuições de melhoria deste trabalho.

Obrigada a todos!

## RESUMO

Nos departamentos de polícia, as apreensões de selos contendo as chamadas *designer drugs* aumentaram nos últimos anos, provavelmente devido à facilidade de se realizar mudanças moleculares nessas substâncias com o objetivo de evitar a legislação e a ação das agências de controle e de segurança pública. O método analítico padrão para determinar essas substâncias em selos é baseado em cromatografia líquida ou gasosa com detecção de espectrometria de massa (LC-MS e GC-MS). Especificamente para a família dos NBOH, é necessário um método GC-MS customizado para evitar uma identificação errada com a substância 2C-I ou a aplicação de cromatografia líquida acoplada a espectrometria de massas com analisador por quadrupolo de tempo de voo (LC-QTOF-MS), que é capaz de realizar sua identificação correta. Nesse contexto, a presente tese tem como objetivo principal a avaliação da espectroscopia de infravermelho próximo (NIR) e infravermelho médio (MIR) para a caracterização e análise de amostras de selos contendo drogas sintéticas para aplicações forenses. Em um primeiro estudo, foram desenvolvidos modelos baseados em SIMCA ou PLS-DA, utilizando espectros NIR das amostras, considerando a existência ou não de drogas nos selos (modelo A) e, uma vez confirmada esta presença, realizou-se a diferenciação de selos contendo NBOMe daqueles que contêm NBOH (modelo B). Estas duas famílias de novas substâncias psicoativas (NPS) são derivadas de feniletilaminas e causam efeitos alucinógenos ainda mais fortes do que a dietilamida do ácido lisérgico (LSD). Esses mesmos modelos PLS-DA foram desenvolvidos com espectros MIR. Os modelos foram propostos como método de triagem para a análise *in loco*, previamente a análise por LC-MS, GC-MS ou LC-QTOF-MS. Os modelos desenvolvidos realizavam uma rotina de detecção de amostras anômalas (*outliers*), fornecendo taxas de eficiência maiores que 96% para PLS-DA. Para o SIMCA (aplicado somente no NIR), as taxas de eficiência foram entre 36 e 92%. A comparação entre o PLS-DA e o SIMCA leva à conclusão de que o PLS-DA apresentou melhores resultados em termos de taxa de eficiência. Entretanto, o SIMCA demonstrou melhor capacidade de detectar classes não incluídas na fase de treinamento. Selos sem drogas foram impregnados com 25E-NBOMe de três formas distintas (gotejamento, imersão e aspersão) para avaliação da distribuição da droga por meio da deconvolução de imagens hiperespectrais, na região do infravermelho médio, por MCR-ALS. Os resultados sugerem que a forma de impregnação pode influenciar na distribuição e na concentração de substâncias no selo. Uma amostra apreendida foi avaliada e os resultados indicaram que os modelos desenvolvidos são capazes de identificar uma substância impregnada e sua dispersão.

## ABSTRACT

In police departments, seizures of blotter papers containing designer drugs have increased in the last years, probably due to the molecularly changes of those substances with the aim of avoiding control and public security agencies. The analytical method for determining these substances in blotter papers is based on liquid or gas chromatography with mass spectrometry detection (LC-MS and GC-MS). Specifically, for NBOH family, it is necessary a customized GC-MS method to avoid a misidentification with the substance 2C-I or the application of liquid chromatography-quadrupole time-of-flight-mass spectrometry (LC-QTOF-MS) which is capable to perform its correct identification. In this context, the present thesis has the purpose of evaluating the Near Infrared and Mid Infrared Spectroscopy (NIR and MIR) for characterization and analysis of synthetic drug in blotter papers for forensic applications. In a first study, multivariate models based on SIMCA or PLS-DA, using NIR spectra of the samples, were developed for classification of seized blotter paper samples, considering the existence or not of drugs in the blotter papers (model A) and, once the drug presence was confirmed, perform the differentiation of blotter papers containing NBOMe from those ones containing NBOH (model B). These two families of new psychoactive substances (NPS) are derivatives of phenethylamines and they cause hallucinogen effects even stronger than the lysergic acid diethylamide (LSD). These same models were developed using its MIR spectra. The models are proposed such as a screening method for *in loco* analysis, prior LC-MS, GC-MS or LC-QTOF-MS analysis. The models have used an outlier detection routine and provided efficiency rates higher than 96% for PLS-DA. For SIMCA (just applied in NIR), obtained efficiency rates were between 36 and 92%. The comparison between PLS-DA and SIMCA lead to the conclusion that PLS-DA has showed better capability in detecting no included classes in training stage. Blotter papers were impregnated with 25E-NBOMe in three different ways (dripping, immersion and sprinkling) for evaluation of drug's distribution by the deconvolution of hyperspectral images in mid infrared region, using MCR-ALS. It is possible to conclude that the impregnation way influences in the distribution and concentration of substances in the blotter paper. A seized sample was evaluated and the results have indicated that the developed models are capable of identifying an impregnated substance and its dispersion.

## SUMÁRIO

1. INTRODUÇÃO E OBJETIVOS .....	1
1.1. Introdução .....	1
1.2. Objetivo geral .....	2
1.2.1. Objetivos específicos.....	2
2. REVISÃO BIBLIOGRÁFICA.....	4
2.1. Química Forense .....	4
2.1.1. Novas substâncias psicoativas .....	7
2.2. Espectroscopia na região do infravermelho .....	11
2.2.1. Espectroscopia NIR.....	12
2.2.2. Espectroscopia MIR.....	14
2.3. Análise multivariada.....	15
2.3.1. Organização dos dados e pré-processamento.....	16
2.3.2. Métodos supervisionados e não supervisionados .....	17
2.3.3. Métodos de discriminação e métodos modelativos .....	18
2.3.4. SIMCA e PLS-DA .....	20
2.3.5. Validação de um modelo .....	22
2.3.6. Identificação de amostras anômalas ( <i>outliers</i> ).....	23
2.3.7. Análise Multivariada de Curvas (MCR) e Imagens Hiperespectrais .....	24
3. MATERIAIS E MÉTODOS .....	28
3.1. Discriminação de drogas sintéticas em selos por espectroscopia no infravermelho próximo (NIR) .....	28
3.1.1. Aquisição de dados .....	28
3.1.1.1. Equipamentos e softwares .....	28
3.1.1.2. Testes iniciais .....	29
3.1.2. Amostras utilizadas e aquisição dos espectros NIR.....	30
3.1.3. Pré-processamento .....	32
3.1.4. Validação cruzada .....	33
3.1.5. Modelos de calibração multivariada .....	33
3.1.6. Rotina de detecção de <i>outliers</i> .....	35
3.1.7. Validação.....	37
3.1.8. Avaliação da robustez do modelo B por permutação de amostras.....	37
3.2. Modelos de classificação por MIR .....	39
3.2.1. Aquisição de dados .....	39
3.2.1.1. Equipamento e <i>software</i> .....	39
3.2.1.2. Modo de aquisição e conversão dos espectros.....	41

3.2.2.	Amostras .....	41
3.2.3.	Escolha do melhor modelo .....	42
3.2.4.	Exclusão de <i>outliers</i> .....	43
3.2.5.	Análise de amostras desconhecidas.....	43
3.3.	Avaliação da distribuição de NPS por imagem hiperespectral e MCR-ALS ..	44
3.3.1.	Equipamento e <i>softwares</i> .....	44
3.3.2.	Amostras simuladas .....	47
3.3.3.	Pré-processamento de sinais, restrições e estimativa inicial.....	49
4.	RESULTADOS E DISCUSSÃO .....	50
4.1.	Identificação de drogas sintéticas em selos por espectroscopia no infravermelho próximo (NIR) .....	50
4.1.1.	Testes iniciais .....	50
4.1.2.	Pré-processamento .....	51
4.1.3.	Discriminação por PLS-DA .....	56
4.1.3.1.	Modelo A: Selos sem drogas <i>versus</i> selos com drogas.....	56
4.1.3.2.	Modelo B: Selos com NBOMe <i>versus</i> selos com NBOH .....	62
4.1.4.	Classificação por SIMCA .....	68
4.1.4.1.	Modelo A: Selos sem drogas <i>versus</i> selos com drogas.....	68
4.1.4.2.	Modelo B: Selos com NBOMe <i>versus</i> selos com NBOH .....	70
4.1.5.	Comparação de resultados por PLS-DA e SIMCA.....	73
4.1.6.	Avaliação da robustez do modelo B por PLS-DA com permutação de amostras.....	74
4.2.	Discriminação de drogas sintéticas em selos por espectroscopia no infravermelho médio (MIR).....	77
4.2.1.	Conjunto de dados e pré-processamento .....	77
4.2.2.	Desenvolvimento dos modelos A e B .....	78
4.2.3.	Análise de amostras desconhecidas.....	83
4.3.	Avaliação da distribuição de NPS por imagem hiperespectral e MCR-ALS ..	84
4.3.1.	Amostras simuladas .....	84
4.3.1.1.	Avaliação dos histogramas dos mapas de concentração e teste de hipótese para as amostras simuladas obtidas gotejamento e imersão.....	95
4.3.2.	Avaliação das imagens de selo contendo 25E-NBOMe e selos não impregnados .....	99
4.3.3.	Avaliação dos modelos para amostra apreendida .....	104
5.	CONCLUSÕES E PERSPECTIVAS .....	110
6.	REFERÊNCIAS BIBLIOGRÁFICAS.....	112
7.	ANEXOS.....	117

## LISTA DE FIGURAS

<b>Figura 1.</b> Estrutura molecular dos NBOMe (X pode ser um átomo de hidrogênio, bromo, cloro, iodo ou ser um grupamento orgânico, como o etil).....	8
<b>Figura 2.</b> Estrutura molecular do NBOH (X pode ser um átomo de bromo, iodo ou um grupo etila).....	9
<b>Figura 3.</b> Esquema de decomposição de uma imagem hiperespectral por MCR-ALS. ....	26
<b>Figura 4.</b> Esquema da seção transversal (à esquerda) e fotografia digital do MicroNIR 1700 (à direita).....	28
<b>Figura 5.</b> Exemplo de aquisição de espectros dos selos com o MicroNIR.....	30
<b>Figura 6.</b> Matriz de dados de absorvância gerada pelo MicroNIR. A legenda dos comprimentos de onda é apresentada na primeira linha e as absorvâncias das 10 replicatas nas dez linhas seguintes.....	30
<b>Figura 7.</b> (a) Fotografia da parte da frente dos selos sem drogas adquiridos pela internet para desenvolvimento do primeiro modelo de classificação multivariada. (b) Fotografia do verso dos selos sem drogas (à direita).....	32
<b>Figura 8.</b> Fluxograma de análise multivariada de selos.....	35
<b>Figura 9.</b> Gráfico de influência Hotelling e resíduos Q com indicação de outliers na área de união dos parâmetros (amostras dentro das áreas delimitadas em vermelho são consideradas outliers e excluídas do conjunto de dados). ....	36
<b>Figura 10.</b> Gráfico de dispersão de amostras com indicação dos limites de classes (retas azuis) e do limite de discriminação (reta vermelha). ....	36
<b>Figura 11.</b> Esquema de permutação das amostras de NBOMe e NBOH para construção dos conjuntos de treinamento e validação para os 10000 modelos (Os números nos cantos inferiores dos quadros que representam as matrizes indicam o número de linhas das matrizes, ou seja, número de amostras multiplicado por 10 replicatas).....	38
<b>Figura 12.</b> (a) Visão geral do equipamento do Instituto de Criminalística da Polícia Civil do Distrito Federal. (b) Posicionamento do selo na célula de leitura do Alpha II. ....	40
<b>Figura 13.</b> (a) Fotografia do microscópio de infravermelho LUMOS com joystick e computador. (b) Visão aproximada da plataforma de suporte e da câmera do equipamento.....	45
<b>Figura 14.</b> (a) Fotografia digital obtida pela câmera CCD do LUMOS do verso de parte de um selo e (b) grade de divisão da imagem obtida para aquisição dos espectros com indicação de sua direção. ....	46

<b>Figura 15.</b> Esquema que representa a forma pela qual se obtém o arquivo “.dpt” de um mapeamento no LUMOS. $l_p$ são as subdivisões da imagem (imagem do pixel) definidas pela grade e $\lambda$ representa os sinais para cada intervalo espectral (833 intervalos, devido à resolução de $8\text{ cm}^{-1}$ ). .....	46
<b>Figura 16.</b> Fotografias do processo de impregnação de 25E-NBOMe nos modos de imersão, aspersão e gotejamento. ....	48
<b>Figura 17.</b> Gráfico de escores das PCs 1 e 2 para 10 replicatas espectrais (5 na parte da frente e 5 no verso dos selos) de 5 amostras de selos apreendidas. Amostra 1 ( $\square$ ), amostra 2 ( $\bullet$ ), amostra 3 ( $\blacktriangledown$ ), amostra 4 ( $\blacktriangle$ ) e amostra 5 (+). ....	51
<b>Figura 18.</b> Espectros brutos na região do infravermelho próximo para amostras que não contêm drogas (branco) com 125 variáveis (comprimentos de onda). ....	52
<b>Figura 19.</b> Espectros brutos na região do infravermelho próximo para amostras que não contêm drogas (branco). ....	53
<b>Figura 20.</b> Espectros brutos na região do infravermelho próximo para as amostras que contêm drogas. ....	53
<b>Figura 21.</b> Espectros pré-processados na região do infravermelho próximo para amostras que não contêm drogas (branco). ....	55
<b>Figura 22.</b> Espectros pré-processados na região do infravermelho próximo para as amostras que contêm drogas. ....	55
<b>Figura 23.</b> Gráfico de RMSEC e RMSECV por número de variáveis latentes para o modelo A. ....	56
<b>Figura 24.</b> Gráfico de influência ( $T^2$ ) versus resíduos para a fase de treinamento no modelo A ( $\bullet$ – selos que contêm drogas; $\blacktriangledown$ – selos que não contêm drogas) antes da detecção de outliers (a), após a primeira etapa de detecção (b), após a segunda (c) e após três etapas de detecção de outliers (d). ....	58
<b>Figura 25.</b> Gráfico de influência ( $T^2$ ) versus resíduos (Q) para o modelo A ( $\bullet$ – selos que contêm drogas; $\blacktriangledown$ – selos que não contêm drogas) antes (a) e após a etapa de detecção de outliers na fase de validação (b). ....	59
<b>Figura 26.</b> (A) Verso de uma das apreensões de selos contendo 25B-NBOH, no qual se observou imagens no verso em comparação com o verso de um selo contendo NBOMe. (B) Espectros dos selos contendo droga que apresentam o verso sem imagens (linhas azuis) e que apresentam imagens (linhas vermelhas). ....	60
<b>Figura 27.</b> Valores estimados de classe para a discriminação por PLS-DA de amostras que contêm ( $\bullet$ ) e que não contêm drogas ( $\blacktriangledown$ ). Símbolos preenchidos indicam a fase de treinamento e vazios indicam a fase de validação. ....	61
<b>Figura 28.</b> Gráfico de RMSEC e RMSECV por número de variáveis latentes para o modelo B. ....	63

<b>Figura 29.</b> Gráfico de influência ( $T^2$ ) versus resíduos para a fase de treinamento no modelo B antes da etapa de detecção de outliers (a), na primeira (b) e na segunda e última etapa (c). Amostras de NBOMe (♦) e NBOH (■).....	64
<b>Figura 30.</b> Gráfico de influência ( $T^2$ ) versus resíduos para o modelo B antes (a) e após a detecção de outliers na fase de validação (b). Amostras de NBOMe (♦) e NBOH (■). .....	65
<b>Figura 31.</b> Valores estimados de classe para a discriminação por PLS-DA de amostras de NBOMe (♦) e NBOH (■). Símbolos preenchidos indicam replicatas na fase de treinamento e símbolos vazios indicam replicatas na fase de validação).....	66
<b>Figura 32.</b> Gráfico de influência ( $T^2$ ) versus resíduos para o modelo B usando espectros de outras drogas como grupo de validação (●). Amostras de NBOMe (♦) e NBOH (■). .....	67
<b>Figura 33.</b> Valores estimados de classe para a discriminação por PLS-DA de amostras NBOMe (♦) e NBOH (■) como conjunto de validação (●) composto por outras drogas. .....	68
<b>Figura 34.</b> Valores de taxa de eficiência das as fases de treinamento e validação de 10000 modelos B por PLS-DA (Em azul, taxas de eficiência da fase de treinamento e em vermelho da fase de validação). ....	74
<b>Figura 35.</b> (a) Espectros ATR-FTIR das amostras apreendidas de NBOMe e NBOH usadas no conjunto de treinamento. (b) Espectros médios do NBOMe (–) e NBOH (–). (c) Visão aproximada dos espectros médios do NBOMe e NBOH de $600\text{ cm}^{-1}$ a $1800\text{ cm}^{-1}$ .....	77
<b>Figura 36.</b> Probabilidade de outlier espectral pela distância de Mahalanobis das amostras usadas no Modelo A por ATR-FTIR.....	79
<b>Figura 37.</b> Valores estimados de classe para discriminação entre amostras contendo NPS ( $\Delta$ – fase de treinamento; $\blacktriangle$ – fase de validação) e amostras que não contêm NPS ( $\circ$ – fase de treinamento; $\bullet$ – fase de validação). (---) Limite de discriminação: 0,52;. 79	
<b>Figura 38.</b> Probabilidade de outlier espectral pela distância de Mahalanobis das amostras usadas no Modelo B por ATR-FTIR.....	80
<b>Figura 39.</b> Valores estimados de classe para discriminação entre amostras contendo NBOH ( $\Delta$ – fase de treinamento; $\blacktriangle$ – fase de validação) e amostras que não contêm NBOMe ( $\circ$ – fase de treinamento; $\bullet$ – fase de validação) . (---) Limite de discriminação: 0,50; (---) Limite superior (1,21) e inferior (-0,21). ....	82
<b>Figura 40.</b> (a) Espectros brutos da amostra impregnada por gotejamento. (b) Espectros pré-processados por SNV da amostra impregnada por gotejamento. ....	85

<b>Figura 41.</b> (a) Espectros puros recuperados para o selo impregnado por imersão. (b) Perfil de concentração da componente 1 e (c) perfil de concentração da componente 2. ....	86
<b>Figura 42.</b> Espectros puros recuperados para o microselo impregnado por gotejamento. C1 é referente ao 25E-NBOMe e C2 é referente ao papel.....	87
<b>Figura 43.</b> Perfis de concentração para o microselo impregnado por gotejamento. C1 é referente ao 25E-NBOMe e C2 referente ao papel. ....	88
<b>Figura 44.</b> Perfil de concentração do 25E-NBOMe para o selo impregnado por gotejamento com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo. Marcação em preto indica a região da parte picotada do microselo. ....	89
<b>Figura 45.</b> (a) Espectros puros recuperados para o microselo impregnado por imersão. (b) Perfil de concentração da componente 1 e (c) perfil de concentração da componente 2. ....	90
<b>Figura 46.</b> Perfil de concentração do 25E-NBOMe para o selo impregnado por imersão com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo.....	91
<b>Figura 47.</b> (a) Espectros puros recuperados para o microselo impregnado com 25E-NBOMe por aspensão. (b) Perfil de concentração da componente 1 e (c) perfil de concentração da componente 2. ....	92
<b>Figura 48.</b> Espectros puros recuperados para a matriz de dados aumentada por gotejamento e imersão. ....	93
<b>Figura 49.</b> Perfis de concentração para o microselo impregnado por (a) gotejamento e (b) imersão a partir da matriz aumentada. C1 é referente ao 25E-NBOMe e C2 referente ao papel.....	94
<b>Figura 50.</b> Perfil de concentração do 25E-NBOMe para o selo impregnado por (a) gotejamento e (b) imersão (matriz aumentada) com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo. ....	95
<b>Figura 51.</b> Histogramas dos perfis de concentração obtidos a partir da deconvolução de dados por MCR das imagens dos selos impregnados por (a) gotejamento e (b) imersão por meio de matrizes simples. ....	96
<b>Figura 52.</b> Histogramas dos perfis de concentração obtidos a partir da deconvolução de dados por MCR das imagens dos selos impregnados por (a) gotejamento e (b) imersão por meio de matriz aumentada. ....	97
<b>Figura 53.</b> Espectros puros recuperados para a matriz aumentada de microselos impregnados por gotejamento e sem impregnação (branco). C1 corresponde ao espectro do papel e C2 ao espectro do 25E-NBOMe.....	99

<b>Figura 54.</b> Perfis de concentração para o microselo impregnado por (a) gotejamento e (b) sem impregnação a partir da matriz aumentada. C1 é referente ao 25E-NBOMe e C2 referente ao papel.....	100
<b>Figura 55.</b> Perfis de concentração para os microselos impregnados por (a) gotejamento e (b) sem impregnação a partir da matriz aumentada com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo. ....	100
<b>Figura 56.</b> Histogramas sobrepostos dos perfis de concentração obtidos a partir da deconvolução de dados por MCR da matriz aumentada dos microselos impregnados por gotejamento (em vermelho) e sem impregnação (em azul). ....	101
<b>Figura 57.</b> Espectros puros recuperados para a matriz aumentada de microselos impregnados por imersão e sem impregnação (branco). C1 corresponde ao espectro do papel e C2 ao espectro do 25E-NBOMe. ....	102
<b>Figura 58.</b> Perfis de concentração para o microselo impregnado por (a) imersão e (b) sem impregnação a partir da matriz aumentada. C1 é referente ao 25E-NBOMe e C2 referente ao papel.....	103
<b>Figura 59.</b> Perfis de concentração para os microselos impregnados por (a) imersão e (b) sem impregnação a partir da matriz aumentada com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo. ....	103
<b>Figura 60.</b> Histogramas sobrepostos dos perfis de concentração obtidos a partir da deconvolução de dados por MCR da matriz aumentada dos microselos impregnados por imersão (em vermelho) e sem impregnação (em azul). ....	104
<b>Figura 61.</b> Espectros puros recuperados para a amostra apreendida contendo 25E-NBOH. C1 é referente ao papel e C2 é referente ao 25E-NBOH. ....	105
<b>Figura 62.</b> Perfis de concentração para a amostra apreendida. C1 é referente ao papel e C2 referente ao 25E-NBOH. ....	106
<b>Figura 63.</b> Espectros puros recuperados para a matriz aumentada amostra apreendida contendo 25E-NBOH e microselo impregnado com 25E-NBOMe por gotejamento. C1 é referente ao papel e C2 é referente ao 25E-NBOH.....	107
<b>Figura 64.</b> Perfis de concentração para o microselo impregnado por (a) gotejamento e (b) para a amostra apreendida da matriz aumentada. C1 é referente ao papel e C2 referente ao 25E-NBOMe/25E-NBOH.....	108
<b>Figura 65.</b> Perfil de concentração do 25E-NBOMe para o microselo impregnado por (a) gotejamento com conversão da escala de cores para concentração em %m/m e (b) perfil de concentração de 25E-NBOH no microselo apreendido com conversão da escala de cores para concentração em %m/m.....	108

## LISTA DE TABELAS

<b>Tabela 1.</b> Quantidades e composições de amostras de selos para os conjuntos de treinamento e validação do primeiro modelo de classificação multivariada.....	31
<b>Tabela 2.</b> Número e conteúdo de amostras utilizadas nas fases de treinamento e validação dos modelos A e B. Números entre parêntesis indicam o número total de espectros para cada grupo de amostras.....	42
<b>Tabela 3.</b> Figuras de mérito para o modelo A utilizando PLS-DA como ferramenta de discriminação.....	62
<b>Tabela 4.</b> Figuras de mérito para o modelo B utilizando PLS-DA como ferramenta de discriminação.....	66
<b>Tabela 5.</b> Figuras de mérito para o modelo A utilizando SIMCA como ferramenta de classificação. ....	70
<b>Tabela 6.</b> Atribuição de classes e número das amostras no modelo B por SIMCA nas fases de treinamento e validação.....	72
<b>Tabela 7.</b> Figuras de mérito para o modelo B utilizando SIMCA como ferramenta de classificação. ....	72
<b>Tabela 8.</b> Médias e desvio padrão das eficiências dos 10000 modelos B por PLS-DA. ....	75
<b>Tabela 9.</b> Número de outliers detectados por distância de Mahalanobis, resíduos espectrais e resíduos na estimativa de classe nos modelos A e B.....	81
<b>Tabela 10.</b> Figuras de mérito dos modelos A e B por ATR-FTIR.....	82
<b>Tabela 11.</b> Resultados por GC-MS/LC-MS e ATR-FTIR para os Modelos A e B. ....	83
<b>Tabela 12.</b> Resumo dos resultados do teste de hipótese de Kolmogorov-Smirnov (teste K-S) para as amostras de gotejamento e imersão por matrizes simples e aumentada. ....	98

## LISTA DE ABREVIATURAS E ACRÔNIMOS

<b>25B-NBOMe</b>	2-(4-bromo-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina
<b>25C-NBOMe</b>	2-(4-cloro-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina
<b>25C-NBOH</b>	2-((2-(4-cloro-2,5-dimetoxifenil)etilamina)metil)fenol
<b>25E-NBOMe</b>	2-(4-etil-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina
<b>25E-NBOH</b>	2-((2-(4-etil-2,5-dimetoxifenil)etilamina)metil)fenol
<b>25I-NBOH</b>	2-((2-(4-iodo-2,5-dimetoxifenil)etilamina)metil)fenol
<b>25I-NBOMe</b>	2-(4-iodo-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina
<b>2C-I</b>	2-(4-iodo-2,5-dimetoxifenil)etan-1-amina
<b>ATR-FTIR</b>	Attenuated Total Reflectance Fourier -Transform Infrared Spectroscopy
<b>C</b>	Componente
<b>CP</b>	Componentes Principais
<b>DOC</b>	2,5-dimetoxi-4-cloroanfetamina
<b>EASI-IMS</b>	Easy Ambient Sonic-Spray Imaging Mass Spectrometry
<b>FIR</b>	Far Infrared
<b>FN</b>	Falso negativo
<b>FP</b>	Falso positivo
<b>FTIR-PAS</b>	Fourier-Transform Infrared Photoacoustic Spectroscopy
<b>G</b>	Gotejamento
<b>GC-MS</b>	Gas Chromatography Mass Spectrometry
<b>HCA</b>	Hierarchical Cluster Analysis
<b>I</b>	Imersão
<b>IC</b>	Intervalo de confiança
<b>k-NN</b>	k-Nearest Neighbor
<b>K-S</b>	Teste de hipótese de Kolmogorov-Smirnov
<b>LC-MS</b>	Liquid Chromatography Mass Spectrometry
<b>LC-QTOF-MS</b>	Liquid Chromatography - Quadrupole Time of Flight Mass Spectrometry
<b>LDA</b>	Linear Discriminant Analysis
<b>LI</b>	Limite inferior
<b>LS</b>	Limite superior
<b>LSD</b>	Dietilamida do ácido lisérgico
<b>MDA</b>	Metilendioxianfetamina
<b>MDMA</b>	Metilendioximetanfetamina
<b>MDPHP</b>	3',4'-metilendioxi- $\alpha$ -pirrolidinoheiofenona
<b>MCR-ALS</b>	Multivariate Curve Resolution - Alternating Least Squares

<b>MIR</b>	Mid Infrared
<b>MPHP</b>	4'-metil- $\alpha$ -pirrolidinohefenona
<b>MSC</b>	Multiplicative Scatter Correction
<b>NIR</b>	Near Infrared
<b>NIRS</b>	Near Infrared Spectroscopy
<b>NPS</b>	New Psychoactive Substances
<b>PCA</b>	Principal Component Analysis
<b>PCDF</b>	Polícia Civil do Distrito Federal
<b>PLS-DA</b>	Partial Least Squares - Discriminant Analysis
<b>Q</b>	Resíduos Q
<b>RMSEC</b>	Root Mean Square Errors of Calibration
<b>RMSECV</b>	Root Mean Square Errors of Cross Validation
<b>SIMCA</b>	Soft Independent Modeling of Class Analogy
<b>SNV</b>	Standard normative variance
<b>SP</b>	Especificidade
<b>SPAL</b>	Seção de Perícias e Análises Laboratoriais
<b>ST</b>	Seletividade
<b>SVD</b>	Singular Value Decomposition
<b>T<sup>2</sup></b>	Influência Hotelling
<b>TEF</b>	Taxa de eficiência
<b>TFN</b>	Taxa de falsos negativos
<b>TFP</b>	Taxa de falsos positivos
<b>TIN</b>	Taxa de inconclusivos
<b>TVN</b>	Taxa de verdadeiros negativos
<b>TVP</b>	Taxa de verdadeiros positivos
<b>V</b>	Verdadeiros (positivos ou negativos)
<b>VL</b>	Variáveis latentes
<b>VN</b>	Verdadeiro negativo
<b>VP</b>	Verdadeiro positivo

# 1. INTRODUÇÃO E OBJETIVOS

## 1.1. Introdução

*Blotter papers* é um termo em inglês que se refere a selos de papel que geralmente são impregnados com substâncias ilícitas, tais como LSD (dietilamida do ácido lisérgico), MDMA (metilenodioximetanfetamina), NBOMe e outras substâncias.<sup>1</sup> No Brasil a apreensão desses materiais pelos departamentos de polícia tem se intensificado nos últimos anos, bem como a substituição de substâncias proscritas por outras substâncias modificadas, as chamadas *design drugs*.<sup>2</sup> Os NBOMe (2-(4-X-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina), onde X é um átomo de bromo, cloro, iodo ou substituinte orgânico como o etil) são drogas sintéticas, derivadas das fenetilaminas, cujo controle é realizado desde 2014 pela atualização da lista F2 (Lista de Substâncias Psicotrópicas de Uso Proscrito no Brasil) da portaria 344 de 1998 da ANVISA.<sup>3</sup> Os NBOMe podem apresentar em sua molécula átomos de bromo, cloro ou iodo (além de grupamentos orgânicos), formando as substâncias específicas 25B-NBOMe, 25C-NBOMe e 25I-NBOMe, respectivamente.<sup>4</sup> Essas substâncias apresentam efeitos alucinógenos semelhantes aos do LSD, porém são utilizadas em concentrações menores, quando comparadas ao LSD, devido a riscos de overdose.

Após a inclusão dos NBOMe na lista de substâncias controladas pela ANVISA, surgiu com maior frequência, no mercado de ilegal, selos impregnados com uma substância molecularmente muito semelhante aos NBOMe. Os NBOH (2-((2-(4-X-2,5-dimetoxifenil)etilamina)metil)fenol, onde X é um átomo de bromo, iodo ou grupo etil) são também derivados das fenetilaminas e diferem em relação aos NBOMe somente pela substituição de um grupo éter ligado a um dos anéis aromáticos da molécula por um grupo OH, gerando um grupo fenol. Tal como os NBOMe, os NBOH apresentam substituições de átomos de halogênios ou grupamento orgânico, sendo mais comum a apreensão de 25B-NBOH, 25I-NBOH e 25E-NBOH.

A análise de rotina destes selos em laboratórios periciais é baseada na utilização de cromatografia líquida ou gasosa com detecção por espectrometria de massas (LC-MS ou GC-MS).<sup>5,6</sup> Porém o método GC-MS, quando aplicado a selos contendo NBOH, pode gerar muitos resultados incorretos, fazendo referência a outra substância proscrita, 2C-I (2,5-dimetoxi-4-iodofenetilamina).<sup>6</sup> Pesquisadores se dedicaram a resolver este problema, adaptando o método de GC-MS ou sugerindo um novo método de detecção, utilizando um equipamento que impede a decomposição do NBOH em 2C-I, a

cromatografia líquida acoplada a espectrometria de massa com analisador por quadrupolo de tempo de voo (LC-QTOF-MS).<sup>5,6</sup>

Considerando estes fatores, um método alternativo de classificação de amostras seria útil de modo a realizar uma triagem, a fim de que selos contendo NBOH pudessem ser previamente discriminados de selos contendo NBOMe, para que ambos fossem destinados ao método mais adequado de análise, gerando um laudo pericial por meio da diminuição das análises cromatográficas. Para tanto, este trabalho se dedicou a associar a espectroscopia no infravermelho próximo (NIR), espectroscopia no infravermelho médio (MIR) e imagens hiperespectrais a ferramentas de análise multivariada para a análise de selos apreendidos contendo substâncias sintéticas como forma de realizar sua classificação e caracterização.

## **1.2. Objetivo geral**

O objetivo geral desta tese é a avaliação da espectroscopia NIR e MIR associada à Quimiometria, bem como o uso de imagens hiperespectrais para a análise de drogas sintéticas em selos visando a classificação dessas drogas e a caracterização de sua homogeneidade.

### **1.2.1. Objetivos específicos**

Considerando o objetivo geral, são objetivos específicos:

1. Empregar um espectrômetro NIR portátil e modelos multivariados, tais como SIMCA ou PLS-DA, para diferenciar selos contendo NPS, apreendidos pela Polícia Civil do Distrito Federal, de selos que não contêm quaisquer substâncias psicoativas impregnadas.
2. Utilizar o mesmo espectrômetro NIR portátil para diferenciar selos que contêm NPS entre as classes NBOMe e NBOH.
3. Avaliar a capacidade de predição dos modelos PLS-DA com um conjunto de validação com características físicas e químicas diferentes das do conjunto de treinamento, por meio dos valores de influência e resíduos das amostras;
4. Avaliar a robustez do modelo PLS-DA que discrimina amostras com NBOMe das amostras com NBOH por meio dos valores de eficiência obtidos com a permutação randômica de amostras em ambas as classes nos conjuntos de treinamento e validação;

5. Comparar os resultados obtidos pelos modelos SIMCA e PLS-DA, considerando suas particularidades, e com o modelo que apresentar melhor eficiência propor um método que possa ser aplicado em rotina em laboratórios forenses e análises em campo.
6. Desenvolver os mesmos modelos por NIR, utilizando a espectroscopia MIR, muito comum em laboratórios de perícia, e aplicá-los em análise de rotina em um laboratório forense.
7. Realizar um estudo da distribuição de drogas sintéticas nos selos por meio da análise por imagem hiperespectral, utilizando equipamentos de obtenção das imagens na região do infravermelho médio para verificar se o modo de incorporação da droga no selo interfere na sua homogeneidade e se essa informação seria relevante para avaliar a similaridade ou correlação de apreensões com a mesma substância.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1. Química Forense

A Química Forense é um ramo que se dedica ao estudo e resolução de casos na área criminal (forense, relativo a foro) com o auxílio de conhecimentos químicos e técnicas analíticas.<sup>7</sup> A Química Forense tem se estabelecido no cenário científico, solucionando problemas práticos e propondo novos métodos de análise de identificação de amostras relacionadas ao tema. Adicionalmente, ela é uma subdivisão das Ciências Forenses, que ainda se dividem em subáreas a depender do material que está sendo analisado, do analito e do que se deseja responder em uma análise.<sup>8</sup>

As técnicas analíticas mais utilizadas no campo forense, para a determinação de analitos em matrizes como fluídos corporais e drogas ilícitas de abuso, são as cromatografias líquida e gasosa, associadas a diferentes tipos de detecção, como a detecção por chama, por espectroscopia e por meio da espectrometria de massas.<sup>9,10</sup> Além dessas técnicas, é crescente a utilização de espectroscopia na região do infravermelho, espectroscopia Raman, bem como técnicas eletroanalíticas para resolução de casos que englobam também as outras subáreas da Química Forense.<sup>11-17</sup> O tipo de técnica a ser utilizada depende da amostra que será analisada, do analito e também do objetivo da análise. Amostras apreendidas em baixa quantidade normalmente necessitam de técnicas de análise não destrutivas ou que requeiram pequena quantidade de amostra, além de técnicas mais rápidas, para favorecer o andamento do processo e para manter a amostra para futuras análises.<sup>18-20</sup>

Devido ao estabelecimento da cromatografia como uma técnica de referência para separação de compostos e posterior análise, esta é comumente empregada em análises de rotina nos laboratórios periciais. A cromatografia, quando associada a uma técnica de identificação de compostos como a espectrometria de massas é considerada uma técnica B pelas recomendações do grupo de trabalho científico para análise de drogas apreendidas (SWGDRUG, do inglês *Scientific Working Group for the Analysis of Seized Drugs Recommendations*), isto é, uma técnica confirmatória, que necessita de técnicas complementares para a determinação de um analito, uma vez que ela confirma uma suspeita proveniente de outra técnica de análise.<sup>21</sup> Nas últimas décadas, as técnicas cromatográficas têm se desenvolvido bastante, bem como os métodos de análise de compostos específicos, de modo que há um aumento na resolução de sinais, na velocidade de análise e menores erros de determinação.<sup>22</sup> Técnicas de detecção baseadas na espectrometria de massas geram resultados extremamente confiáveis e de alta resolução, quando usada, por exemplo, a cromatografia gasosa acoplada a

espectrometria de massas com analisador de quadrupolo por tempo de voo (GC-QTOF-MS do inglês *Gas Chromatography Quadrupole Time of Flight Mass Spectrometry*).<sup>20,23</sup>

Mesmo com estes avanços, que podem melhorar uma análise forense, a cromatografia, assim como outros tipos de técnicas analíticas como a eletroforese, é uma técnica minimamente destrutiva, na qual a amostra precisa ser consumida para gerar o resultado final, mesmo que em pequenas quantidades. Em alguns casos, a quantidade de amostra pode impossibilitar a análise, devido à limitação de algumas técnicas para análises em baixíssimas concentrações ou quantidades. A quantidade de amostra pode influenciar no tempo necessário para a resolução de um caso forense, principalmente quando as técnicas analíticas ainda não são desenvolvidas o suficiente para uma determinada amostra.

Outra desvantagem dessas técnicas analíticas está relacionada aos equipamentos utilizados que, normalmente, são de grande porte e necessitam de aparelhagem anexa que os mantém funcionando. No caso da cromatografia, os cromatógrafos são dependentes de *no breaks* para que se evite a perda dos dados e da análise em caso de queda de energia. Além disso, o ambiente de análise precisa ter temperatura controlada, favorecendo a necessidade de aparelhos de ar-condicionado. Existem cromatógrafos que possuem amostragem automática, existindo um módulo anexo ao aparelho que pode torná-lo ainda maior. A eletroforese necessita de fontes com alta diferença de potencial para realização de uma separação, bem como outros equipamentos e materiais que tornam esta técnica também uma técnica de bancada, exclusivamente utilizada em laboratórios.<sup>24</sup>

Essas características geralmente tornam uma análise forense mais longa, requerendo mais tempo por parte da equipe que irá realizá-la, e com maior custo associado (referente aos materiais utilizados, manutenção de equipamentos, transporte de amostras apreendidas etc.). Independentemente destas técnicas e métodos serem considerados de referências para análises forenses, há uma necessidade de desenvolvimento de métodos alternativos que possam realizar as mesmas determinações, com o mesmo nível de confiança, ou até melhores e em um menor período de tempo.

Nos últimos anos, há um aumento no desenvolvimento de técnicas e métodos alternativos à cromatografia e outras técnicas destrutivas que podem gerar também resultados confiáveis, porém com a vantagem de se manter a amostra intacta ou pouco consumida, bem como tornar a análise rápida e propícia de ser realizada *in situ*, que no caso de drogas, seria no local da apreensão.<sup>25,26</sup> Técnicas espectroscópicas, como espectroscopia na região do infravermelho, ultravioleta, visível e Raman têm sido empregadas no desenvolvimento de métodos pouco destrutivos, rápidos e portáteis,

não só para análise forense, mas também para outros tipos de análise que requeiram estas vantagens analíticas.<sup>11,19,27</sup>

A espectroscopia na região do ultravioleta foi utilizada por Martins e colaboradores<sup>11</sup>, em colaboração com a Polícia Federal, para a discriminação de uísques falsificados e classificação de diferentes marcas, com o auxílio de métodos de calibração multivariada. Os autores foram capazes de discriminar as falsificações com eficiências maiores que 93%, utilizando Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA, do inglês *Partial Least Squares – Discriminant Analysis*), bem como erros de, no máximo, 1,5% para a classificação por marcas. Este tipo de trabalho fornece a institutos de perícia, métodos pouco destrutivos, rápidos e eficientes para resolução de problemas criminais muito comuns. Uma equipe bem treinada para utilizar o método poderia facilmente aplicá-lo em suas análises de rotina.

Álvarez e colaboradores<sup>18</sup> desenvolveram um método de classificação multivariada utilizando a espectroscopia fotoacústica no infravermelho médio por transformada de Fourier (FTIR-PAS, do inglês *Fourier-Transform Infrared Photoacoustic Spectroscopy*) para diferenciar marcas de propelentes utilizados na fabricação de bombas caseiras. Como no trabalho citado no parágrafo anterior, a utilização de uma técnica espectroscópica é eficiente para este tipo de análise, uma vez que não há destruição da amostra, comparando-se com o método cromatográfico de referência, e também há uma maior rapidez na realização do procedimento. Os autores conseguiram fazer a diferenciação de quatro marcas de propelentes com 100% e 97% de eficiência pelas ferramentas quimiométricas de classificação SIMCA e PLS-DA, respectivamente. É importante considerar que as técnicas espectroscópicas de análise, devido à pouca seletividade no sinal, muitas vezes necessitam de ferramentas multivariadas para interpretação dos resultados, assunto a ser abordado na seção 2.4.

No campo da análise de drogas, há também um avanço do uso da espectroscopia como alternativa a métodos cromatográficos de análise. Pereira e colaboradores<sup>28</sup> desenvolveram um método de discriminação multivariada de amostras de novas substâncias psicoativas (NPS, do inglês *New Psychoactive Substances*) por meio da aquisição de espectros de infravermelho – por espectroscopia no infravermelho por transformada de Fourier com reflectância total atenuada (ATR-FTIR, do inglês *Attenuated Total Reflectance Fourier-transform Infrared Spectroscopy*) diretamente nas amostras. O método indicou uma eficiência de discriminação de 95,8%.

A análise de entorpecentes é uma das áreas mais exploradas no meio forense brasileiro devido ao avanço no uso e tráfico dessas substâncias. Os institutos de perícia espalhados pelo Brasil se dedicam à análise de drogas, gerando laudos periciais com muita frequência. Adicionalmente, as técnicas e métodos de análise para estas

substâncias são numerosas e diversificadas, uma vez que há diferentes tipos de analitos neste campo de atuação. Há uma crescente necessidade de desenvolvimento de métodos de análise alternativos que forneçam resultados mais rápidos, precisos e exatos, ainda mais devido ao surgimento, nos últimos anos, das chamadas drogas desenho (do inglês *designer drugs*) ou novas substâncias psicoativas.<sup>2</sup>

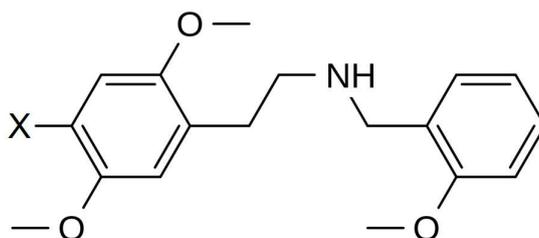
### **2.1.1. Novas substâncias psicoativas**

As novas substâncias psicoativas (NPS) são um grupo de compostos sintéticos provenientes da modificação molecular de outros entorpecentes, os quais apresentam efeitos alucinógenos intensos.<sup>29-31</sup> Este grupo é formado por substâncias que não estavam incluídas na lista de compostos controlados pelas convenções de 1961 e 1971 das Nações Unidas. Pertencem a este grupo de substâncias os canabinóides sintéticos, catinonas sintéticas, triptaminas, derivados de anfetaminas e feniletilaminas, bem como outros compostos orgânicos semelhantes.<sup>32</sup> As novas substâncias psicoativas têm se espalhado pelo mundo como drogas alternativas, isto é, compostos com efeitos semelhantes, porém ainda desconhecidos ou não proscritos pelas agências de controle sanitário.<sup>33,34</sup> No Brasil, a ANVISA (Agência de Vigilância Sanitária) é o órgão que regulamenta o controle de substâncias entorpecentes, através da Portaria 344/98, cujos anexos são atualizados periodicamente. Esta portaria regulamenta o controle especial de substâncias e fármacos, além de compostos psicoativos.<sup>3</sup> Devido às crescentes modificações moleculares de substâncias já controladas pela portaria, há uma dificuldade de controle e atualizações periódicas dos anexos da portaria. No ano de 2014, a agência incluiu na sua lista de substâncias entorpecentes controladas algumas substâncias da família dos NBOMe, sendo, a partir desta data, crime a produção e tráfico desses compostos.<sup>35</sup> Para algumas classes de compostos, como os canabinóides sintéticos, existem anexos da portaria 344/98 que usam classes análogas que englobam grande parte das moléculas modificadas.<sup>36</sup>

No grupo das feniletilaminas se encontra um conjunto de substâncias popularmente chamado de NBOMe.<sup>4</sup> Descobertos em 2003 por Ralph Heim, os NBOMe surgiram no mercado como drogas alternativas ao LSD (dietilamida do ácido lisérgico), apresentando efeitos alucinógenos similares, porém mais intensos que o LSD, mesmo em baixas doses (50 a 250 µg).<sup>1,4,37</sup> Existem relatos que descrevem casos fatais para o consumo desses entorpecentes e que indicam elevada toxicidade destes compostos.<sup>38,39</sup> Este conjunto de substâncias agonistas interage, em concentrações nanomolares<sup>37</sup>, com o receptor de serotonina 5-HT<sub>2A</sub>, desencadeando diversos processos psíquicos e patológicos no corpo, dentre eles, a euforia, náuseas, pânico e

convulsões. Os NBOMe são metabolizados nos rins e sua overdose pode desencadear processos de falência renal.<sup>40,41</sup>

Os NBOMe, também conhecidos por “boom”, “n-bomb”, e “pandora” são moléculas constituídas por um grupo feniletilamina, com substituições de metóxidos e de um átomo de halogênio (Cl, Br ou I) ou grupamentos orgânicos como o etil, ligado a um grupo benzila, também com substituição de metóxido.<sup>41</sup> Dependendo do tipo de halogênio presente na molécula, esta recebe um nome diferente. A molécula contendo o átomo de cloro é conhecida por 25C-NBOMe (2-(4-cloro-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina), a que contém o átomo de bromo é chamada de 25B-NBOMe (2-(4-bromo-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina), a molécula contendo iodo é conhecida por 25I-NBOMe (2-(4-iodo-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina) e a molécula contendo etil é conhecida por 25E-NBOMe (2-(4-etil-2,5-dimetoxifenil)-N-[(2-metoxifenil)metil]etanamina).<sup>41</sup> A estrutura molecular dos diferentes tipos de NBOMe é apresentados na figura 1.



**Figura 1.** Estrutura molecular dos NBOMe (X pode ser um átomo de hidrogênio, bromo, cloro, iodo ou ser um grupamento orgânico, como o etil).

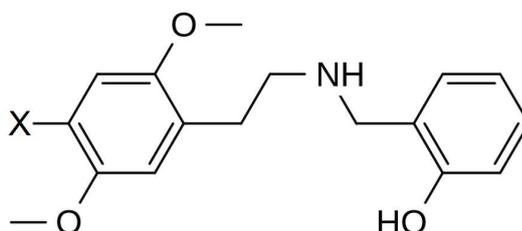
Estas substâncias são utilizadas por via oral ou sublingual, intravenosa, vaginal, retal e também podem ser fumadas.<sup>41</sup> Há relatos de utilização dessas drogas também por via ocular. As vias de acesso destas substâncias no organismo costumam ser regiões altamente vascularizadas, capazes de absorvê-las, direcionando-as diretamente à corrente sanguínea e, por consequência, às específicas áreas do cérebro.<sup>42</sup> Os riscos de utilização deste tipo de *designer drugs* são relacionados a seus efeitos tóxicos diretos, como confusão mental, taquicardia, alucinações, agitação, agressividade, hipertermia, convulsões e acidose metabólica. Adicionalmente, o risco de overdose é eminente, uma vez que uma concentração bem pequena dessas substâncias ( $0,81 \mu\text{g} \cdot \text{kg}^{-1}$ ) já é suficiente para levar a este estado e até mesmo à morte.<sup>1</sup>

Os NBOMe são normalmente apreendidos dissolvidos e impregnados em selos de papel, conhecidos popularmente por *blotter papers*.<sup>1</sup> Este é o mesmo suporte utilizado inicialmente no mercado ilegal para o consumo ilícito de LSD.<sup>43</sup> Os *blotter*

*papers* voltaram ao mercado ilegal agora impregnados não só por LSD, mas também por outros NPS, tal como os NBOMe. Estes selos de papel são apreendidos pelos departamentos de polícia civil e federal por todo o país, devido à proibição da venda e uso dessas substâncias. Com a criminalização periódica de vários compostos psicoativos, tem surgido no mercado ilegal outras substâncias derivadas que, por um determinado momento, não são controlados pelas agências sanitárias, de modo que propiciam a sua disseminação por um maior período nas ruas, antes da atualização da lista de substâncias controladas.

É exatamente esse processo que os NBOMe sofreram nos últimos anos. Observou-se na Polícia Civil do Distrito Federal (PCDF), um aumento nas apreensões de selos contendo NBOMe em relação a um decréscimo na apreensão de selos de LSD. Porém, com a atualização da lista de substâncias controladas no Brasil em 2014, pela RDC N.º 6 da ANVISA<sup>44</sup>, os NBOMe passaram também a ter uma distribuição controlada. Por consequência, outros derivados psicoativos surgiram no mercado ilegal de modo a substituir estes compostos.

Tem sido observado nos últimos cinco anos na Polícia Civil do Distrito Federal, um aumento nas apreensões de selos impregnados com substâncias muito semelhantes a compostos da família dos NBOMe, o 25B-NBOH e o 25I-NBOH. Estas substâncias são similares em efeito e em estrutura, havendo, ao invés do substituinte metóxido em um dos anéis benzênicos, uma substituição de hidroxila, gerando um grupamento fenólico. O 25B-NBOH (2-((2-(4-bromo-2,5-dimetoxifenil)etilamina)metil)fenol), 25I-NBOH (2-((2-(4-iodo-2,5-dimetoxifenil)etilamina)metil)fenol) e o 25E-NBOH (2-((2-(4-etil-2,5-dimetoxifenil)etilamina)metil)fenol) são também compostos pertencentes às feniletilaminas, sendo o 25I-NBOH derivado da substância halucinógena 2C-I (2-(4-iodo-2,5-dimetoxifenil)etan-1-amina), como o 25I-NBOMe.<sup>6</sup> Esta nova família de compostos foi descoberta poucos anos após os NBOMe, como substituição legal para eles<sup>5</sup>, e sua estrutura molecular é apresentada na figura 2.



**Figura 2.** Estrutura molecular do NBOH (X pode ser um átomo de bromo, iodo ou um grupo etila).

Esta simples modificação na molécula dos NBOMe foi suficiente para burlar as agências de controle de substâncias ilícitas pelo mundo, porém, por meio de algumas técnicas analíticas, a identificação da substância pode ser realizada. Esta modificação molecular é perceptível na espectrometria de massas, utilizada normalmente como técnica de detecção associada à cromatografia em laboratórios de perícia. Em espectroscopia também é possível analisar essas novas substâncias, devido à mudança de substituição em um dos anéis benzênicos, de metóxido para hidroxila. Estes dois grupos de drogas são os mais encontrados em selos apreendidos pela PCDF nos últimos anos, e estas são o foco para o desenvolvimento dos métodos desse trabalho.

Existem na literatura diferentes métodos analíticos para determinação de NBOMe e NBOH em selos. A maioria dos métodos utilizados se baseia em cromatografia com detecção por espectrometria de massas ou detecção direta por massas.<sup>43,45</sup> Técnicas cromatográficas são amplamente utilizadas devido as suas características como robustez, exatidão e precisão<sup>24</sup>, uma vez que para gerar laudos periciais, se fazem necessárias estas figuras de mérito.

Recentemente, peritos da Polícia Civil do Distrito Federal, em conjunto com pesquisadores perceberam que os resultados na determinação de 25I-NBOH por GC-MS geravam uma identificação incorreta, isto é, o espectro de massas obtido indicava a presença da substância 2C-I ao invés de indicar a presença do 25I-NBOH. Ao se avaliar a concentração da substância 2C-I nos selos, Arantes e colaboradores<sup>6</sup> notaram que esta estava muito abaixo das quantidades normalmente determinadas, fato que levou a um estudo mais detalhado dessas amostras. Utilizando a cromatografia líquida acoplada a espectrometria de massa quadrupolo de tempo de voo (LC-QTOF-MS), uma técnica de maior resolução e que evita a degradação térmica de NBOH durante a injeção, os pesquisadores foram capazes de realizar a identificação correta de 25I-NBOH. Considerando este fato, é possível que algumas análises realizadas em departamentos de polícia não tenham sido exatas inicialmente, levando a determinação errada do composto, uma vez que foi utilizado o método padrão por GC-MS em vez de um método GC-MS personalizado para evitar esse erro<sup>5</sup> ou o método por LC-QTOF-MS.

Muitos trabalhos são desenvolvidos na área forense com o auxílio da Quimiometria, a ser tratada na seção 2.3. A utilização de técnicas espectroscópicas associadas às ferramentas de análise multivariada tem fornecido alternativas de análise ou métodos de triagem em laboratórios acadêmicos e forenses.

## 2.2. Espectroscopia na região do infravermelho

A espectroscopia na região do infravermelho é uma técnica de identificação de compostos que são formados por moléculas assimétricas, e que apresentam um momento dipolo que se modifica com a frequência absorvida.<sup>46</sup> Esta técnica é baseada no uso de uma região do espectro eletromagnético com energia suficiente para ocasionar transições de energia vibracionais e rotacionais, as quais fornecem informações acerca da estrutura molecular do composto.<sup>46</sup>

A região espectral do infravermelho se divide em infravermelho próximo (NIR, do inglês *Near Infrared*), infravermelho médio (MIR, do inglês *Mid Infrared*) e infravermelho distante (FIR, do inglês *Far Infrared*). Toda a região vibracional do infravermelho compreende comprimentos de onda que vão de 0,78 a 30  $\mu\text{m}$ . O infravermelho próximo é a parte do infravermelho mais energética, uma vez que engloba menores comprimentos de onda. O inverso ocorre para o infravermelho distante, que apresenta maiores comprimentos de onda, próximos à energia rotacional, fornecida pela região eletromagnética das micro-ondas.

Com a absorção de energia na região do infravermelho, é possível gerar o espectro de infravermelho, um gráfico que une bandas ou sinais de absorção, relacionando-os com o número ou comprimento de onda. Este espectro é gerado por meio de deformações conhecidas por estiramento e deformação (para infravermelho próximo e médio) e rotacional (para infravermelho distante), que existem em diferentes modos a depender da região de comprimento de onda. Adicionalmente, uma molécula pode absorver diferentes quantidades de energia na região do infravermelho, gerando as chamadas bandas vibracionais de combinação, que são a soma de bandas que interagem entre si.<sup>46</sup>

O uso do infravermelho para identificação de compostos se baseia na aplicação das espectroscopias no infravermelho próximo e médio, devido à magnitude de energia associada a essas regiões. A análise por meio de NIR e MIR é realizada utilizando-se equipamentos conhecidos por espectrômetros ou espectrofotômetros de infravermelho. Os dois tipos mais comuns de espectrômetros de infravermelhos são os dispersivos e os de transformada de Fourier (FT, do inglês, *Fourier transform*).<sup>46</sup>

Os espectrômetros dispersivos de feixe duplo geram dois feixes de radiação, um que interage com a amostra e outro utilizado como referência. Após os feixes passarem por um conjunto de espelhos, são direcionados para a amostra e referência e, em seguida, para o monocromador, que varia a frequência da radiação que chega ao detector. Esta radiação resultante é amplificada e registrada como espectro. No caso dos espectrômetros dispersivos, devido à mudança da frequência do feixe de radiação,

o espectro obtido está no domínio da frequência, isto é, gera-se um sinal de transmitância por frequência ou número de onda.<sup>24</sup>

Por outro lado, os espectrômetros por FT geram o chamado interferograma, que é uma resposta de intensidade pelo tempo, ou seja, um sinal que está no domínio do tempo. Por meio da transformada de Fourier, o interferograma pode ser convertido em espectro, com domínio na frequência. As maiores vantagens do espectrômetro FT em relação ao dispersivo são a velocidade e sensibilidade. Um espectrômetro deste tipo pode fazer aquisição de vários interferogramas em um curto espaço de tempo, quando comparado aos dispersivos, além de gerar sinais com maior razão sinal/ruído.<sup>24</sup>

O modo de aquisição do espectro de infravermelho também pode variar. O modo mais comum é o de transmitância, no qual se observa a diminuição da intensidade do feixe ao passar pela amostra. Existe também o modo de reflectância, que se divide em reflectância dispersiva e reflectância total atenuada (ATR, do inglês *Attenuated Total Reflectance*). O modo de reflectância se baseia na incidência de radiação sobre a amostra, que a absorve e reflete a radiação resultante de volta ao detector. Espectros de reflectância costumam ser apresentados no formato absorbância × comprimento de onda e suas bandas são geralmente mais largas em comparação àquelas geradas em um espectro de transmitância.<sup>24</sup>

O uso da espectroscopia no infravermelho para caracterização de compostos é muito comum na área forense, uma vez que esta técnica, segundo as recomendações SWGDRUG, é classificada como A, isto é, uma técnica confirmatória para identificação de substâncias.<sup>21</sup> As seções seguintes abordarão alguns fundamentos da espectroscopia no infravermelho próximo e no médio, uma vez que ambas as regiões foram utilizadas na elaboração desta tese.

### **2.2.1. Espectroscopia NIR**

A espectroscopia na região do infravermelho próximo (NIRS, do inglês *Near Infrared Spectrometry*) é uma técnica de espectroscopia vibracional para análise química por meio da interação da radiação eletromagnética na região do infravermelho próximo com a amostra a ser analisada, fornecendo ao analista informações qualitativas e quantitativas.<sup>47</sup> A região do infravermelho próximo compreende os comprimentos de onda de 750 a 2500 nm ( $13300$  a  $4000$   $\text{cm}^{-1}$ ).<sup>47</sup> Esta região espectral, que está entre o infravermelho médio e o visível, frequentemente tem energia mais que suficiente para promover moléculas a estágios vibracionais, porém não tem energia suficiente para causar sua excitação eletrônica.<sup>47</sup> A técnica NIRS apresenta vantagens como velocidade de análise, integridade da amostra mantida, robustez, alta penetração da

radiação, mínimo preparo de amostra, entre outras.<sup>47,48</sup> Essas características fizeram com que esta técnica fosse amplamente utilizada, gerando resultados eficientes, quando associada a métodos de tratamento de dados e de análise multivariada, que serão abordados nas próximas seções.<sup>49</sup>

A resposta instrumental gerada pela espectroscopia NIR é fundamentada pela teoria do oscilador anarmônico.<sup>47</sup> Para uma molécula diatômica em seu estado vibracional fundamental, a absorção de energia específica para esta região espectral propiciará a mudança de seu estado vibracional, uma vez que a energia está em função do deslocamento dos átomos da molécula.<sup>47</sup> Pelo modelo do oscilador harmônico, somente algumas transições são permitidas, porém o observado para a região NIR indica que, nesta região, transições não permitidas, como sobretons e combinações de bandas, ocorrem.<sup>47</sup> De todo modo, o modelo do oscilador anarmônico exemplifica bem o funcionamento vibracional de uma molécula.

Em relação ao estado vibracional fundamental, na região do NIR, as bandas de absorção para uma molécula são bandas de sobretons, isto é, com uma frequência maior que a fundamental, e bandas de combinação, relacionadas à junção de duas ou mais vibrações moleculares.<sup>47</sup> A região do NIR é capaz de fornecer um espectro característico da amostra, uma vez que gera resposta instrumental para esta como um todo, considerando a sua composição.<sup>50</sup> Basta que a amostra seja composta por moléculas que apresentem ligações do tipo C-H, S-H, N-H e O-H para que a técnica NIRS possa ser utilizada.<sup>47,51</sup> Contudo, é importante ressaltar que, apesar do espectro NIR ser capaz de identificar a amostra, esse espectro é muito menos característico do que o obtido na região do infravermelho médio, uma vez que as bandas de cada ligação estão sempre sobrepostas entre si, de forma que o uso da Quimiometria, a ser abordada na seção 2.3, é necessária em praticamente todas as aplicações.<sup>50,52</sup>

A espectroscopia na região do infravermelho próximo é comumente aplicada para análise de alimentos, como grãos e farinha, medicamentos e hidrocarbonetos.<sup>51-53</sup> Esta técnica tem se estabelecido nestes campos de atuação principalmente por suas características anteriormente citadas, e também pela sua portabilidade, com o desenvolvimento de equipamentos NIR portáteis.<sup>50,53</sup> Como essas amostras são conhecidas, elas são usadas para gerar modelos contra os quais novas amostras são testadas. Levando este fato em consideração, a aplicação de NIR em análises forenses é considerada um desafio, devido à variabilidade de amostras que são apreendidas, de modo que modelos desenvolvidos devem ser frequentemente atualizados.

Risoluti e colaboradores<sup>19</sup> desenvolveram um método para classificação de canabinóides sintéticos e feniletilaminas usando espectroscopia na infravermelho próximo e PCA como ferramenta quimiométrica. O objetivo desse trabalho foi

desenvolver um procedimento de triagem para amostras apreendidas de selos e amostras simuladas de ervas. Os autores excluíram dois intervalos de comprimentos de onda de todo o espectro, que eles consideraram não ter influência nos resultados da classificação. Uma análise exploratória foi realizada utilizando amostras simuladas e amostras apreendidas e os resultados mostraram a possibilidade de prever classes NPS apesar da complexidade da matriz. Mesmo com um método de triagem rápido e eficiente, não é comum a presença de um equipamento NIR nos laboratórios forenses.

### 2.2.2. Espectroscopia MIR

Os fundamentos teóricos da espectroscopia vibracional na região do infravermelho médio são basicamente os mesmos da espectroscopia NIR, porém, no primeiro caso, os espectros obtidos possuem mais bandas e bandas mais finas, de modo que a caracterização da amostra é bem mais eficiente. Isto se dá devido aos vários modos de deformações moleculares causadas pela interação do composto com a radiação.

A região espectral do infravermelho médio está entre 4000 a 400  $\text{cm}^{-1}$ . Usualmente, um espectro de infravermelho médio é apresentado na forma de transmitância  $\times$  número de onda, já que existe uma relação diretamente proporcional entre número de onda e frequência/energia.<sup>46</sup> Esta é uma técnica muito utilizada para identificação de compostos orgânicos, principalmente, devido à quantidade de informação espectral gerada a partir das vibrações dessas moléculas. Devido as suas potencialidades, o infravermelho médio é uma técnica muito explorada no ambiente forense, gerando resultados confirmatórios para a composição de drogas.<sup>21</sup>

Na literatura, é comum a aplicação de espectroscopia no infravermelho médio na análise de drogas, dentre elas, os *blotter papers*. Pereira e colaboradores<sup>28</sup> utilizaram a espectroscopia no infravermelho médio com reflectância total atenuada (ATR-FTIR) associada à PCA e PLS-DA para realizar classificação direta de NPS em amostras de selos. Os autores foram capazes de discriminar amostras contendo NPS de amostras que não continham drogas com 95,8% de nível de eficiência e classificaram três tipos de NBOMe com resultado médio de eficiência de 88,9% na fase de validação. Esse trabalho encontrou um problema de discriminação ao lidar com amostras em branco (sem drogas) e amostras de LSD. Um aditivo contido em alguns papéis de selos leva a uma classificação errada no método de discriminação usando amostras contendo NPS e amostras do branco. A presença deste aditivo foi considerada significativa durante a avaliação quimiométrica, não mostrando bons resultados nesse método. Este trabalho foi considerado pelos autores como um método de triagem rápida, de baixo custo e

eficiente, mas dependente de um equipamento ATR-FTIR que tem uma portabilidade limitada e dificilmente permite análises *in situ*.

Quando associada à análise multivariada, a espectroscopia no infravermelho médio também pode fornecer resultados de análise mais abrangentes, isto é, não somente relacionados à identificação de compostos, mas também a sua classificação e dispersão em um suporte, podendo ser uma ferramenta de triagem e *screening*.

### **2.3. Análise multivariada**

A Quimiometria é um ramo da Química, Matemática e Estatística que se dedica à resolução de problemas químicos nos quais o número de variáveis medidas ou observadas para uma determinada amostra é grande, de modo a dificultar a interpretação dos dados. Em se tratando de medidas instrumentais, a quantidade de dados gerados pode ser tão elevada que pode dificultar a determinação da propriedade de interesse. Para solucionar estes casos, a Quimiometria pode ser aplicada, possibilitando, através de cálculos matemáticos, a diminuição do número de variáveis, de modo que as propriedades químicas das amostras possam ser melhor interpretadas.<sup>54</sup>

Com o desenvolvimento dos computadores modernos, houve um avanço nas medidas químicas instrumentais, que não só possibilitou a obtenção de mais informações acerca de uma amostra, mas também gerou certa dificuldade em lidar com esse conjunto maior de dados.<sup>55</sup> Houve, portanto, uma necessidade de se desenvolver novas ferramentas de análise de dados para organização e trabalho com essa grande quantidade de informação química gerada. No final da década de 1960, diversos estudos foram publicados em periódicos da área de Química Analítica sugerindo ferramentas matemáticas e estatísticas que pudessem auxiliar na resolução de problemas químicos multivariados.<sup>54</sup> A partir desses pesquisadores, Isenhour, Jurs, Kowalski e Reilley, as bases da Quimiometria (a qual ainda não era por este nome definida) foram inicialmente formadas. Porém, somente na década de 1970 este ramo de análise química multivariada foi definido como Quimiometria por Svant Wold.<sup>54</sup>

Atualmente existem pacotes computacionais que facilitam a utilização das ferramentas quimiométricas pelo analista.<sup>54</sup> Os métodos quimiométricos vêm se desenvolvendo bastante desde sua criação e suas aplicações são muito diversificadas, isto é, a Quimiometria é aplicada em vários campos de pesquisa, como o campo forense, alimentício, de metabolômica, de autenticação de materiais, entre outros.<sup>54</sup> Quando uma técnica de análise não fornece de imediato resultados claros e exatos, os

métodos quimiométricos podem ser aplicados de modo a facilitar a interpretação dos resultados.

### 2.3.1. Organização dos dados e pré-processamento

Previamente à análise multivariada dos dados, estes precisam ser organizados em matrizes, sendo que na convenção mais comum, estas são compostas pelos dados das amostras (linhas da matriz) e suas variáveis (colunas da matriz). Muitas vezes, antes dessa análise, é necessária a realização de pré-processamento de dados, para eliminação de interferências originadas pela grande variabilidade dos dados, devido a suas origens, variação na linha de base de espectros, intensidade de sinal, grande diferença no patamar dos valores de resposta para cada amostra, ruído espectral etc.<sup>54</sup> Os métodos de pré-processamento de dados podem ser aplicados nas amostras, isto é, nas linhas da matriz de dados, ou nas variáveis, colunas na matriz de dados. A escolha dos pré-processamentos que serão aplicados deve ser cautelosa e coerente para que estes não comprometam o sinal obtido e invalidem a análise.<sup>54</sup>

No caso de dados espectrais, diferentes pré-processamentos podem ser utilizados, e dependem da análise a ser realizada a escolha do melhor pré-processamento ou combinação de pré-processamentos. Um tipo de tratamento de dados nas linhas é o alisamento por Savitsky-Golay, no qual entre dois pontos dos sinais, determinados na janela de alisamento, é gerado um polinômio que suaviza o ruído espectral naquela região. O alisamento tem por objetivo eliminar o ruído espectral, uma vez que se considera que variáveis próximas estão relacionadas e contêm informações similares, sendo possível reduzir o ruído sem perda de informação significativa. Mas dependendo do tamanho da janela de realização do algoritmo, o alisamento pode suprimir sinais relevantes para a análise multivariada.<sup>56</sup> As derivadas também são pré-processamentos de dados nas linhas que eliminam os deslocamentos constantes e as inclinações nos espectros, que podem surgir devido a características das amostras que levam a um espalhamento da radiação. É usual utilizar primeira ou segunda derivada para dados espectrais, uma vez que este tipo de pré-processamento elimina deslocamentos aditivos e multiplicativos dos espectros e enfatiza pequenas diferenças de sinal, propiciando maior distinção entre espectros.<sup>54,56</sup>

Além disso, existem métodos de normalização, como a padronização normal de sinal (SNV, do inglês *Standart Normal Variate*) e a correção multiplicativa de espalhamento (MSC, do inglês *Multiplicative Scatter Correction*) que corrigem a variação de linha de base, eliminando efeitos aditivos e multiplicativos nos sinais.<sup>54,56</sup> O SNV corrige eventuais efeitos de variação na fonte de radiação, caminho óptico e linha

de base devido ao espalhamento de radiação, enquanto o MSC promove a diminuição do espalhamento multiplicativo da radiação, observado em espectros inclinados.

Em relação ao pré-processamento de dados nas colunas (variáveis), o pré-processamento mais utilizado em espectros é o de centrar dados na média. É calculado o valor médio em uma coluna e essa média é subtraída de todos os outros valores da coluna em questão. Este método não prejudica a informação contida nos dados, mas somente realiza uma mudança dos eixos, trasladando a origem destes para o centro da dispersão dos dados de treinamento, de modo que possam ser melhor observadas as diferenças presentes entre as linhas da matriz (amostras).<sup>54,56</sup> Na prática, isso faz com que essas diferenças entre os espectros agora apresentem uma maior variância, uma vez que a média dos espectros foi subtraída do conjunto de dados.

Após realizar o correto pré-processamento dos dados, diferentes modelos multivariados podem ser aplicados, com o intuito de analisar a amostra de maneira qualitativa ou quantitativa a depender do objetivo da análise.

### **2.3.2. Métodos supervisionados e não supervisionados**

Os modelos multivariados podem ser divididos em modelos supervisionados e modelos não supervisionados. Os não supervisionados geram modelos mais primários, que são utilizados no reconhecimento de padrões e na análise exploratória de dados.<sup>54</sup> Isto significa que esses modelos são capazes de identificar as semelhanças e diferenças entre amostras, e até mesmo o comportamento destas em relação a um determinado parâmetro de medida. Se encaixam nesse grupo a análise por componentes principais (PCA, do inglês *Principal Component Analysis*) e a análise por agrupamentos hierárquicos (HCA, do inglês *Hierarchical Cluster Analysis*).<sup>54,57</sup>

O PCA explica o comportamento das amostras por meio combinação das variáveis originais para gerar um novo conjunto de variáveis, chamadas de componentes principais, as quais apresentam ortogonalidade entre si, o que geralmente reduz consideravelmente o número e variáveis (colunas) da matriz de dados.<sup>54</sup> Isto é, novos eixos são gerados, de modo que cada um tenta explicar a informação contida na matriz original, eliminando-se informação redundante, diminuindo-se o erro na projeção através da maximização da variância explicada pelas componentes principais.

A decomposição por PCA é realizada gerando-se matrizes chamadas de pesos e escores, a partir da expressão apresentada na equação 1.

$$\mathbf{X} = \mathbf{TL}^T \quad (\text{Equação 1})$$

Sendo  $\mathbf{X}$  ( $i, j$ ) a matriz original de dados com  $i$  linhas referentes às amostras e  $j$  colunas referentes às variáveis originais,  $\mathbf{T}$  ( $i, A$ ) é a matriz de escores, a qual indica a relação entre amostras e  $\mathbf{L}$  ( $A, j$ ) é a matriz de pesos, que indica a relação entre variáveis, onde  $A$  é o número de componentes principais escolhidos para a modelagem.<sup>54</sup> A redução de variáveis propicia ao analista maior clareza para interpretação do conjunto de dados, por meio da projeção destes em um subespaço com um número menor de variáveis.

Os métodos não supervisionados se dedicam a reconhecer os padrões semelhantes em amostras de modo a agrupá-las em um subespaço gerado pela decomposição matricial. Por outro lado, os métodos supervisionados já possuem informações de grupos ou concentrações definidas, isto é, estes métodos geram modelos matemáticos baseados na informação de que cada amostra pertence a uma classe pré-definida ou possui determinada concentração. Ao se utilizar o conhecimento prévio das classes na análise não supervisionada de dados, é possível desenvolver modelos supervisionados de classificação.<sup>54,57</sup> Esses modelos podem empregar diferentes medidas de similaridade entre as amostras. Métodos de classificação multivariada são desenvolvidos por meio de modelos supervisionados que realizam a separação das amostras em classes através de uma etapa de modelagem ou treinamento. Estes métodos são assim chamados devido ao fato de que as amostras utilizadas em classes pré-definidas geram informação que auxiliam na classificação de novas amostras.<sup>54</sup>

A discriminação ou modelagem de classes, feita por esses modelos, ocorre a partir de um conjunto de amostras conhecidas chamadas de conjunto de treinamento. Essas amostras são modeladas a fim de gerar uma regra de classificação, ou seja, um modelo empírico a fim de que amostras que não foram utilizadas na fase de treinamento possam ser analisadas por meio dessa regra de classificação ou modelo definido previamente.<sup>57</sup> Antes que seja realizada a classificação das novas amostras desconhecidas, é necessário que o modelo seja validado com um outro conjunto de amostras conhecidas, chamadas de conjunto de teste ou de validação. Isto é essencial para se avaliar a capacidade de predição do modelo ou sua eficiência.<sup>57</sup> Havendo resultados eficazes, o modelo pode então ser utilizado na classificação das novas amostras de interesse.

### **2.3.3. Métodos de discriminação e métodos modelativos**

Os métodos supervisionados utilizados para classificação de amostras podem ser divididos ainda em métodos de discriminação e métodos modelativos. Essa subdivisão surge a partir da forma com a qual se classifica uma amostra ou grupo de

amostras de outro(os) grupos(os). Quando a delimitação entre as classes é definida por um espaço fechado de classe, construído com base nas informações amostrais e um nível de confiança pré-determinado, tem-se os métodos modelativos, também conhecidos por classificadores de uma classe.<sup>58,59</sup> Dentre os métodos modelativos mais empregados na literatura está o SIMCA (do inglês *Soft Independent Modeling of Class Analogy*). Métodos modelativos têm sido recomendados como mais apropriados para a certificação da autenticidade de amostras, de modo que são classificadas autênticas somente aquelas que correspondem à classe de interesse.

Já no caso dos métodos de discriminação, conhecidos também como classificadores de duas classes ou multiclassses, o delimitador destas é construído com base na contribuição de todas as classes, sendo que o modelo deve sugerir se as amostras são ou não pertencentes a determinada classe.<sup>58</sup> Se enquadram neste grupo os métodos LDA (do inglês *Linear Discriminant Analysis*), PLS-DA (do inglês *Partial Least Squares - Discriminant Analysis*) e *k*-NN (do inglês, *k-Nearest Neighbor*), bem como outros não abordados nessa tese. Um dos maiores problemas dos métodos de discriminação é a sua dificuldade em classificar amostras que não se assemelham com nenhuma classe pré-definida na etapa de treinamento, uma vez que é formado um perfil padrão para cada classe. Devido a isso, há autores que recomendam a utilização de métodos modelativos quando se desconhece as características das novas classes que não puderam ser incluídas na fase de treinamento.<sup>60</sup> Todos os métodos modelativos podem ser utilizados como métodos de discriminação, porém o inverso não é possível. Contudo, há estudos que tentam desenvolver métodos de discriminação modificados, isto é, adaptados com vertentes modelativas, de modo a unir as principais vantagens de ambos.<sup>58,61</sup>

Existem diferenças de eficácia entre os métodos modelativos e de discriminação citados. Oliveri e Downey<sup>58</sup> discutiram em seu trabalho a funcionalidade dos métodos modelativos e métodos de discriminação para definição de autenticidade de amostras. Os autores concluíram que, para fins de autenticidade, na qual uma classe (autêntica) é relevante para classificação, os métodos modelativos apresentam melhores resultados em comparação aos métodos de discriminação, que necessitam de, ao menos, duas classes para que haja sentido na classificação. O uso indevido, isto é, não recomendado, de métodos de discriminação para autenticação de amostras, pode, em alguns casos, fornecer resultados tendenciosos que podem comprometer a sua habilidade de discriminação, devido, principalmente, à variação do perfil das novas amostras. Este fator é considerado importante na aplicação de métodos multivariados em análise forense, uma vez que, com o passar dos anos, a apreensão de drogas pode ter um perfil diferente.

#### 2.3.4. SIMCA e PLS-DA

A maioria dos métodos supervisionados lineares derivam do PCA, como é o caso do SIMCA, que é um modelo de classificação por modelagem, e do PLS-DA, que é um método de discriminação. O método SIMCA gera modelos de classificação probabilística, realizando modelagem de cada classe por PCA, uma vez que ele se baseia na probabilidade de uma amostra pertencer a uma classe específica, de modo que para cada uma delas há um modelo PCA próprio que define seus limites. Para a criação de um modelo SIMCA, é realizada a decomposição em PCA para cada classe a ser modelada. Essas componentes são essenciais para modelar a distribuição da classe em um subespaço definido por essas PCs (Componentes principais, do inglês *Principal Components*). Isso faz com que o modelo SIMCA seja capaz de classificar amostras em uma ou mais classes ou até mesmo em nenhuma, a depender da proximidade da amostra aos subespaços definidos pelas PCs das diferentes classes.<sup>54</sup>

O método PLS-DA realiza a discriminação entre amostras. Na prática, é realizada uma regressão entre os dados instrumentais (variáveis) em relação a um vetor que divide as amostras em duas classes: uma que é considerada classe alvo, com valor de classe igual a 1, e outra contendo as amostras de todas as outras classes, com valor de classe igual a zero. Quando existirem diferentes classes são construídos diferentes modelos PLS-DA para cada classe, discriminando uma em relação às outras.<sup>54,57,62</sup> Os valores estimados obtidos tentam explicar o modelo com um número reduzido de variáveis, chamadas de variáveis latentes. Para se calibrar ou treinar PLS-DA, bem como qualquer tipo de análise supervisionada, é necessária uma matriz ( $\mathbf{X}$ , equação 2) que contenha as informações de cada amostra de treinamento que será a base do modelo gerado. As linhas nesta matriz são as diferentes amostras e as colunas são suas variáveis. O PLS-DA faz com que os dados sejam decompostos, de modo que o comportamento das amostras seja explicado por somente algumas variáveis latentes, isto é, variáveis cujas informações sobre a amostra sejam mais relevantes no modelo.<sup>54,57,62</sup> As equações 2 e 3 representam a decomposição matricial realizada pelo método PLS. Os valores estimados para as amostras após a decomposição por PLS, bem como o limite de discriminação gerado, indicaram a eficiência de discriminação do modelo.

$$\mathbf{X} = \mathbf{TL}^T + \mathbf{E} \quad (\text{Equação 2})$$

$$\mathbf{y} = \mathbf{Tq}^T + \mathbf{f} \quad (\text{Equação 3})$$

Onde  $\mathbf{X}$  (i, j) é a matriz contendo os dados instrumentais das amostras de treinamento,  $\mathbf{T}$  (i, A) é a matriz escores (do inglês *scores*),  $\mathbf{L}$  (A, j) e  $\mathbf{q}$  (A, 1) são as matrizes de pesos (do inglês *loadings*), sendo A o número de variáveis latentes,  $\mathbf{E}$  (i, j) e  $\mathbf{f}$  (i, 1) são as matrizes de erros de  $\mathbf{X}$  e  $\mathbf{y}$ , respectivamente.

A fim de se proceder com uma escolha coerente de componentes principais ou variáveis latentes para um modelo PCA, PLS-DA ou SIMCA, utilizam-se métodos de validação cruzada na calibração do modelo. A validação cruzada consiste na construção do modelo multivariado por meio da retirada de uma ou mais amostras da matriz  $\mathbf{X}$  para o cálculo inicial do modelo. Na iteração seguinte do modelo essas amostras são repostas e é feita a retirada de outra(s). Este processo se repete quantas iterações sejam necessárias para a participação de todas as amostras do conjunto de treinamento.<sup>54</sup>

Diferentes métodos de validação cruzada podem ser utilizados. O mais comum é o método “deixar uma de fora”, mais conhecido por *leave one out*, o qual retira um espectro/linha da matriz por vez em cada iteração.<sup>54</sup> Há métodos que fazem a retirada de um grupo de amostras/espectros para a validação cruzada. Estes normalmente são utilizados quando há replicatas no conjunto de dados. O método “blocos contíguos”, mais conhecido por *contiguous blocks*, realiza a retirada de todas as replicatas de uma amostra para a validação cruzada por iteração, isto é, a cada iteração do cálculo uma amostra composta por todas as suas replicatas é retirada do grupo de treinamento e passa a fazer parte do grupo de validação cruzada.<sup>57</sup> O método “janelas venezianas” ou *venetian blinds*, realiza a retirada de uma replicata de cada amostra, de modo a uní-las em um conjunto de validação cruzada, a cada iteração.<sup>57</sup> Em casos que as amostras (constituídas por replicatas) do conjunto de treinamento são heterogêneas entre si, a escolha de um método como *venetian blinds*, que considera essa heterogeneidade, é essencial para a eficiência da etapa de validação cruzada.

A validação cruzada é um passo essencial para a escolha do posto do modelo. O posto é o número de variáveis que podem descrever bem um modelo com baixos valores de erros.<sup>54</sup> Para PCA e SIMCA, por exemplo, o posto é o número de componentes principais a ser selecionado. Já no PLS-DA, o posto é o número de variáveis latentes. As amostras na validação cruzada geram as raízes de erro médio quadrático (RMSECV, do inglês *Root Mean Squared Errors of Cross Validation*), cujos valores vão diminuindo à medida que se aumenta o posto do modelo. O posto escolhido é normalmente aquele no qual a diminuição do RMSECV começa a cessar e apresente um valor próximo das raízes de erro médio quadrático de calibração (RMSEC, do inglês *Root Mean Squared Errors of Calibration*), obtido para as amostras na ausência da

validação cruzada. A escolha do posto é fundamental para evitar o sobreajuste ou o subajuste do modelo.

O sobreajuste está relacionado a uma escolha de posto maior do que o necessário, tornando o modelo muito restrito para classificar novas amostras. Um número grande de variáveis latentes ou componentes principais escolhidos aproxima o modelo ao conjunto de dados original, de modo que o modelo só será capaz de explicar a variabilidade de dados de amostras muito semelhantes ao conjunto de calibração. Por outro lado, o subajuste leva o modelo a explicar a variabilidade de muitas amostras, até mesmo aqueles que diferem muito do conjunto de calibração, devido ao número de posto ser menor do que deveria.

### 2.3.5. Validação de um modelo

Após a fase de treinamento do modelo, é importante que este seja testado com um conjunto de amostras distintas, isto é, amostras não utilizadas na fase supracitada. Esta é a chamada fase de teste ou validação, a qual deve ter uma eficiência relativamente semelhante à eficiência da fase de treinamento para que o modelo seja considerado coerente e útil. A validação do modelo final é dada pela avaliação de algumas figuras de mérito, como as taxas de falso positivo (TFP) e falso negativo (TFN), a sensibilidade, especificidade e eficiência. A taxa de falso positivo (equação 4) está relacionada à porcentagem de amostras que não são da classe de interesse, porém são nela classificadas. A taxa de falso negativo (equação 5) se relaciona à porcentagem de amostra da classe de interesse que não são nela classificadas. A sensibilidade (TST) está relacionada à quantidade de amostras que são pertencentes a uma classe e são corretamente classificadas nela. Estas amostras são chamadas de verdadeiros positivos (VP), e sua taxa pode ser calculada segundo a equação 6. A especificidade (TSP) está relacionada à quantidade de amostras que não pertencem a uma classe e não são classificadas nela. Estas amostras são chamadas de verdadeiros negativos (VN), e sua taxa pode ser calculada segundo a equação 7. A eficiência de uma fase (TEF) é dada pela subtração de TFP e TFN de 100, como apresentado na equação 8.

$$TFP = \frac{FP}{FP+VN} \times 100 \quad (\text{Equação 4})$$

$$TFN = \frac{FN}{FN+VP} \times 100 \quad (\text{Equação 5})$$

$$TST = \frac{VP}{FN+VP} \times 100 \quad (\text{Equação 6})$$

$$TSP = \frac{VN}{VN+FP} \times 100 \quad (\text{Equação 7})$$

$$TEF = 100 - (TFP + TFN) \quad (\text{Equação 8})$$

### 2.3.6. Identificação de amostras anômalas (*outliers*)

De modo a garantir elevada eficiência de um modelo, é necessário avaliar a existência de amostras cujo comportamento difere consideravelmente do conjunto de dados. As amostras anômalas, também chamadas de *outliers*, são amostras cujas medidas diferem de maneira significativa das medidas para as outras amostras pertencentes ao conjunto de dados. O fenômeno de aparição de *outliers* pode estar relacionado a diferenças na amostra, problemas instrumentais e, até mesmo, erros de medida. Quando incluídas no conjunto de dados de treinamento ou calibração, as amostras anômalas interferem na definição das componentes principais ou das variáveis latentes, podendo contribuir para o aumento dos erros de predição.

Uma amostra considerada como *outlier* normalmente apresenta valores de resíduos de decomposição (Q) e influência (Hotelling  $T^2$ ) significativamente maiores quando comparados aos valores das amostras do conjunto de dados de treinamento, o que pode ser determinado através de limites em um nível de confiança pré-determinado. Em análise multivariada, estes dois parâmetros (resíduos Q e  $T^2$  Hotelling) podem ser utilizados para eliminação dos *outliers*.

A estatística Hotelling ( $T^2$ ) é utilizada para avaliação da influência das medidas no conjunto de dados. A influência ou *leverage* indica o quanto uma amostra  $i$  está distante do centro do conjunto de dados. Quanto maior essa influência, maior a capacidade da amostra  $i$  em prejudicar a calibração de um modelo. A influência Hotelling  $T^2$  é dada pelo quadrado da distância de Mahalanobis até o centro de dispersão dos dados, como apresentado nas equações 9 e 10.<sup>56</sup>

$$S = \frac{(T^T T)}{(I-1)} \quad (\text{Equação 9})$$

$$T_i^2 = t_i \text{inv}(S) t_i^T \quad (\text{Equação 10})$$

Onde  $S$  é a matriz de variância-covariância dos escores,  $I$  é o número de amostras,  $T$  é a matriz de escores no número de componentes ou de variáveis latentes escolhido e  $t$  é o valor de escore da amostra  $i$ .

Já os resíduos  $Q$  se referem à soma quadrática dos erros gerados pela comparação dos dados estimados em relação aos dados experimentais, e seu cálculo é apresentado na equação 11.<sup>56</sup>

$$Q_i = e_i e_i^T \quad (\text{Equação 11})$$

Estes parâmetros, associados ao seu limite de confiança, são frequentemente utilizados para a exclusão de amostras que podem prejudicar o desenvolvimento do modelo. Os limites para os parâmetros Q e T<sup>2</sup> também podem ser calculados a partir de equações que consideram a distribuição qui-quadrado X<sup>2</sup> (Q) e a distribuição de Fisher (T<sup>2</sup>) apresentadas no manual do PLS\_Toolbox 4.0.<sup>56</sup>

Além disso, é possível utilizar os limites superior e inferior de classes. Estes limites também podem ser calculados com base nos erros dentro de uma classe e fornecem uma faixa aceitável de valores estimados de classe para as amostras, valores esses nos quais estas amostras não são consideradas *outliers*. O cálculo para esses limites são apresentados no trabalho de Soares e colaboradores e também no trabalho de Martins e colaboradores e se baseiam no valor crítico do teste t descrito na norma da ASTM E1655-05 com correção de viés no nível de 99,9% de confiança.<sup>11,50,63</sup> Amostras que apresentarem resíduos que ultrapassem esse valor são consideradas *outliers*.

### 2.3.7. Análise Multivariada de Curvas (MCR) e Imagens Hiperespectrais

A análise multivariada de curvas (MCR, do inglês *Multivariate Curve Resolution*) é uma ferramenta de análise multivariada que se dedica à deconvolução de sinais para estimar espectros puros e perfis de intensidade relativa de um conjunto de dados obtidos para uma mistura. Para um conjunto de dados de uma mistura, contendo sinais espectrais, a ferramenta MCR é capaz de gerar um modelo aditivo bilinear contendo as contribuições dos componentes da mistura.<sup>64</sup> É comum se utilizar métodos iterativos no desenvolvimento de modelos de análise multivariada, pois estes métodos costumam ser mais robustos à presença de ruídos nos dados e gerar estimativas mais próximas aos espectros e intensidades relativas esperadas.<sup>65</sup> Quando esse método iterativo para MCR é realizado por meio de mínimos quadrados alternantes, a ferramenta de análise é denominada MCR-ALS (do inglês, *Multivariate Curve Resolution with Alternating Least-Squares*). A equação 12 representa a forma pela qual ocorre a resolução de curvas no MCR.

$$D = CS^T + E \quad (\text{Equação 12})$$

Em que **D** é a matriz original, contendo os dados da análise, **C** é a estimativa da matriz de intensidades relativas, **S** é estimativa da matriz de espectros puros (contendo o número de espectros relacionado ao número de diferentes componentes no modelo) e **E** é a matriz de erro.

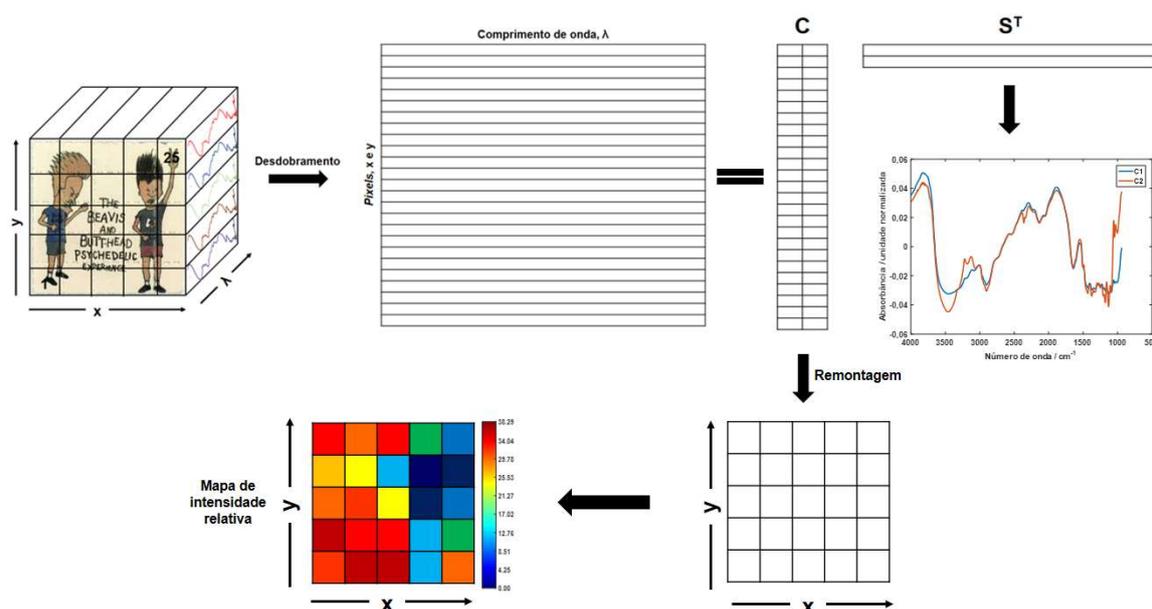
As iterações, durante o desenvolvimento do modelo MCR, iniciam a partir de uma estimativa inicial. A estimativa inicial pode consistir de uma matriz contendo os espectros puros dos componentes da mistura, quando existe conhecimento *a priori* de quem são esses componentes e se dispõe de padrões para cada um deles. Quando não se dispõe dos espectros dos componentes puros, as estimativas iniciais são obtidas por meio de algoritmos matemáticos, tal como o *Purity*, que buscam os espectros dos dados que mais se aproximariam de espectros puros. Este algoritmo seleciona as colunas da matriz original com variáveis mais puras, obedecendo o número de componentes selecionado para ser calculado para o conjunto de dados analisado. Gera-se uma estimativa inicial da intensidade relativa ou do perfil espectral (**C** ou **S**) e, após isso, se calcula o outro perfil com os dados originais, até se obter uma melhor estimativa da variância dos dados.<sup>65</sup> Além disso, caso o número de componentes da mistura não for conhecido, se faz necessário uso de ferramentas numéricas, como a decomposição de valores singulares (SVD, do inglês *Singular Value Decomposition*) que indicam o número de componentes.<sup>66</sup>

Existem semelhanças entre o MCR e o PCA, mas neste primeiro, a ortogonalidade entre as componentes não é imposta como no PCA. No MCR, é necessário que os componentes da mistura apresentem uma relação linear entre suas concentrações e intensidade de sinal, isto é, a Lei de Beer-Lambert deve ser seguida, uma vez que a decomposição matricial realizada pelo MCR gera combinações lineares das matrizes de intensidade relativa (**C**) e de espectros puros (**S**).<sup>65</sup> Desse modo, deve modelar os perfis de intensidade relativa (**C**) e os espectros puros (**S**) de acordo com as propriedades químicas, espectroscópicas e matemáticas naturais dos dados, ou seja, deve-se levar em consideração alguns aspectos dos dados espectrais, como, por exemplo, o fato de não existir valores negativos para concentração.<sup>66</sup>

Portanto, ao se realizar uma decomposição por MCR, existem algumas opções de restrições que podem ser impostas para “guiar” ou restringir o número de soluções possíveis do modelo. Uma delas é a restrição de não-negatividade, que impede o surgimento de valores negativos durante a decomposição. Essa restrição pode ser imposta para as intensidades relativas (**C**) e também para os espectros (**S**) a depender da natureza dos dados. Além da restrição de não-negatividade, outras restrições comuns que podem ser impostas são: unimodalidade, que permite somente um valor máximo por perfil, sendo geralmente aplicada para dados cromatográficos; e *closure*, que leva em consideração um balanço de massa, no qual a soma das concentrações dos componentes é um valor constante. Estas restrições, junto com a restrição de não-negatividade, são as mais aplicadas para dados espectrais.<sup>66</sup>

O MCR-ALS é um modelo bastante aplicado para a resolução de dados de mapeamento de imagens. O objetivo do uso de MCR para estes casos é promover a separação dos sinais presentes nos espectros das diferentes espécies químicas (componentes) que compõem a área mapeada, de modo que se possa analisar a forma com a qual estas espécies se distribuem na região. Para tanto, aplica-se a decomposição por MCR nas chamadas imagens hiperespectrais.<sup>67</sup>

Imagens hiperespectrais podem ser definidas como um cubo de dados obtido a partir de medidas espectrais de uma superfície bidimensional, de modo que as dimensões x e y do cubo se referem aos *pixels* da imagem e a dimensão z se refere ao espectro obtido para cada *pixel*, como é apresentado na figura 3.



**Figura 3.** Esquema de decomposição de uma imagem hiperespectral por MCR-ALS.

As imagens hiperespectrais consistem em um cubo de dados que pode ser submetido a uma deconvolução por diferentes modelos quimiométrico. Quando a deconvolução é realizada por MCR, as variáveis redundantes são eliminadas, obtendo-se novas imagens que indicam as intensidades relativas de cada componente separado pelo modelo, com informação seletiva e específica.<sup>67</sup>

O uso de imagens hiperespectrais se iniciou por volta dos anos de 1970 com o imageamento remoto e aplicação em outros campos, como na ciência de alimento, área farmacêutica e também nas ciências forenses.<sup>67</sup> Na literatura, várias são as aplicações de imagens hiperespectrais para avaliação ou identificação de substâncias proibidas, resíduos de disparo de arma de fogo e avaliação da falsificação de documentos.<sup>68-71</sup> Na avaliação de distribuição de substâncias proibidas em selos, a aplicação de imagens hiperespectrais ainda é discreta. Destaca-se o trabalho desenvolvido por Moraes e

colaboradores<sup>72</sup>, no qual quatro amostras de selos apreendidos, contendo 25I-NBOH, foram submetidos à análise de imagem por espectrometria de massas ambiente com ionização por *sonic-spray* (EASI-IMS, do inglês *Easy Ambient Sonic-Spray Imaging Mass Spectrometry*), considerada uma técnica pouco destrutiva, por manter a integridade da amostra. Os resultados apresentados pelos autores, para apenas uma amostra das quatro avaliadas, indicaram que a distribuição de 25I-NBOH no selo foi uniforme, a partir dos sinais do íon de massa/carga 414,05, característico do 25I-NBOH protonado, presente no espectro de massas obtido da superfície do selo. Os autores não realizaram análise quantitativa e a técnica aplicada não necessitou de análise multivariada, uma vez que o imageamento foi realizado com base em um pico característico do espectro de massas.

Nesta tese, será apresentada uma outra abordagem de uso de imagem hiperespectral para avaliar a distribuição de 25E-NBOMe em selos, utilizando espectroscopia no infravermelho médio, que é uma técnica que não requer preparo de amostra e também não é destrutiva, associada ao MCR para a resolução dos sinais.

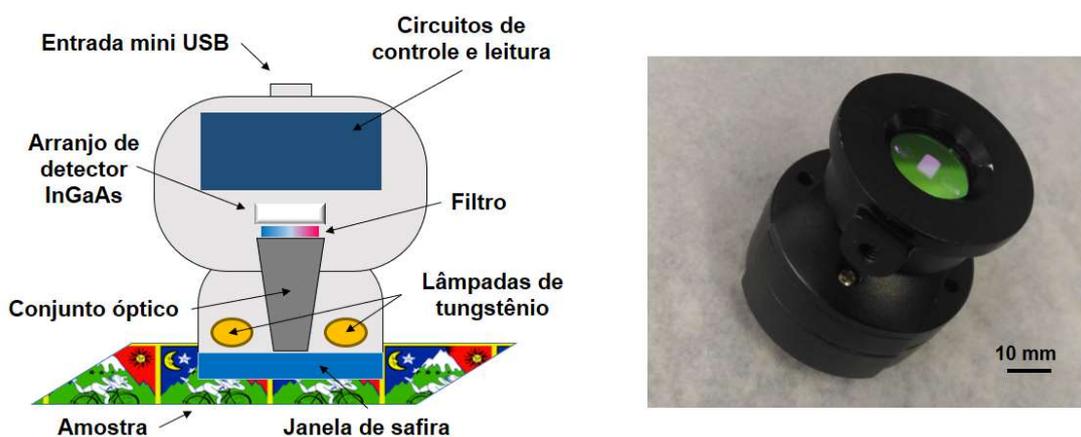
### 3. MATERIAIS E MÉTODOS

#### 3.1. Discriminação de drogas sintéticas em selos por espectroscopia no infravermelho próximo (NIR)

##### 3.1.1. Aquisição de dados

##### 3.1.1.1. Equipamentos e softwares

Para a aquisição dos espectros na região do infravermelho próximo, foi utilizado o espectrômetro portátil MicroNIR1700 (JDSU) com região espectral de 908 a 1690 nm (11013,22 a 5917,16  $\text{cm}^{-1}$ ). Este espectrômetro tem dimensões de 50 mm de diâmetro por 45 mm de altura, com massa de apenas 64 g.<sup>53</sup> Ele é dotado de uma fonte de radiação de tungstênio, capaz de emitir toda a região espectral em questão.<sup>73</sup> O detector do MicroNIR é composto por um arranjo de fotodiodos de InGaAs (Arseneto de Gálio e Índio) com 128 *pixels* e cerca de 6 nm de resolução. O equipamento pode fornecer resposta nos modos de absorbância, reflectância e transmitância. Como padrão de referência, para uma reflectância maior que 97%, utiliza-se um material à base de politetrafluoroetileno (PTFE) sinterizado (Avian WCR 2066). Na figura 4 são apresentados um esquema da seção transversal do MicroNIR e uma fotografia digital do equipamento.



**Figura 4.** Esquema da seção transversal (à esquerda) e fotografia digital do MicroNIR 1700 (à direita).

A aquisição dos espectros é realizada por meio do *software* do instrumento (MicroNIR 1. 6. 1), compatível com Windows 7, 8, 8.1 e 10. Toda a manipulação de dados foi feita com o programa MATLAB, versão R2012b (The MathWorks, Natick, MA, USA). Os modelos de calibração multivariada foram desenvolvidos utilizando-se a plataforma PLS Toolbox, versão 7. 0. 3 (Eigenvector Technologies, Manson, MA, USA).

Antes do início da aquisição de espectros, as lâmpadas de tungstênio permaneciam ligadas por cerca de dez minutos para aquecimento e estabilização do sinal (temperatura = ~50 °C). Foram definidos um tempo de integração de 2000  $\mu$ s e uma média de 100 varreduras por espectro. Realizava-se a aquisição do espectro para o padrão de reflectância e também para o escuro, direcionando o detector para qualquer ponto que não gerasse resposta por reflectância difusa.

### **3.1.1.2. Testes iniciais**

Previamente ao início de aquisição dos espectros das amostras utilizando o MicroNIR, observou-se que não havia necessidade de pré-tratamento das amostras. É importante salientar que o equipamento era colocado sobre os selos, em diferentes pontos, para a aquisição dos espectros. Entre amostras diferentes, limpava-se a base metálica e a janela de safira do MicroNIR que entrava em contato com os selos com álcool etílico hidratado 92,8 °INPM.

A aquisição dos espectros costumava ser realizada na parte da frente dos selos, que usualmente contém imagens coloridas referentes desenhos, gravuras artísticas, efeitos psicodélicos, entre outros. Também se realizava aquisição espectral na parte do verso dos selos, normalmente sem qualquer desenho. O número de replicatas foi inicialmente definido como sendo cinco em ambos os lados. É importante salientar que somente apreensões com mais de cinco microssselos unidos, isto é, sem estarem destacados, foram selecionadas, uma vez que a janela de detecção do MicroNIR tem uma área restrita, e um conjunto de selos muito pequeno poderia levar à absorção da superfície na qual os selos foram colocados. Os microssselos são definidos como a menor parte destacada dos selos, com área aproximada de 1 cm<sup>2</sup>. A figura 5 apresenta a relação entre o tamanho dos selos utilizados e o MicroNIR.



**Figura 5.** Exemplo de aquisição de espectros dos selos com o MicroNIR.

Tendo em vista os resultados obtidos nesses testes iniciais, que serão apresentados na seção 4.1.1, definiu-se um novo procedimento para a análise dos selos. Para cada amostra, composta por uma apreensão distinta de selos da Polícia Civil do Distrito Federal (PCDF), definiu-se um número de 10 replicatas espectrais, adquiridas aleatoriamente por toda região do verso dos selos. Os valores de absorvância para cada um dos 125 comprimentos de onda medidos pelo MicroNIR para todas as replicatas eram salvos em arquivo formato CSV, sendo que os sinais para cada replicata eram dispostos em colunas, como é apresentado na figura 6.

	H	I	J	K	L	M	N	O	P	Q
1	908.1	914.294	920.489	926.683	932.877	939.072	945.266	951.46	957.655	963.849
2	0.07462	0.07641	0.07772	0.07822	0.08174	0.07579	0.07436	0.07445	0.07451	0.07326
3	0.07482	0.07585	0.07477	0.07601	0.08071	0.0745	0.07201	0.0714	0.07193	0.07156
4	0.07233	0.073	0.07611	0.07428	0.07994	0.07432	0.071	0.07184	0.07254	0.07041
5	0.07272	0.07171	0.07662	0.07537	0.07976	0.0745	0.07244	0.07345	0.07231	0.07052
6	0.06639	0.06865	0.06964	0.069	0.07271	0.06764	0.06411	0.06577	0.06493	0.06345
7	0.06392	0.06351	0.0659	0.0654	0.06869	0.06252	0.06148	0.06242	0.06024	0.05903
8	0.06706	0.06656	0.06578	0.06685	0.06981	0.06344	0.0611	0.06164	0.06019	0.05794
9	0.07592	0.07482	0.07649	0.07686	0.08087	0.07563	0.07338	0.07474	0.07384	0.07045
10	0.08108	0.08224	0.08574	0.08455	0.08935	0.08345	0.08196	0.08385	0.08364	0.08009
11	0.08715	0.0872	0.08998	0.08988	0.09408	0.08818	0.08716	0.0875	0.08653	0.08435

**Figura 6.** Matriz de dados de absorvância gerada pelo MicroNIR. A legenda dos comprimentos de onda é apresentada na primeira linha e as absorvâncias das 10 replicatas nas dez linhas seguintes.

### 3.1.2. Amostras utilizadas e aquisição dos espectros NIR

Todas as amostras utilizadas neste trabalho foram apreendidas pela PCDF, analisadas na Seção de Perícias e Análises Laboratoriais (SPAL) do Instituto de Criminalística (IC) da PCDF por meio de métodos de referência.<sup>5</sup> Foram selecionadas 101 amostras apreendidas entre os anos de 2009 e 2017, que apresentavam, no mínimo, 5 microselos unidos, devido ao tamanho da janela do detector do MicroNIR, como já explicado anteriormente. Do verso de cada amostra, foram adquiridos 10 espectros em pontos variados para aumentar a representatividade. Das 101 amostras, 15 apresentavam em seus laudos periciais conter 25B-NBOMe. Outras 4 continham 25C-NBOMe e 17 continham 25I-NBOMe. Estas três substâncias ilícitas foram agrupadas em um conjunto denominado NBOMe. Outras 51 amostras continham 25I-NBOH e 3 continham 25B-NBOH, as quais foram agrupadas em um conjunto denominado NBOH. Por fim, 11 amostras de um grupo denominado OUTROS apresentavam conteúdo variado (1 de acetaminofeno, 1 de 2C-I, 2 de LSD, 5 de DOC e 2 de MDMA com LSD), diferente dos grupos anteriores. Todas as amostras foram previamente analisadas por GC-MS ou LC-QTOF-MS na SPAL. A tabela 1 apresenta as quantidades de amostras utilizadas nas fases de treinamento e validação dos modelos.

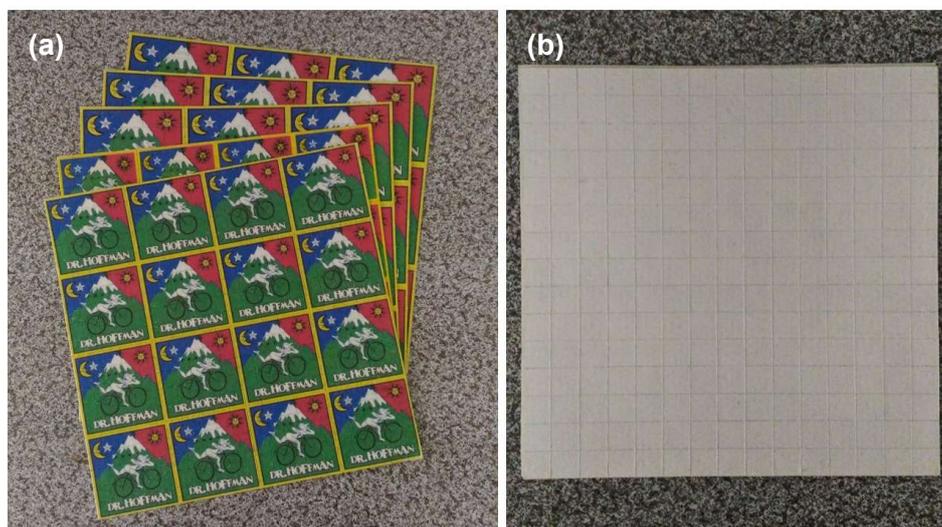
**Tabela 1.** Quantidades e composições de amostras de selos para os conjuntos de treinamento e validação do primeiro modelo de classificação multivariada.

Amostra	Quantidade		Composição *
	Treinamento	Validação	
NBOMe	24	12	25B-NBOMe, 25C-NBOMe e 25I-NBOMe
NBOH	35	19	25B-NBOH e 25I-NBOH
Outros	7	4	2C-I, acetaminofeno, LSD, DOC, MDMA

\*Determinada por métodos de referência.<sup>5,6</sup>

Como forma de se obter uma classe de selos que não contém nenhum tipo de NPS, foram adquiridas pela internet cinco folhas de selos sem nenhum tipo de substância impregnada para utilização como branco. As folhas de selos utilizadas como branco consistiam em três lotes distintos, sendo três folhas de um mesmo lote e outras duas de dois lotes distintos. De cada folha de selos foram adquiridos 50 espectros no verso de forma aleatória. Esses espectros foram utilizados para o desenvolvimento do primeiro modelo de classificação, a ser explicado na seção 3.1.5. Parte destes selos também foi submetida à análise por LC-QTOF-MS para confirmar a ausência de

substâncias impregnadas. Na figura 7 são apresentados os selos sem drogas adquiridos para elaboração do primeiro modelo. Apesar de serem de três lotes distintos, todos eles têm uma ilustração e tamanho de selos que diferem ligeiramente entre os lotes. A ilustração nos selos é referente a Albert Hoffmann, químico suíço que descobriu os efeitos do LSD em 1943.<sup>74</sup>



**Figura 7.** (a) Fotografia da parte da frente dos selos sem drogas adquiridos pela internet para desenvolvimento do primeiro modelo de classificação multivariada. (b) Fotografia do verso dos selos sem drogas (à direita).

Além desses selos, cinco folhas de papel mata-borrão de três lotes distintos e duas folhas de filtro quantitativo (Unifil, 85 g/m<sup>2</sup>) de dois lotes diferentes foram inseridas na classe branco. De cada papel foram adquiridos 50 espectros. É importante destacar que a composição, espessura e textura dos papéis mata-borrão e de filtro utilizados para compor a classe branco são similares às dos selos adquiridos pela internet. Este fato justifica a inclusão desse grupo de amostras na classe para aumentar a variabilidade dos dados e a representatividade do conjunto.

### 3.1.3. Pré-processamento

Diferentes tipos de pré-processamento e diferentes combinações destes foram testados, dentre eles: Alisamento pelo algoritmo Savitsky-Golay, primeira e segunda derivada, SNV, MSC e dados centrados na média. Esses tipos de pré-processamentos são frequentemente utilizados em dados espectrais no infravermelho. O pré-processamento ou a combinação de pré-processamentos que apresentou menores taxas de erros de classificação foi escolhido para compor o modelo.

### 3.1.4. Validação cruzada

O conjunto de dados das amostras foi composto por um número de espectros múltiplo de 10, devido à quantidade de replicatas espectrais das amostras apreendidas. Devido a isso, escolheu-se como método de validação cruzada as janelas venezianas (do inglês *venetian blinds*), com um número de janelas referente ao número de espectros do conjunto de treinamento dividido por 10. Os valores de RMSECV obtidos após a validação cruzada foram os parâmetros principais para a escolha de variáveis latentes para o modelo.

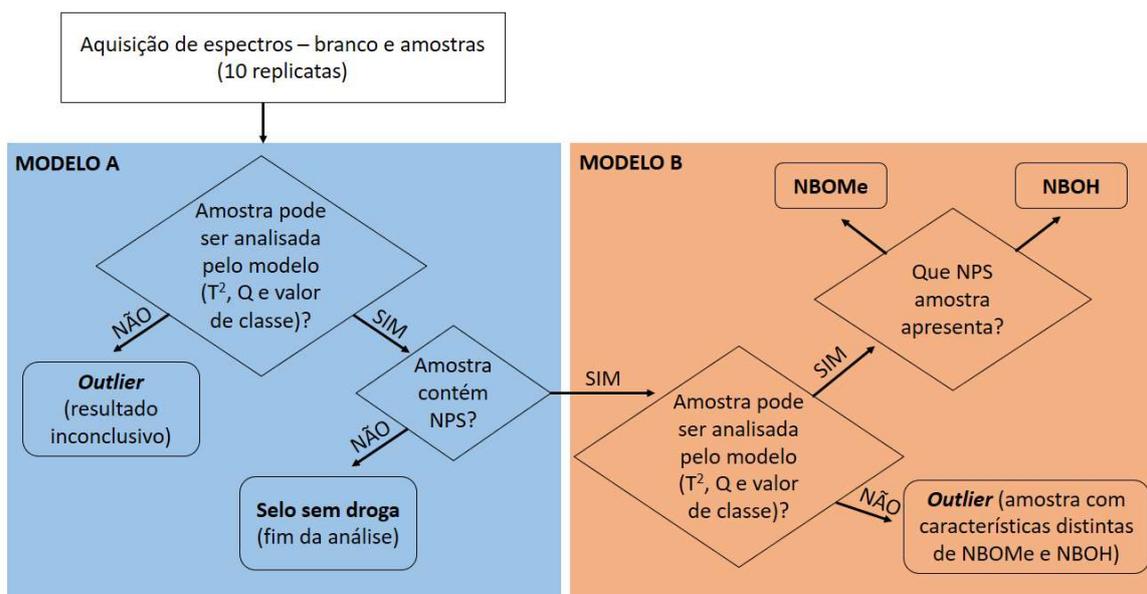
### 3.1.5. Modelos de calibração multivariada

Os métodos quimiométricos de classificação utilizados neste trabalho foram o PLS-DA e o SIMCA. Dois modelos diferentes foram elaborados para essas duas ferramentas quimiométricas. As amostras foram organizadas em ordem cronológica de apreensão, de modo que a cada duas que eram separadas para o conjunto de treinamento, uma era selecionada para o conjunto de validação. Essa divisão foi realizada considerando-se cerca de dois terços de todas as amostras para a fase de treinamento e um terço para a fase de validação. O primeiro modelo de classificação – **modelo A** – é composto por duas classes, amostras que contêm drogas e amostras que não contêm drogas (branco). A classe que contêm drogas foi composta por 1010 espectros, referentes às 101 amostras (36 de NBOMe, 54 de NBOH e 11 de outras drogas). Essas amostras foram divididas em duas partes, uma para compor a fase de treinamento do modelo e outra para compor a fase de validação, como é apresentado na tabela 1 da seção 3.1.2. A classe que não contém drogas foi composta pelos 600 espectros obtidos dos selos adquiridos pela internet e de papel filtro quantitativo, sendo que 350 espectros foram utilizados na fase de treinamento e 250 espectros na fase de validação do modelo. No conjunto de treinamento, utilizaram-se espectros de cinco lotes distintos (100 espectros de duas folhas de selos do lote que contém três folhas e 50 espectros de um dos lotes contendo uma só folha e 200 espectros de quatro papéis de filtro quantitativo de lotes diferentes). Como conjunto de validação utilizaram-se espectros de quatro lotes distintos (50 espectros da terceira folha de selos do lote que contém três folhas, 50 espectros do outro lote que contém só uma folha e que não foi utilizado na fase de treinamento e outros 150 espectros de três papéis de filtro quantitativo dos lotes não utilizados na fase de treinamento). Essa divisão entre as duas fases do modelo foi realizada de modo a garantir a representatividade dos dois conjuntos de dados.<sup>11,50</sup>

O segundo modelo de classificação – **modelo B** – também foi composto por duas classes, a classe de NBOMe e a classe de NBOH. Todas as amostras que contém drogas do modelo A foram utilizadas no desenvolvimento do modelo B, com exceção das 11 amostras de selos contendo outras drogas. A classe dos NBOMe foi composta por 360 espectros, referentes às 36 amostras de NBOMe. Já a classe dos NBOH foi composta por 540 espectros de 54 amostras de NBOH. Da mesma forma que no modelo A, essas amostras foram divididas em duas partes para compor as fases de treinamento e validação, respeitando a proporção de dois terços e um terço, respectivamente. É importante destacar que quando uma amostra foi destinada ao conjunto de treinamento ou validação, todas as suas replicatas foram incluídas nesse mesmo conjunto.

Além dos conjuntos de validação descritos nos parágrafos anteriores, o modelo B também teve sua eficiência avaliada com relação a amostras com características distintas daquelas presentes no conjunto de treinamento, o que corresponde a uma validação com uma classe não modelada na fase de treinamento. Esta avaliação é essencial para se concluir se o modelo gerado é capaz de identificar uma classe de características distintas das amostras presentes na fase de treinamento como anômala, através dos valores de resíduos Q, Hotelling  $T^2$  (influência) e erros nos valores estimados de classe. No modelo B, esse novo conjunto validação foi composto pelas amostras de selos contendo outras drogas (LSD, MDMA e DOC).

A figura 8 indica um fluxograma para a aplicação dos modelos A e B por PLS-DA em análises forenses. Em uma situação forense hipotética, esses modelos seriam utilizados como métodos de triagem nas apreensões de selos. A aplicação do modelo B é dependente do resultado obtido no modelo A. Inicialmente, seria definido se a amostra de selo contém ou não droga (modelo A) e, em seguida, as amostras contendo drogas seriam analisadas no modelo B, para serem classificadas como NBOMe, NBOH ou contendo outras drogas, no caso de serem observados valores de Hotelling  $T^2$  (influência) e resíduos Q acima de seus limites, considerando 99,9% de confiança.

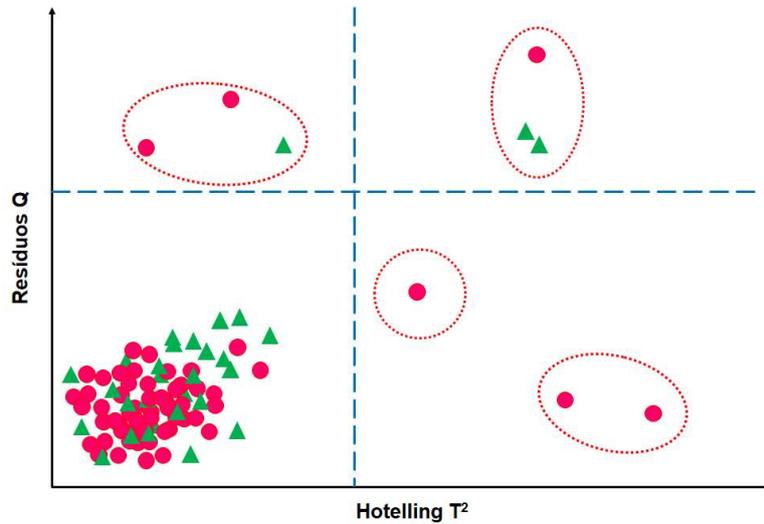


**Figura 8.** Fluxograma de análise multivariada de selos.

### 3.1.6. Rotina de detecção de outliers

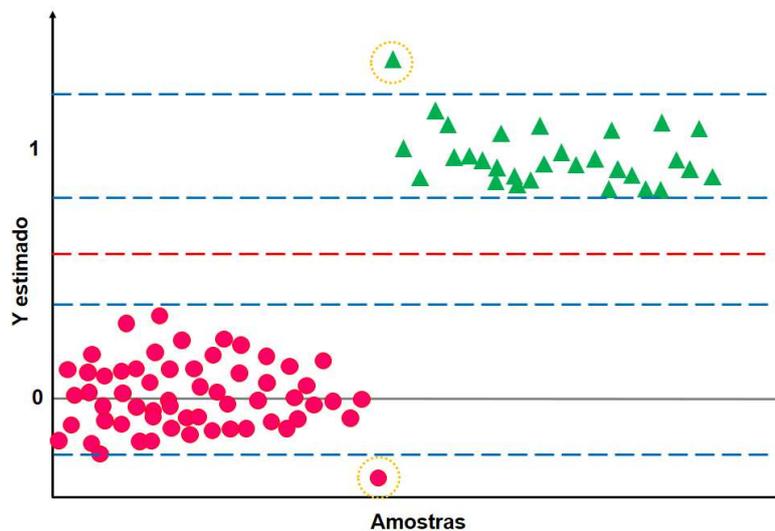
Devido a heterogeneidade das amostras de selos, um procedimento de eliminação de *outliers* se fez necessário. Para tanto, utilizou-se uma rotina de detecção de *outliers* desenvolvida por Soares e colaboradores<sup>50</sup>, que foi adaptada para este trabalho.

No trabalho de Soares e colaboradores, a área de *outliers* foi definida pela intersecção dos parâmetros influência Hotelling  $T^2$  e resíduos Q no gráfico gerado por eles. Deste modo, são amostras anômalas as medidas que se localizam no quadrante superior direito do gráfico, que é definido por um nível de confiança. Para este trabalho, a restrição de amostras anômalas foi ampliada para se certificar que amostras anômalas não fossem analisadas nos modelos desenvolvidos, o que poderia gerar resultados falso positivos ou negativos. Para tanto, considerou-se como área de *outliers*, a área de união dos parâmetros  $T^2$  e Q, referente aos quadrantes superiores esquerdo e direito e também ao inferior direito, que correspondem às regiões cujos valores de  $T^2$  ou Q são maiores que seus valores limites. Na figura 9 é apresentado um gráfico ilustrativo que exemplifica o uso dos parâmetros  $T^2$  e Q para a retirada de *outliers*.



**Figura 9.** Gráfico de influência Hotelling e resíduos Q com indicação de *outliers* na área de união dos parâmetros (amostras dentro das áreas delimitadas em vermelho são consideradas *outliers* e excluídas do conjunto de dados).

Além desse critério de exclusão de *outliers*, os limites de classe calculados também foram utilizados nos modelos desenvolvidos. A figura 10 apresenta outro gráfico ilustrativo que indica a forma com a qual os limites de classe delimitam o espaço de valores estimados. As amostras delimitadas em laranja consistem em prováveis *outliers*, uma vez que seus valores de classe estimados estão além dos limites das classes.



**Figura 10.** Gráfico de dispersão de amostras com indicação dos limites de classes (retas azuis) e do limite de discriminação (reta vermelha).

Neste trabalho, qualquer tomada de decisão sobre uma amostra (se é um *outlier*, falso positivo ou negativo) é tomada com base no resultado da maioria das replicatas das amostras. Ou seja, se seis ou mais replicatas fornecem um determinado resultado, este é considerado para a amostra como um todo.

### 3.1.7. Validação

Para validação dos métodos, nas fases de treinamento e validação foram avaliadas as seguintes figuras de mérito: taxa de falsos positivos, taxa de falsos negativos e a taxa de eficiência. Adicionalmente, foi avaliada a quantidade de amostras anômalas eliminadas durante as três etapas de exclusão de *outliers* na fase de treinamento e na única exclusão na fase de validação.

Para os modelos desenvolvidos por SIMCA, foi também incluída a taxa de amostras com resultado inconclusivo (TIN), uma vez que esta ferramenta pode classificar uma mesma amostra em várias classes. A taxa de inconclusivos para cada classe é definida pela seguinte equação:

$$TIN = \frac{IN}{IN+T_C} \times 100 \quad (\text{Equação 13})$$

Onde IN é o número de amostras com resultado inconclusivo e T<sub>C</sub> é o número total de amostras da classe, previamente definido.

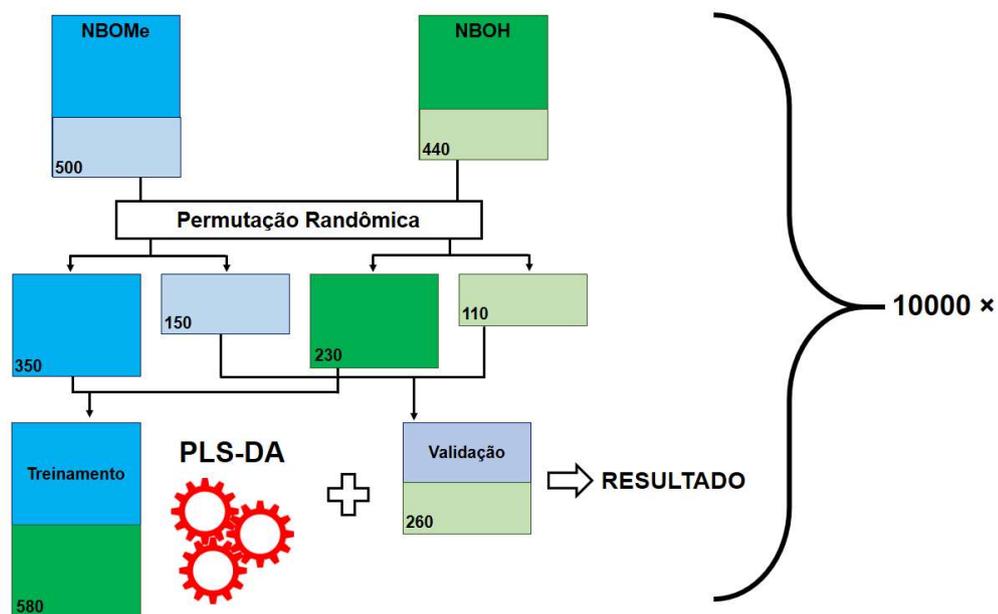
É importante considerar que as amostras/apreensões que não foram classificadas em nenhuma classe são consideradas falsos negativos, enquanto as amostras classificadas em ambas as classes são consideradas inconclusivas, gerando diferentes taxas de resultado inconclusivo a depender da classe. Adicionalmente, ao se calcular as taxas de eficiências das classes, utilizando SIMCA, considerou-se também a influência da taxa de resultados inconclusivos, de modo que a eficiência, para este caso, é dada pela equação 14.

$$TEF = 100 - (TFP + TFN + TIN) \quad (\text{Equação 14})$$

### 3.1.8. Avaliação da robustez do modelo B por permutação de amostras

Tendo em vista que cada amostra de NBOMe e NBOH consiste de uma apreensão distinta da PCDF, que possui em geral figuras nos selos diferentes entre si e que podem ter passado por processos de incorporação das drogas distintos, tem-se

uma grande heterogeneidade nessas amostras. Por essa razão, existe uma possibilidade de que a taxa de eficiência medida no modelo B possa depender da forma como os conjuntos de treinamento e validação foram divididos. A fim de avaliar a robustez do modelo B por PLS-DA (Discriminação entre NBOMe e NBOH) quanto a essa divisão, foi elaborado um programa em linguagem MATLAB que realizou permutações entre as amostras pertencentes ao conjunto de treinamento e as do conjunto de validação (sendo cada amostra constituída dos 10 espectros de suas replicatas). A princípio, como exposto na seção 3.1.5, as amostras foram organizadas em ordem cronológica de apreensão e divididas em um conjunto de treinamento e outro de validação. Participaram deste processo de avaliação somente as amostras que não foram excluídas como *outliers*. Deste modo, modelos PLS-DA diferentes foram gerados a partir de 10000 conjuntos de treinamento e validação permutados randomicamente, contendo 23 e 11 amostras da classe NBOMe e 35 e 15 amostras da classe NBOH, respectivamente para treinamento e validação. Cada modelo gerado forneceu um valor de eficiência para a fase de treinamento e para a fase de validação. A figura 11 apresenta um esquema que ilustra a realização das permutações para montagem dos conjuntos de treinamento e validação do modelo B.



**Figura 11.** Esquema de permutação das amostras de NBOMe e NBOH para construção dos conjuntos de treinamento e validação para os 10000 modelos (Os números nos cantos inferiores dos quadros que representam as matrizes indicam o número de linhas das matrizes, ou seja, número de amostras multiplicado por 10 replicatas)

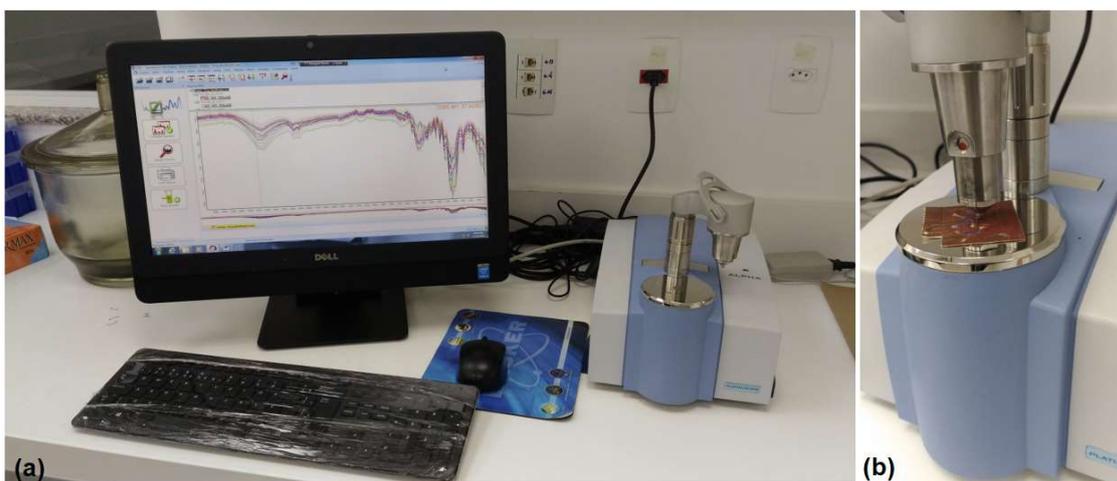
## **3.2. Modelos de classificação por MIR**

### **3.2.1. Aquisição de dados**

#### **3.2.1.1. Equipamento e *software***

A aquisição dos espectros na região do infravermelho médio foi realizada por meio do espectrômetro ATR-FTIR Alpha II (Bruker), equipado com uma fonte de tecnologia própria (CenterGlow™), acessório de ATR com cristal de diamante e detector de sulfato de triglicina deuterado (DTGS) compreendendo uma região de número de onda de 4000 a 600  $\text{cm}^{-1}$ . Este tipo de equipamento é comumente encontrado em laboratórios forenses, devido a sua facilidade de uso, bem como a eficiência na identificação de compostos, devido às bibliotecas de espectros atualizadas que podem ser utilizadas. As análises de selos foram conduzidas com resolução de 4  $\text{cm}^{-1}$  e 24 *scans* por espectro. Todos os espectros adquiridos foram convertidos para absorbância previamente à análise multivariada.

As amostras de apreensões utilizadas para a aquisição de espectros no MIR apresentavam números de microsselos diversificados e eram, até mesmo, bem menores do que aquelas utilizadas na análise por NIR. A utilização de amostras com menor número de microsselos foi possível devido ao menor tamanho da janela de aquisição do Alpha II. Deste modo, amostras contendo no mínimo dois microsselos, já poderiam ser utilizadas. Estas amostras foram posicionadas com o verso voltado para o cristal de ATR, de modo que a obtenção do espectro fosse realizada na face com ausência de imagens. De cada amostra de selos, foram obtidos 5 espectros em pontos distintos e aleatórios. A escolha deste número menor de replicatas em relação ao estudo anterior se deu, principalmente, pelo tempo total da análise, que abrange o tempo que o equipamento leva para a obtenção de cada espectro (cerca de 30 s), abertura do acessório, reposicionamento da amostra, fechamento do acessório, salvamento dos arquivos e limpeza do acessório para a próxima amostra, perfazendo aproximadamente 15 minutos para a análise de uma amostra realizando cinco replicatas. Um número maior de espectros poderia tornar este processo muito mais longo, devido ao número de amostras utilizadas nesta parte do trabalho, a ser abordado na seção 3.2.2. Na figura 12 é apresentado o equipamento utilizado nas análises por infravermelho médio.



**Figura 12.** (a) Visão geral do equipamento do Instituto de Criminalística da Polícia Civil do Distrito Federal. (b) Posicionamento do selo na célula de leitura do Alpha II.

O espectrômetro Alpha II é acompanhado do *software* OPUS (Bruker, versão 7.5), que permite a aquisição de espectros, análise com bibliotecas referentes à área forense e análises multivariadas, caso a extensão QUANT2 esteja habilitada no *software*. Este *software* foi utilizado na elaboração de modelos de análise multivariada das amostras nesta parte do trabalho, de modo a se criar um procedimento operacional padrão que pudesse ser aplicado nas análises de rotina do laboratório da PCDF para a triagem dos selos. Os modelos foram elaborados por PLS-DA, com algumas mudanças em relação àqueles desenvolvidos por NIR, devido à região espectral, ao *software* e ao modo de se realizar a exclusão de *outliers*.

O *software* OPUS é dotado da extensão QUANT2 para análise multivariada. Esta extensão pode ser utilizada para gerar modelos PLS com um conjunto de dados já disponível no computador ou no ambiente do *software*. Como o QUANT2 só realiza a análise regressão multivariada (PLS), a determinação do limite de discriminação para os modelos PLS-DA foi calculada manualmente através do uso de uma planilha eletrônica, onde os cálculos foram realizados conforme descrito por Soares e colaboradores.<sup>50</sup> Por meio de uma simples planilha eletrônica, os resultados de predição para as amostras testadas foram relacionados com duas classes: Branco e NPS (modelo A) e NBOMe e NBOH (Modelo B).

Nesta extensão QUANT2 é possível carregar os dados de treinamento e programar o desenvolvimento de vários modelos diferentes, variando o pré-processamento utilizado e região espectral. Este processo é realizado de forma automática no decorrer de algumas horas. O *software* desenvolve modelos variando tipos e combinações de pré-processamento, bem como regiões espectrais de trabalho.

Após desenvolvê-los, os vários modelos diferentes são organizados em modo crescente de erros de predição para as amostras de treinamento.

Os modelos são então avaliados, e o que apresenta pré-processamento ou combinação de pré-processamentos com menor erro de previsão e, simultaneamente, um número menor de variáveis latentes é escolhido. Além disso, deve-se observar as regiões espectrais, a fim de que sejam coerentes com as bandas de infravermelho que diferenciam as classes. Apesar dos diversos estudos que apontam que a seleção de variáveis fornece modelos com menores erros,<sup>75,76</sup> para conjuntos com amostras muito heterogêneas e com sobreposição espectral significativa, os melhores modelos desenvolvidos podem ser aqueles que englobam toda a região espectral.

### **3.2.1.2. Modo de aquisição e conversão dos espectros**

Apesar do modo de aquisição espectral pelo Alpha II ser por reflectância total atenuada, os espectros são armazenados no modo de transmitância, como é comum observarmos para o infravermelho médio. Porém, para o desenvolvimento dos modelos todos os espectros obtidos foram convertidos para o modo de absorbância no próprio *software* do equipamento.

### **3.2.2. Amostras**

Para o desenvolvimento dos modelos A e B por infravermelho médio foram selecionadas 154 amostras apreendidas no período entre 2009 e 2018, sendo que muitas dessas amostras já tinham sido utilizadas para o desenvolvimento dos modelos por NIR. Foram excluídas desse conjunto as amostras que apresentavam imagens no verso dos selos, e que foram consideradas *outliers* nos modelos desenvolvidos por NIR. Todas essas amostras foram analisadas por GC-MS ou LC-QTOF-MS, tendo suas composições identificadas e confirmadas por laudo. Das 156 amostras, 61 correspondiam a NBOMe, 78 a NBOH e 15 correspondiam a outras drogas (DOC, LSD e MDMA), conforme descrito na tabela 2.

**Tabela 2.** Número e conteúdo de amostras utilizadas nas fases de treinamento e validação dos modelos A e B. Números entre parêntesis indicam o número total de espectros para cada grupo de amostras.

Amostra	Treinamento		Validação		Conteúdo
	A	B	A	B	
NBOME		40 (200)		21 (105)	25B-NBOMe, 25C-NBOMe, 25I-NBOMe
NBOH	102 (510)	51 (255)	52 (260)	27 (135)	25B-NBOH, 25I-NBOH
Outras drogas		11 (55)		4 (20)	LSD, DOC, MDMA
Branco	6 (150)	---	3 (75)	---	Sem droga

Para o modelo A, também foram utilizadas as cinco amostras de selos sem droga adquiridas pela internet. É importante lembrar que a ausência de substâncias impregnadas foi confirmada pela análise por GC-MS. Além destas amostras, duas folhas de papel mata-borrão de dois lotes distintos e duas folhas de papel de filtro (Unifil, 85 g/m<sup>2</sup>), também de lotes distintos, foram inseridas no conjunto de dados para fazer parte da classe sem drogas. Todas as amostras foram divididas em um grupo de treinamento (contendo cerca de dois terços das amostras) e validação (contendo cerca de um terço das amostras).

No modelo B, somente as amostras apreendidas e que apresentaram substâncias impregnadas foram utilizadas. Estas foram divididas em duas classes, NBOMe e NBOH e também nos dois conjuntos, treinamento e validação, conforme foi feito para o modelo A. Para a análise de rotina, 19 amostras apreendidas que não fizeram parte do conjunto de treinamento e validação foram analisadas pelos modelos A e B e, depois, os resultados obtidos na análise multivariada foram confrontados aos resultados obtidos por GC-MS e LC-QTOF-MS para estas amostras.

### 3.2.3. Escolha do melhor modelo

De modo semelhante ao realizado na região do infravermelho próximo, desenvolveram-se dois modelos de discriminação por PLS-DA, os modelos A e B, com os mesmos propósitos de discriminação. Como citado na seção 3.2.1.1, vários modelos foram gerados pela extensão QUANT2 do OPUS. Estes foram organizados em modo crescente de erro de predição. A escolha do melhor modelo (tanto A quanto B) se baseou nos menores valores de erros, associados ao conjunto de pré-processamento de dados mais coerente, à região espectral e o número de variáveis latentes do modelo.

### 3.2.4. Exclusão de *outliers*

Após a escolha dos melhores modelos (A e B) desenvolvidos na extensão QUANT, foi realizada a exclusão de *outliers*. Na fase de treinamento dos modelos, foram realizadas duas etapas de exclusão de *outliers* e na fase de validação, apenas uma. Os *outliers* foram identificados por meio de três critérios: distância de Mahalanobis, resíduos espectrais e resíduos na estimativa de classe, conforme indica o manual QUANT do OPUS.<sup>77</sup> Para cada amostra, o modelo fornece o valor de cada um desses três critérios. E, para cada critério há um valor limite que é calculado. Estes valores limites são essenciais para caracterizar uma amostra como sendo ou não um *outlier*.

Quando os valores de probabilidade de resíduos espectrais são plotados contra as distâncias de Mahalanobis das amostras, observa-se um gráfico semelhante ao gráfico de influência e resíduos ( $Q \times T^2$ ), com valores limites que definem a região dos *outliers*. Para a probabilidade de resíduos espectrais, esse valor é de 99% e valor limite para a distância de Mahalanobis é definido pelo software no desenvolvimento de cada modelo. Contudo, não constam na documentação do software o nível de confiança adotado para a distância de Mahalanobis. Se uma amostra apresenta, ao menos, valor de um dos critérios acima dos limites definidos para eles no modelo, ela é considerada um *outlier* e excluída do conjunto dentro das rodadas de exclusão.

Todas as amostras utilizadas nas fases de treinamento e validação fornecem valores para estes três critérios, de modo que aquelas que apresentam valores maiores que os limites para cada um são identificadas e eliminadas manualmente. Estes critérios também são utilizados em análises de rotina a fim de determinar se uma amostra apresenta um espectro compatível com o conjunto de dados utilizado na fase de treinamento, caso o resultado seja inconclusivo.

### 3.2.5. Análise de amostras desconhecidas

Os modelos desenvolvidos por ATR-FTIR foram submetidos à análise de rotina utilizando um procedimento operacional padrão (POP) elaborado para aplicação do método na PCDF (Anexo I). Dezenove amostras que não participaram das fases de treinamento e validação foram utilizadas para verificação da eficiência na aplicação em rotina no laboratório da PCDF. Os espectros para cada uma dessas amostras foram adquiridos e convertidos para absorvância no próprio *software*. Após isso, eles foram carregados como grupo de validação nos modelos otimizados (isto é, modelos com *outliers* já excluídos), gerando valores de predição, distância de Mahalanobis e resíduos espectrais. Estes valores foram então transferidos para a planilha eletrônica

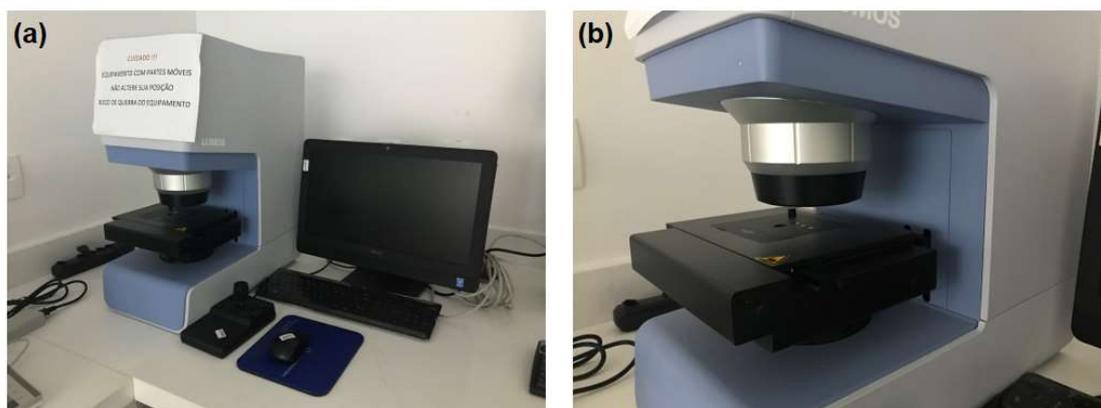
desenvolvida, que apresentou os resultados de classificação como branco ou NPS (modelo A), NBOMe ou NBOH (modelo B) e *outlier* (resultado inconclusivo em ambos os modelos).

### **3.3. Avaliação da distribuição de NPS por imagem hiperespectral e MCR-ALS**

Os exames preliminares de amostras apreendidas são atividades corriqueiras nos departamentos de polícia. Várias características são analisadas visualmente nas apreensões e estas são registradas também nos laudos após as análises. Os exames preliminares de selos normalmente estão relacionados com a identificação da quantidade de microsselos, tamanho e imagens impressas. A análise por LC ou GC-MS é então realizada com um ou dois microsselos coletados na amostra.

#### **3.3.1. Equipamento e softwares**

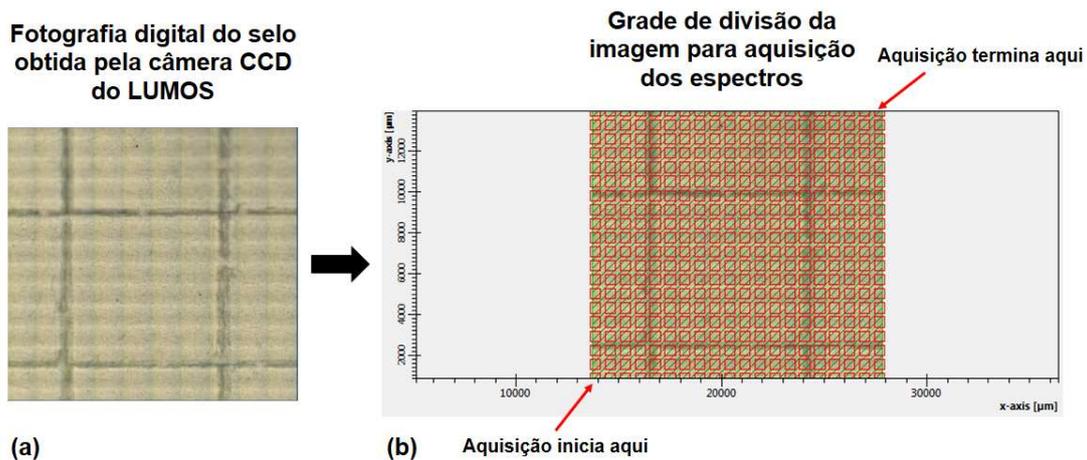
Considerando que a distribuição das substâncias nos selos não é homogênea, e para aumentar a sensibilidade na análise, um exame preliminar que considere o mapeamento do selo, de modo a indicar áreas com maior concentração de substâncias impreganadas, pode ser uma alternativa viável. Para tanto, utilizou-se o equipamento LUMOS (Bruker), que é um microscópio autônomo com espectrômetro de infravermelho integrado que realiza mapeamento em três modos de aquisição: reflectância difusa, reflectância total atenuada e transmitância. Este equipamento é dotado de uma câmera CCD (*Charge-Coupled Device*) digital que captura imagens em alta resolução com ampliação de até 32 vezes. Possui um detector fotocondutivo MCT (Mercúrio, Cádmiio e Telúrio) que gera sinal conforme a variação da condutividade do material dependendo da radiação a ele exposta. O equipamento LUMOS está disponível no Instituto de Criminalística da Polícia Civil do Distrito Federal, e uma imagem dele é apresentada na figura 13.



**Figura 13.** (a) Fotografia do microscópio de infravermelho LUMOS com *joystick* e computador. (b) Visão aproximada da plataforma de suporte e da câmera do equipamento.

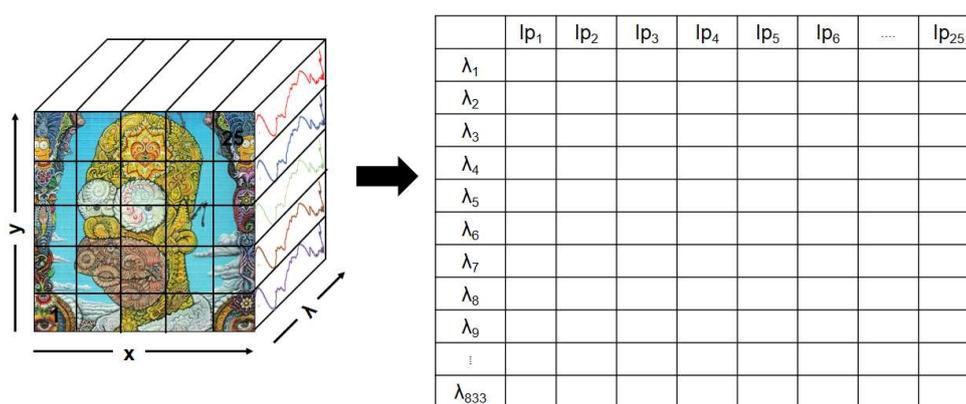
O LUMOS é controlado pelo *software* OPUS (Bruker), apresentando todas as opções de controle que o Alpha II, utilizado para o desenvolvimento dos modelos por infravermelho médio, bem como a disponibilidade de bibliotecas de espectros de infravermelho. O LUMOS pode, assim como o Alpha II realizar medidas de reflectância total atenuada, mas o objetivo principal de seu uso está no mapeamento de superfícies. Para tanto, este equipamento conta com uma câmara que deve ser preenchida com nitrogênio líquido, de modo que a o detector se mantenha resfriado durante todo o processo de leitura, que, dependendo da configuração do analista, pode levar algumas horas. No caso deste trabalho, o tempo máximo de operação do equipamento para uma amostra foi de cerca de 5 horas, com aquisição de 625 espectros em um mapeamento.

Inicialmente, uma imagem digital é obtida da superfície da amostra. Para que o mapeamento seja realizado, deve-se definir uma região a ser mapeada nesta superfície, por meio de marcação de pontos de fronteira. Após este passo, é definido o número de *scans* para cada espectro e a resolução espectral. Estes parâmetros foram definidos como 32 *scans* e  $8\text{ cm}^{-1}$  de resolução. Esta escolha foi baseada em testes iniciais que não indicaram altos níveis de ruído espectral e também no tempo total de aquisição de dados. Quanto maior a resolução e número de *scans*, maior o tempo de aquisição. Em seguida, é definida a resolução da imagem, isto é, quantos espectros serão obtidos em toda a superfície a ser mapeada, por meio de uma grade fixada ao se selecionar divisões em x e y da imagem digital previamente obtida. A obtenção dos espectros se inicia no primeiro pixel do lado inferior esquerdo da imagem, seguido da esquerda para a direita, de baixo para cima até a última divisão definida. Este processo é apresentado na figura 14.



**Figura 14.** (a) Fotografia digital obtida pela câmera CCD do LUMOS do verso de parte de um selo e (b) grade de divisão da imagem obtida para aquisição dos espectros com indicação de sua direção.

Pela figura 14 é possível perceber que o tamanho de um microselo é de aproximadamente  $8000 \mu\text{m} \times 8000 \mu\text{m}$ , de modo que um selo completo, contendo 25 microselos, apresenta um tamanho de  $40 \text{ mm} \times 40 \text{ mm}$ . O arquivo gerado após o mapeamento do selo pode ser armazenado em vários formatos, porém foi utilizado o formato “.dpt” neste trabalho, para que estes pudessem ser exportados e utilizados em ambiente MATLAB. A figura 15 apresenta o formato final do arquivo gerado, contendo os espectros para cada divisão da imagem nas colunas de uma matriz. É importante salientar que a aquisição de espectros foi realizada no verso dos selos. A figura 15 apresenta um esquema com a parte da frente do selo por, meramente, ilustração.



**Figura 15.** Esquema que representa a forma pela qual se obtém o arquivo “.dpt” de um mapeamento no LUMOS.  $lp$  são as subdivisões da imagem (imagem do pixel) definidas pela grade e  $\lambda$  representa os sinais para cada intervalo espectral (833 intervalos, devido à resolução de  $8 \text{ cm}^{-1}$ ).

Após a obtenção destes arquivos, utilizou-se uma função desenvolvida em ambiente MATLAB para que estes pudessem ser convertidas em um cubo de dados. Isto se fez necessário porque a organização na forma de matriz no qual os espectros foram exportados no LUMOS não era compatível com a interface de análise hiperespectral utilizada, o HYPER-Tools, versão 1. Esta interface é um pacote de livre acesso executado em ambiente MATLAB, desenvolvida por Mobaraki e Amigo.<sup>78</sup> Esta interface realiza a análise multivariada de imagens hiperespectrais, por meio de diferentes modelos de decomposição. Devido à compatibilidade com o HYPER-Tools, toda a análise de imagens foi realizada utilizando-se a versão R2016a do MATLAB.

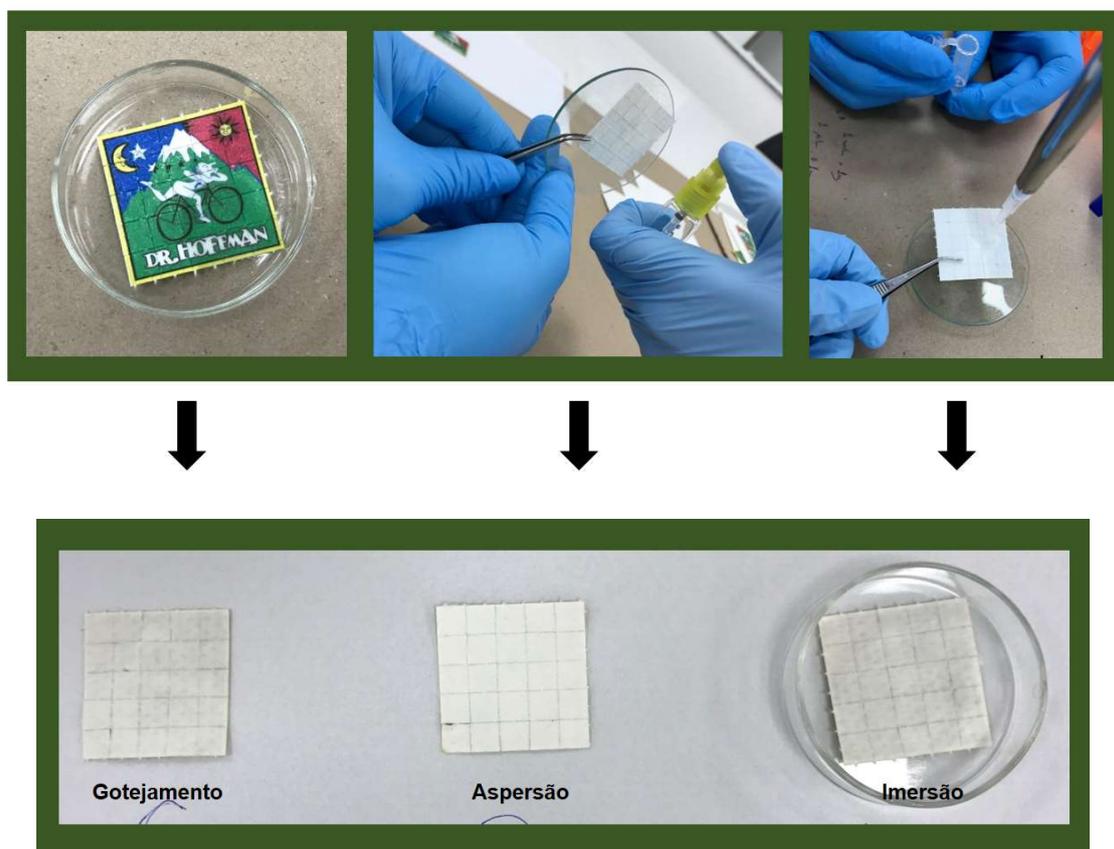
### **3.3.2. Amostras simuladas**

Para avaliação da distribuição de drogas em um conjunto de microsselos (ou microsselo) em amostras de conteúdo conhecido, se fez necessária a produção de amostras simuladas. Para tanto, três amostras contendo 25 microsselos cada (5×5) sem drogas foram impregnadas com 25E-NBOMe (Cayman Chemicals) de três modos diferentes: por gotejamento, imersão e aspersão. Estes modos de impregnação foram escolhidos por serem os mais prováveis para a adição dessas substâncias em selos.

Para as três amostras, pesaram-se 39,9 mg de 25E-NBOMe, que foram dissolvidas em 1,6 mL de metanol (Dinâmica), fornecendo uma solução de concentração 24,9 g mL<sup>-1</sup>. A massa de 25E-NBOMe utilizada foi baseada na experiência de peritos criminais da PCDF ao analisar selos. É comum observar uma concentração deste tipo de substância de 2 a 5% m/m em cada microsselo (cerca de 500 µg por microsselo)<sup>39</sup>, fornecendo cerca de 12,5 mg de substância num conjunto contendo 25 microsselos, uma vez que a massa média de cada selo é de 400 mg. O volume utilizado para dissolução da substância foi testado anteriormente, de modo que fosse um volume suficiente para dissolver esta quantidade de substância, mas não muito elevado a ponto de encharcar as amostras, correndo-se o risco de haver perda de material durante a produção das amostras simuladas. Um selo passou pelo processo de gotejamento somente com o solvente (metanol) com sua posterior evaporação para atuar como amostra de controle (branco).

Após homogeneização da solução com o auxílio de um agitador vortex, estas foram separadas em três partes iguais e impregnadas nos selos. No modo de gotejamento, utilizou-se uma micropipeta de 10 a 100 µL para auxiliar a administração do conteúdo nos microsselos. O volume da solução foi totalmente aplicado nos selos, gotejando-se microsselo por microsselo, a fim de se obter uma distribuição homogênea da substância. No modo imersão, todo o volume da solução foi colocado sobre uma placa de Petri e o

selo foi imerso nessa solução de uma só vez até completa absorção. Já no modo aspersão, o volume total da solução foi disposto em um pequeno borrifador. O selo foi posicionado a cerca de 5 cm do borrifador e toda a solução foi aspergida sobre o selo. É importante destacar que, tanto no modo de imersão como no de aspersão, é provável ter ocorrido perda de material durante a impregnação, mesmo com todos os esforços para evitá-la. Após a impregnação das três amostras, estas foram acomodadas sobre uma superfície para completa secagem em temperatura ambiente e, em seguida, foram encaminhadas para análise hiperespectral. Todo este processo de impregnação está ilustrado na figura 16.



**Figura 16.** Fotografias do processo de impregnação de 25E-NBOMe nos modos de imersão, aspersão e gotejamento.

Um último conjunto de 25 microsselos foi utilizado como branco, passando pelo gotejamento de solvente (metanol) e secagem em temperatura ambiente. De todos os quatro selos (gotejamento, imersão, aspersão e branco) foram obtidas oito imagens hiperespectrais, sendo duas para cada selo. Uma das imagens foi obtida com uma grade de 25×25 (625 espectros) tentando abranger todo o selo e a outra imagem foi obtida com uma grade de 20×20 (400 espectros) abrangendo um único microsselo.

### 3.3.3. Pré-processamento de sinais, restrições e estimativa inicial

A interface Hyper-Tools fornece um grupo restrito de pré-processamento de sinais para a análise das imagens. Nesta interface estão disponíveis os pré-processamentos mais comuns para dados espectrais: SNV, MSC, alisamento e derivadas. Esses vários pré-processamentos foram testados e foi escolhido aquele que apresentou melhores resultados nas análises das imagens.

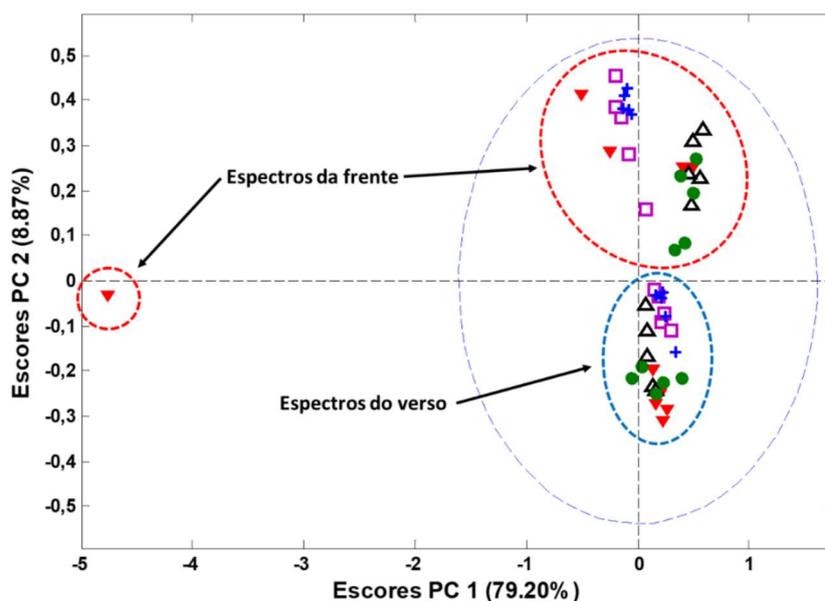
A única restrição aplicada aos modelos foi a restrição de não-negatividade nos perfis de concentração, sendo permitido valores negativos somente nos espectros puros recuperados, devido à característica do pré-processamento espectral escolhido. Além disso, o algoritmo *Purity* foi utilizado para a obtenção das estimativas iniciais dos espectros puros de cada componente.

## **4. RESULTADOS E DISCUSSÃO**

### **4.1. Identificação de drogas sintéticas em selos por espectroscopia no infravermelho próximo (NIR)**

#### **4.1.1. Testes iniciais**

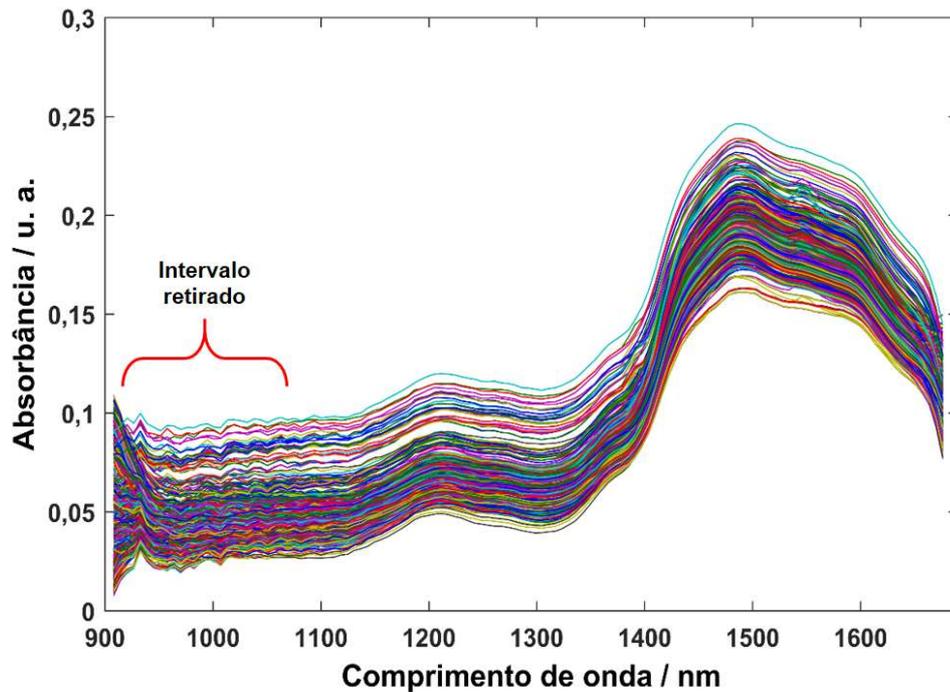
Utilizando os parâmetros definidos nos testes iniciais de aquisição de espectros dos selos (seção 3.1.2), observou-se uma falta de eficiência na discriminação das classes de NPS. Uma análise preliminar por PLS-DA revelou que as cinco replicadas da parte do verso se agrupavam com valores de classe distantes das cinco replicatas da parte da frente, fato que indica que existe uma diferença relevante entre esses espectros. Uma explicação para essa diferença é que a tinta da parte da frente dos selos fornece informação espectral adicional que diferenciava replicatas de uma mesma amostra que foram obtidas em faces diferentes do selo. Na figura 17 é apresentada a dispersão obtida pelas duas primeiras componentes principais de um modelo PCA construído com apenas 5 amostras contendo 10 espectros medidos para cada um, sendo 5 replicatas obtidas na frente e 5 no verso dos selos. É possível observar que, de cada amostra, 5 replicatas se agrupam com valores de escores diferentes das outras 5, o que é causado pela diferença nos espectros entre a parte da frente e do verso dos selos. Isto indica que, pelos escores do modelo PCA, as diferenças espectrais entre a frente e o verso dos selos são maiores do que as diferenças entre as amostras/apreensões. Ao se desenvolver o modelo por PLS-DA, essa diferença tornou-se mais evidente de modo a influenciar negativamente a discriminação das classes. Devido a isso, optou-se por obter somente os espectros do verso do selo, aumentando o número de replicatas para 10, com o intuito de melhorar a representatividade das amostras.



**Figura 17.** Gráfico de escores das PC1 e 2 para 10 replicatas espectrais (5 na parte da frente e 5 no verso dos selos) de 5 amostras de selos apreendidas. Amostra 1 ( $\square$ ), amostra 2 ( $\bullet$ ), amostra 3 ( $\blacktriangledown$ ), amostra 4 ( $\blacktriangle$ ) e amostra 5 ( $+$ ).

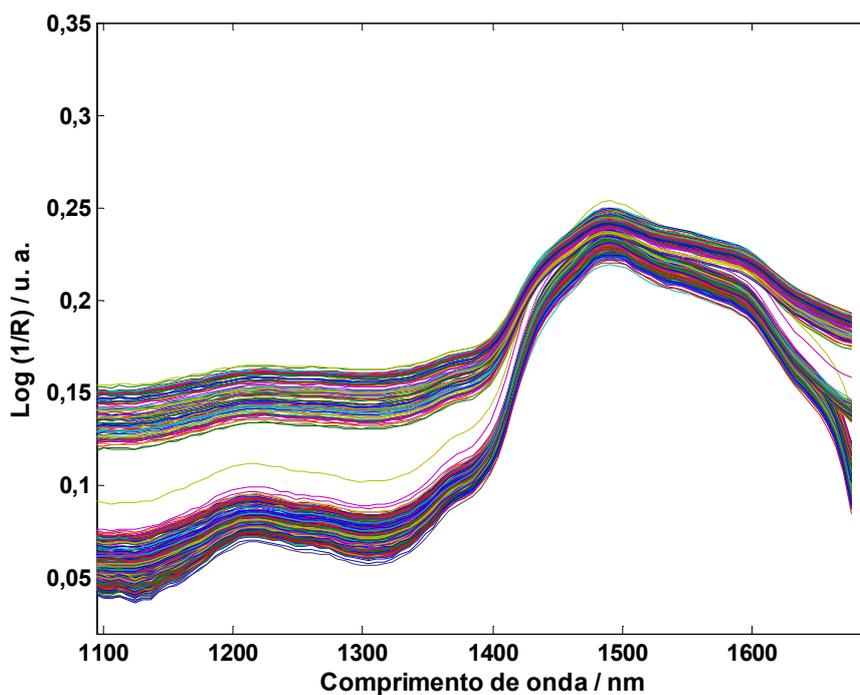
#### 4.1.2. Pré-processamento

Após a aquisição dos espectros, observou-se uma grande contribuição de ruído na parte inicial da região espectral do MicroNIR (908 a 1100 nm) para as amostras. A fim de se eliminar a contribuição de ruído nos dados, o qual pode prejudicar a discriminação entre as amostras pelos modelos, os primeiros 30 comprimentos de onda foram omitidos e o número de variáveis/comprimentos de onda passou de 125 para 95. São apresentados, na figura 18, os espectros para os selos sem drogas com toda a região espectral do MicroNIR, com uma ênfase na região que foi omitida devido ao ruído espectral.

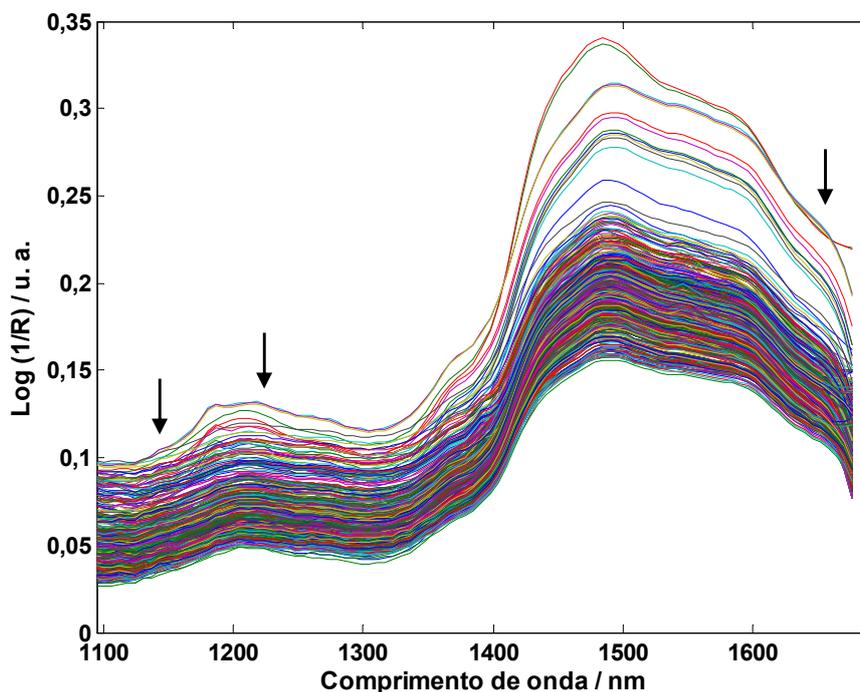


**Figura 18.** Espectros brutos na região do infravermelho próximo para amostras que não contêm drogas (branco) com 125 variáveis (comprimentos de onda).

Na figura 19 são apresentados os espectros brutos para os selos que não contêm drogas (branco) e na figura 20 os espectros para as amostras apreendidas, com conteúdo variado de substâncias ilícitas. Das duas figuras foram retiradas a parte espectral correspondente ao ruído mais intenso (30 primeiros intervalos espectrais), de modo que ambas apresentam 95 intervalos espectrais. É possível observar que a resposta espectral para os selos (contendo ou não as drogas) apresenta bandas largas e com padrão semelhante, uma vez que a maior contribuição espectral se refere ao papel. As mudanças espectrais observadas se referem à algumas variações de intensidade da absorvância, à ligeira variação da posição de determinadas bandas e até mesmo ao aparecimento de novas bandas de baixa intensidade. Há ainda, na figura 18, grupos de espectros com intensidades diferentes na região entre 1200 e 1500 nm. Apesar de serem espectros de papel (selos que não contêm substâncias impregnadas), estes constituem tipos diferentes (selos adquiridos na internet e papéis de filtro), de modo que sutis variações no sinal espectral são observadas.



**Figura 19.** Espectros brutos na região do infravermelho próximo para amostras que não contêm drogas (branco).



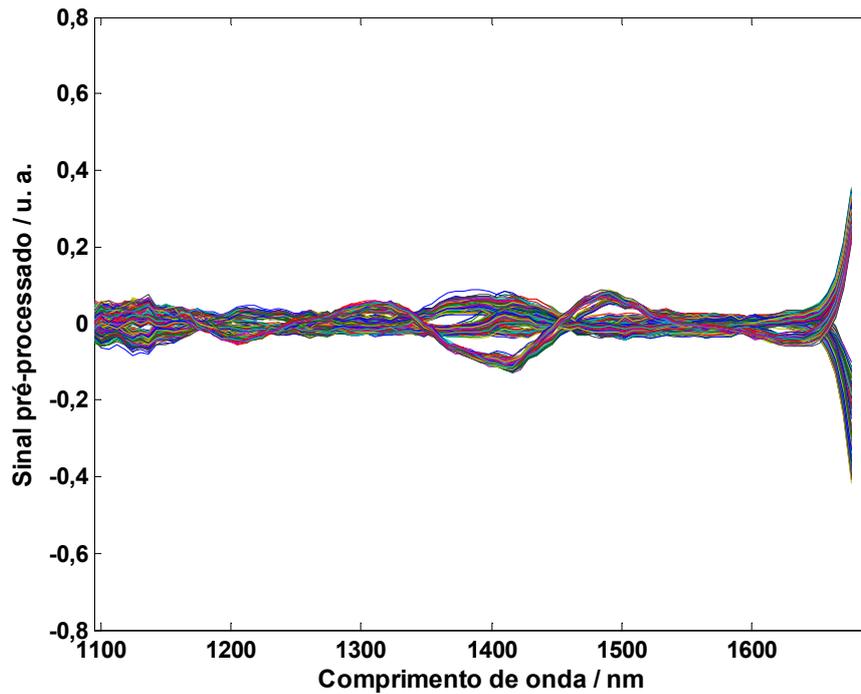
**Figura 20.** Espectros brutos na região do infravermelho próximo para as amostras que contêm drogas.

Em relação ao aparecimento de bandas, na figura 20, por exemplo, alguns poucos espectros apresentam uma pequena banda por volta de 1100 e 1200 nm (indicado por setas na figura) com intensidade um pouco maior que os outros espectros

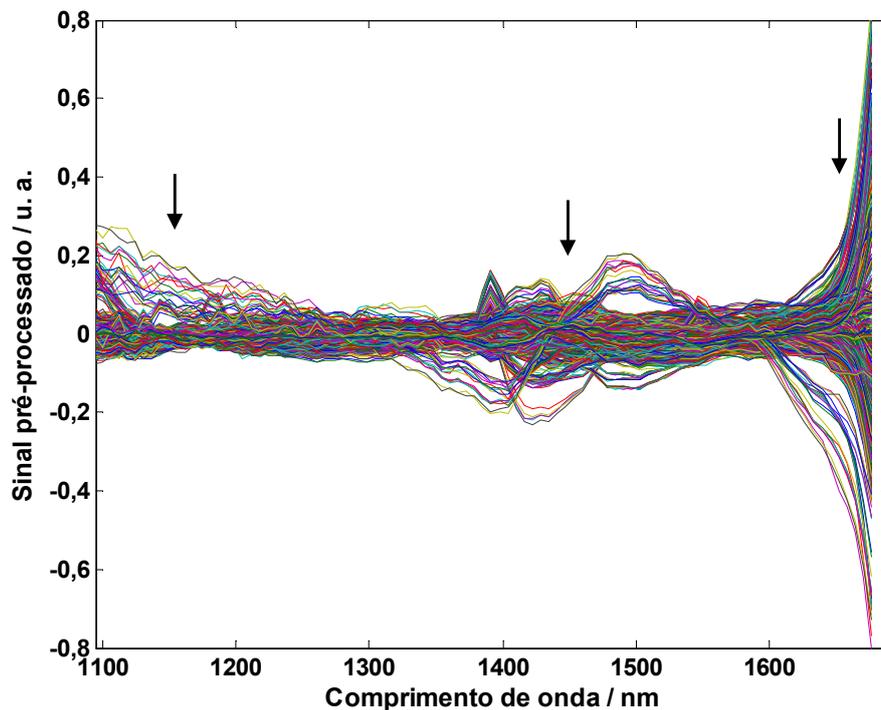
nessa mesma região. Esta banda se refere ao estiramento simétrico e assimétrico de C-H de metila. Porém, deve-se considerar que essa mudança é muito sutil e que não aparece em um número significativo de amostras. Também na mesma região, até cerca de 1300 nm (indicado por setas na figura), há o estiramento simétrico de C-H de metileno. As bandas de anéis aromáticos (referentes às drogas) se encontram na região entre 1680 a 1800 nm (indicado por setas na figura). Estas observações são resultado do sobretom de bandas mais comuns, referentes às ligações de carbono com outros átomos, como o hidrogênio. A ligação O-H apresenta sobretom de 760 a 1200 nm e de 1500 a 2200 nm. Contudo, essa banda não pode ser considerada específica para poder ser usada como marcadora, pois tanto a celulose presente nos selos, como as moléculas das drogas apresentam grupos O-H. Os grupos de aminas (referentes às drogas) apresentam sobretoms entre 1530 e 1650, 1030 e 1150 nm, regiões que podem ser consideradas mais específicas para a droga. Álcoois primários e secundários (celulose) apresentam sobretoms entre 1400 e 1500 nm.

Esse comportamento espectral, referente à presença de sobretoms, é característico da região entre 780 a 1900 nm do infravermelho próximo, que possui menor sensibilidade, quando comparado à região de bandas de combinação (2000 a 2500 nm) e ao infravermelho médio e gera pouca informação específica da estrutura molecular dos componentes de uma amostra. Como a diferença entre os espectros de diferentes amostras é pouco perceptível, a associação desta técnica analítica com métodos de calibração multivariada se faz imprescindível para a discriminação dessas amostras.

Inicialmente diferentes tipos de pré-processamentos foram testados para o conjunto de dados espectrais. Neste trabalho, optou-se pela utilização do SNV e também dados centrados na média, uma vez que esta combinação de pré-processamentos apresentou melhores resultados de discriminação, comparando-se os resultados finais dos modelos. A realização de primeira e segunda derivadas sobre o conjunto de dados forneceu resultados bons, porém não melhores que a realização da combinação de pré-processamentos supracitados. Os espectros pré-processados das amostras do branco e das apreendidas são apresentados nas figuras 21 e 22, respectivamente.



**Figura 21.** Espectros pré-processados na região do infravermelho próximo para amostras que não contêm drogas (branco).



**Figura 22.** Espectros pré-processados na região do infravermelho próximo para as amostras que contêm drogas.

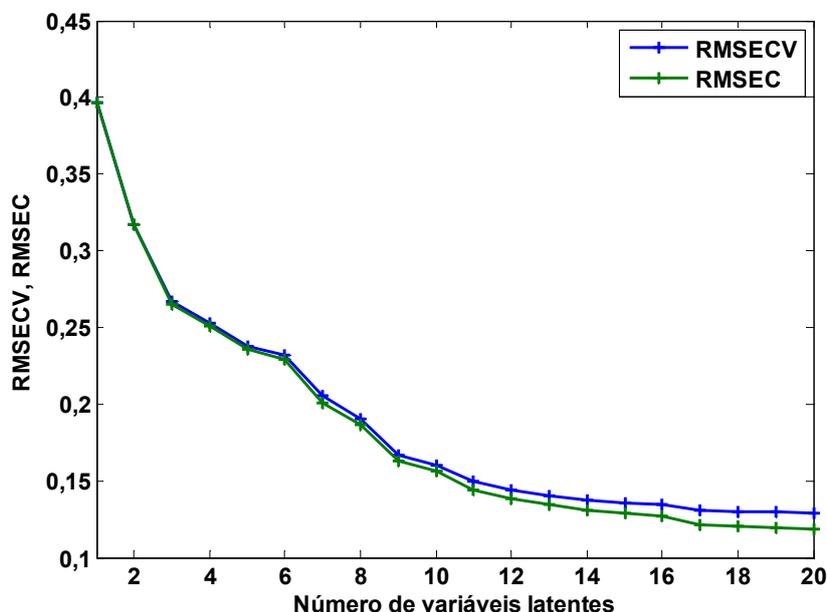
É possível se observar na figura 22 que, com a realização dos pré-processamentos, algumas diferenças podem ser notadas nos espectros das amostras

que contêm drogas. Na região entre 1100 e 1200 nm (indicado por setas na figura), pode-se observar uma variação de sinal para alguns espectros, bem como na região entre 1300 e 1500 nm (indicado por setas na figura) e também após os 1600 nm (indicado por setas na figura). Adicionalmente, o conjunto espectral referente as amostras apreendidas (figura 20) apresenta algumas diferenças em relação ao conjunto espectral para as amostras do branco (figura 19), o que sugere que uma discriminação dessas classes deve ser possível, principalmente devido às variações presentes em torno de 1100 e 1150 nm e também na região entre 1400 e 1500 nm.

### 4.1.3. Discriminação por PLS-DA

#### 4.1.3.1. Modelo A: Selos sem drogas versus selos com drogas

Os modelos de identificação por PLS-DA foram desenvolvidos após etapas de detecção de *outliers* e seleção do número de variáveis latentes que apresentavam menores valores de RMSECV. Na figura 23 é apresentado o gráfico de RMSECV pelas variáveis latentes. É possível observar que os valores de RMSECV diminuem com o aumento do número de variáveis latentes, e, próximo da variável latente 11, há uma estabilização do RMSECV, fato que justifica a escolha de 11 variáveis latentes.

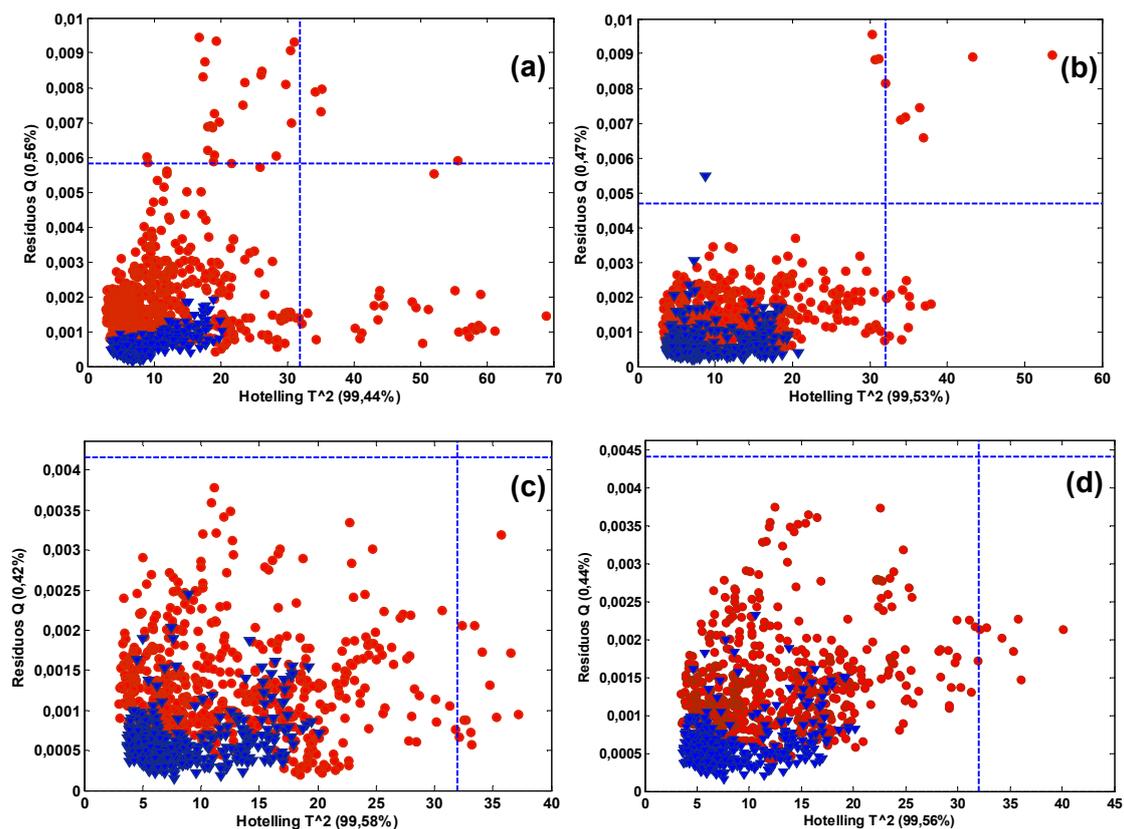


**Figura 23.** Gráfico de RMSEC e RMSECV por número de variáveis latentes para o modelo A.

A detecção de *outliers* foi realizada utilizando-se uma rotina de detecção desenvolvida por Soares e colaboradores.<sup>50</sup> Após a etapa de pré-processamento, no

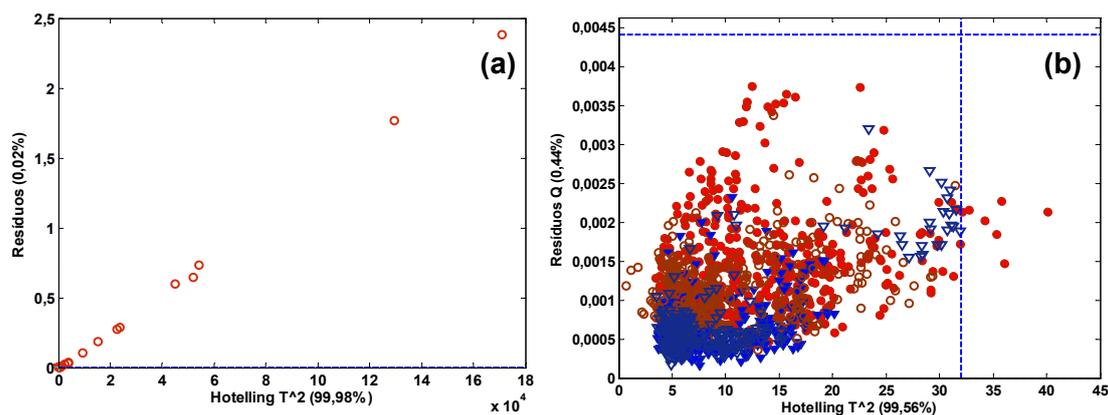
conjunto de dados da fase de treinamento do modelo A foram detectados inicialmente 58 *outliers* (Figura 24(a)). Na segunda rodada de detecção de *outliers*, 17 foram detectados (Figura 24(b)). Na última, 13 foram detectados (Figura 24(c)). Ao todo, na fase de treinamento do modelo, 88 replicatas foram eliminadas do conjunto de treinamento. Dessas 88, 87 são relativas a drogas e apenas uma replicata relativa ao branco. Esse número equivale a cerca de 9 amostras distintas, considerando que, para definir uma amostra/apreensão como *outliers*, 6 ou mais replicatas devem ser excluídas). Neste caso, das 88 replicatas, 60 foram relacionadas a 6 amostras apreendidas, pois todas as 10 replicatas dessas amostras foram excluídas (1 amostra contendo 25C-NBOMe, 2 contendo DOC, 1 contendo 25B-NBOMe, 1 contendo 25I-NBOH e 1 contendo acetaminofeno). As outras 27 replicatas continham alguns espectros de sete outras amostras (3 de 25B-NBOMe, 1 de 25I-NBOMe, 2 de 25I-NBOH e 1 de DOC).

Na figura 24(d) é apresentado o gráfico de influência Hotelling ( $T^2$ ) e resíduos (Q) para todas as replicatas do conjunto de treinamento que não foram consideradas *outliers* pela rotina de detecção. É possível observar que a área de união do gráfico (todos os quadrantes, à exceção do inferior esquerdo) apresenta uma pequena quantidade de replicatas, isto é, quase todas as replicadas com altos valores de resíduos, ou altos valores de influência ou, até mesmo, altos valores de ambos, foram eliminadas. É possível observar na figura 24(d) que, ao final do procedimento para exclusão de *outliers*, somente sete replicatas da fase de treinamento ainda apresentavam valores de Hotelling  $T^2$  levemente superiores ao seu limite de confiança de 99%. A presença dessas replicatas é condizente com o fato de se ter um conjunto de amostras altamente heterogêneo. Caso essas replicatas forem excluídas e o modelo recalculado, observa-se ainda algumas replicatas fora dos limites com o mesmo comportamento. Portanto, segundo a norma ASTM E1655-12,<sup>63</sup> os dados parecem apresentar o efeito “bola de neve” (ao se eliminar *outliers* novas amostras continuam a ser identificadas nos modelos reconstruídos). Por recomendação dessa norma, após a terceira exclusão o modelo foi considerado otimizado e os eventuais espectros fora dos limites nesse modelo não foram eliminados. Esse procedimento de exclusão de amostras anômalas propicia maior eficiência do modelo na classificação, uma vez que medidas com valores muito discrepantes em relação ao conjunto de dados foram retiradas e não influenciam no desenvolvimento do modelo.



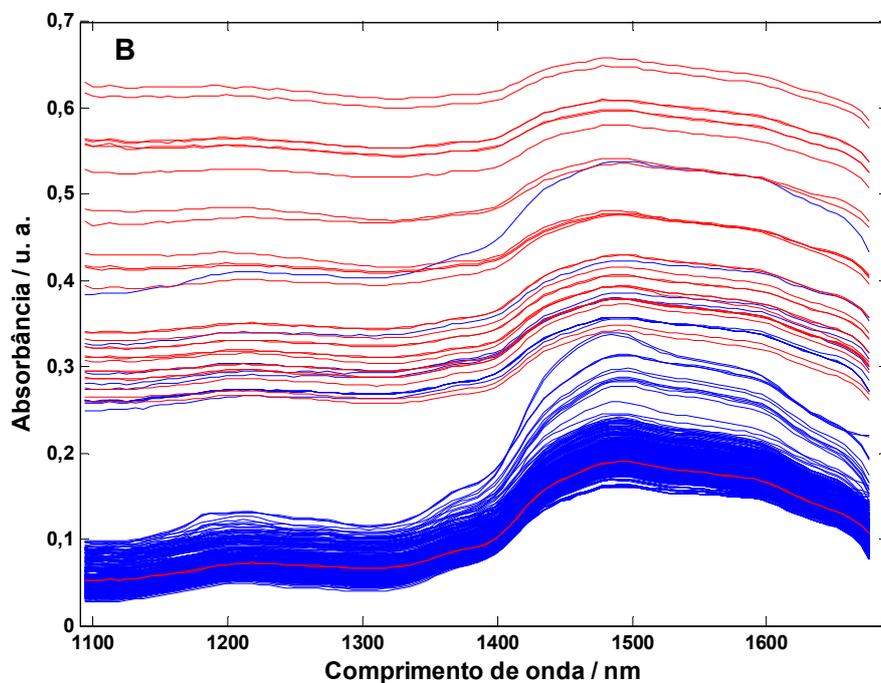
**Figura 24.** Gráfico de influência ( $T^2$ ) versus resíduos para a fase de treinamento no modelo A (● – selos que contêm drogas; ▼ – selos que não contêm drogas) antes da detecção de *outliers* (a), após a primeira etapa de detecção (b), após a segunda (c) e após três etapas de detecção de *outliers* (d).

Na fase de validação do modelo, foram excluídas pela rotina de detecção de *outliers* 128 replicatas, por meio de uma só etapa de exclusão. Neste conjunto de replicatas excluídas se encontravam quatro amostras apreendidas, as quais tinham como conteúdo a substância 25B-NBOH (38 replicatas). As replicatas dessas amostras, diferentemente das outras apreensões, apresentaram valores muito elevados de influência ( $T^2$ ) e resíduos ( $Q$ ), que podem ser observados na figura 25(a). Os valores de  $T^2$  e  $Q$  para estas replicatas foram tão discrepantes que as replicatas das outras amostras não aparecem claramente no gráfico devido à escala. Os outros espectros excluídos se referem ao branco (26 espectros), cinco outras amostras contendo NPS (50), e alguns poucos espectros de quatro outras amostras contendo 25B-NBOMe e 25I-NBOH (14). Após a exclusão desses *outliers*, é possível observar as replicatas das amostras com menores valores de influências e resíduos que se concentraram no primeiro quadrante do gráfico, juntamente com as replicatas da fase de calibração que não foram excluídas. O gráfico da figura 25(b) apresenta resultados para influência e resíduos das replicatas utilizadas nas fases de treinamento e validação.



**Figura 25.** Gráfico de influência ( $T^2$ ) versus resíduos (Q) para o modelo A (● – selos que contêm drogas; ▼ – selos que não contêm drogas) antes (a) e após a etapa de detecção de *outliers* na fase de validação (b).

Uma explicação para o fato da rotina de detecção de *outliers* ter excluído as replicatas das quatro amostras constituídas por 25B-NBOH é a grande diferença do verso dos selos dessas amostras em relação a todas as outras que fazem parte do conjunto de dados. O verso dos selos dessas quatro amostras não é isento de imagens e coloração, como apresentado na figura 7(b), para os selos do branco. Os selos dessas quatro amostras também apresentavam imagens escuras no verso, que contribuíram para a grande diferença dessas amostras em relação ao conjunto de dados da fase de treinamento e validação. Esta diferença pôde também ser observada pelo perfil espectral dessas amostras cujo verso apresentaram imagens escuras. Na figura 26A é apresentada uma fotografia do verso de uma das quatro amostras anômalas de selos contendo 25B-NBOH, comparado com o verso de um selo de NBOMe comum, com o verso isento de imagens e coloração. Os espectros dessas amostras são apresentados na figura 26B, na qual se observa uma grande variação de linha de base e intensidades.



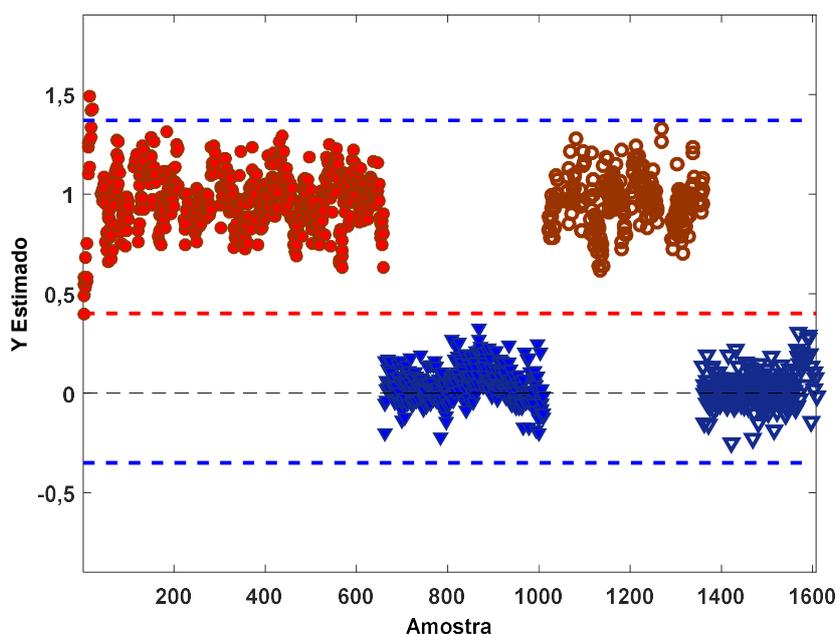
**Figura 26.** (A) Verso de uma das apreensões de selos contendo 25B-NBOH, no qual se observou imagens no verso em comparação com o verso de um selo contendo NBOMe. (B) Espectros dos selos contendo droga que apresentam o verso sem imagens (linhas azuis) e que apresentam imagens (linhas vermelhas).

Este resultado indica uma limitação do modelo gerado devido à heterogeneidade em relação ao aspecto físico das amostras. Se estas são muito diferentes do conjunto utilizado na fase de treinamento do modelo, são consideradas amostras anômalas e são excluídas do conjunto de dados. Portanto, concluiu-se que amostras de apreensões que apresentarem figuras no verso não podem ser analisadas pelo método NIR. Mesmo

assim estas amostras ainda foram aplicadas nos modelos a fim de se avaliar se eles continuam as identificando como *outliers*.

A rotina de detecção de *outliers* da fase de treinamento gerou os valores -0,35 e 1,37 como limites inferior e superior para os valores de classe estimados, respectivamente. Esses limites delimitam a área de distribuição dos valores de classe estimados para cada uma das replicatas, de modo que as amostras que ultrapassarem esses limites também são consideradas *outliers*.

Na figura 27 é apresentada a distribuição das replicatas das amostras da classe que não contém drogas (triângulos azuis) e da classe que as contém (círculos vermelhos) por meio dos valores estimados de classe para as fases de treinamento e validação. A linha vermelha no gráfico da figura 27 equivale a aproximadamente 0,40 e indica o limite de discriminação entre as duas classes. É possível observar pela imagem que nenhuma replicata foi classificada como falso positivo ou negativo, gerando uma discriminação com 100% de eficiência para o modelo A em ambas as fases do modelo.



**Figura 27.** Valores estimados de classe para a discriminação por PLS-DA de amostras que contêm (●) e que não contêm drogas (▼). Símbolos preenchidos indicam a fase de treinamento e vazios indicam a fase de validação.

Os valores estimados de classe para as amostras do modelo A apresentam distribuições diferentes, o que está relacionado à natureza dessas amostras. A classe que contém drogas (círculos vermelhos) apresenta uma distribuição mais dispersa em torno do valor 1, que é o valor do vetor  $y$  referente à presença dessas substâncias. Isto ocorre devido à variedade de tipos de drogas (NBOMe, NBOH, LSD, MDMA, entre

outros) nesta classe. De maneira oposta, a classe que não contém drogas (triângulos azuis) apresenta uma dispersão mais concentrada em torno do valor 0 do vetor  $y$ , referente às amostras que não contém estas substâncias.

Na tabela 3 são apresentadas as figuras de mérito para o modelo A. Para este modelo, seleccionaram-se 11 variáveis latentes, conforme discutido na figura 23. Uma vez que a dispersão das amostras de treinamento e validação, na figura 27, são semelhantes, não há evidências de sobreajuste.

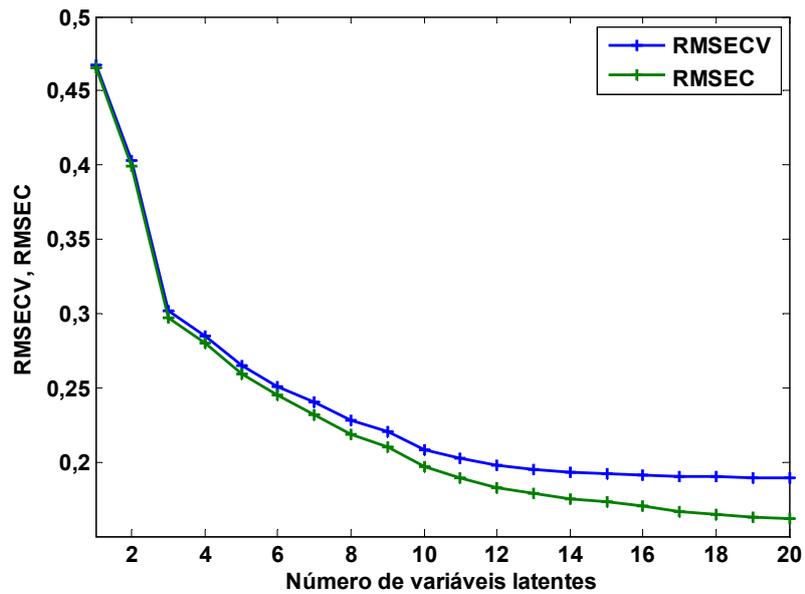
**Tabela 3.** Figuras de mérito para o modelo A utilizando PLS-DA como ferramenta de discriminação.

Modelo A				
Fase	Classe	TFP (%)	TFN (%)	TEF (%)
Treinamento	Branco	0	0	100
	Drogas	0	0	100
Validação	Branco	0	0	100
	Drogas	0	0	100

TFP: Taxa de falsos positivos; TFN: Taxa de falsos negativos; TEF: Taxa de eficiência.

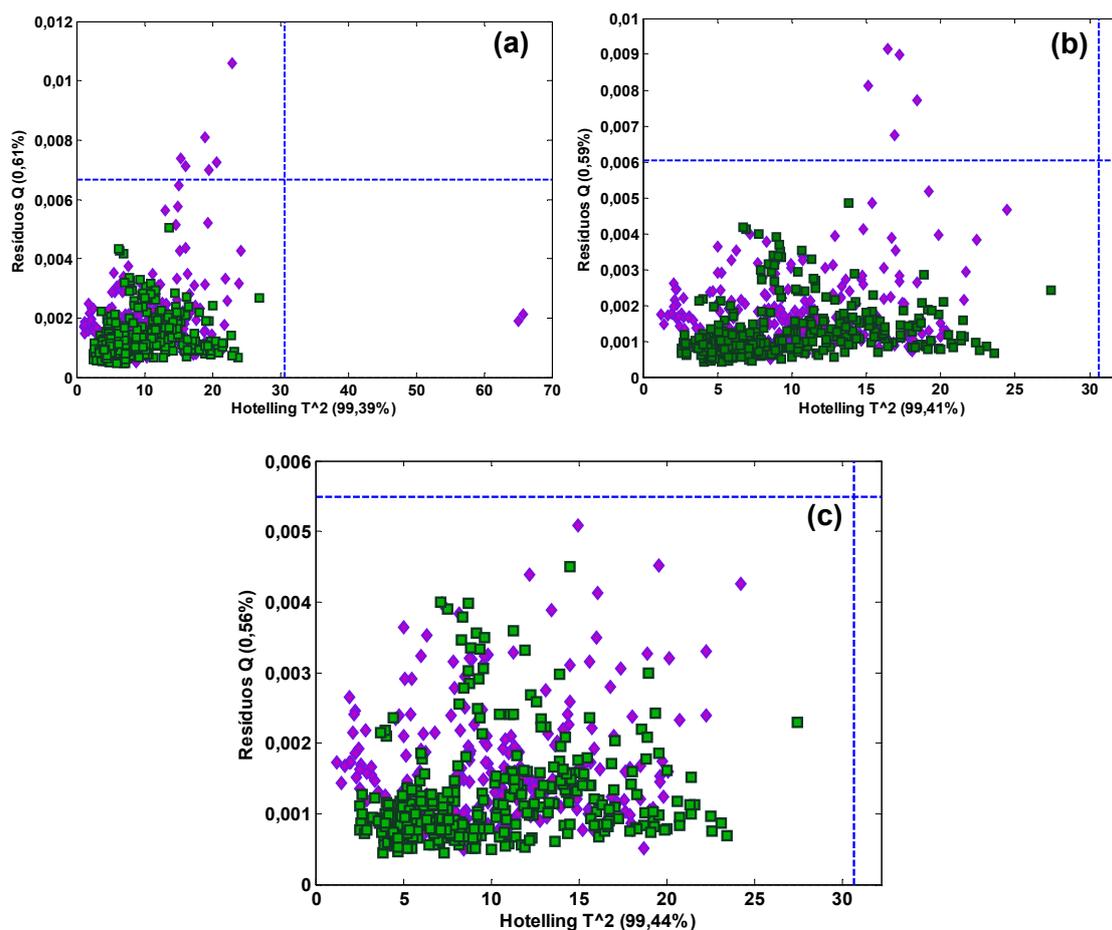
#### 4.1.3.2. Modelo B: Selos com NBOMe versus selos com NBOH

O modelo B foi elaborado de forma semelhante ao modelo A, porém com diferentes conjuntos de treinamento e validação. Da mesma forma que o modelo A, a escolha das variáveis latentes foi baseada nos valores de RMSECV e a etapa de detecção de *outliers* com base em  $Q$ ,  $T^2$  e nos limites para os valores estimados de classe. Observando-se o gráfico de RMSECV em função do número de variáveis latentes, este primeiro valor não apresenta grandes variações a partir de 10 variáveis latentes, sendo este número escolhido para o modelo, conforme apresentado na figura 28.



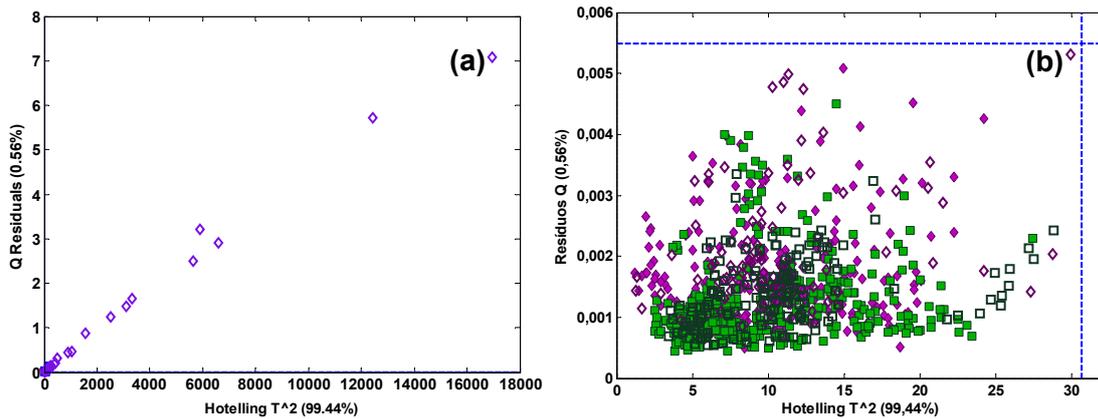
**Figura 28.** Gráfico de RMSEC e RMSECV por número de variáveis latentes para o modelo B.

Após o pré-processamento dos dados, foi obtido o gráfico de influência Hotelling ( $T^2$ ) e resíduos (Q) para as amostras da fase de treinamento (figura 29(a)). A rotina de detecção de *outliers* propiciou a exclusão de 13 replicatas, em apenas duas etapas de exclusão (figura 29(b) e 29(c)). Na fase de treinamento, após a etapa de detecção de *outliers*, obteve-se o gráfico de influência Hotelling ( $T^2$ ) e resíduos (Q) apresentado na figura 29(c). Neste caso, pode-se observar que na área de união do gráfico não há replicata alguma, de modo que todas as replicadas com altos valores de resíduos, de influência e de ambos, foram eliminadas.



**Figura 29.** Gráfico de influência ( $T^2$ ) versus resíduos para a fase de treinamento no modelo B antes da etapa de detecção de *outliers* (a), na primeira (b) e na segunda e última etapa (c). Amostras de NBOMe (◆) e NBOH (■).

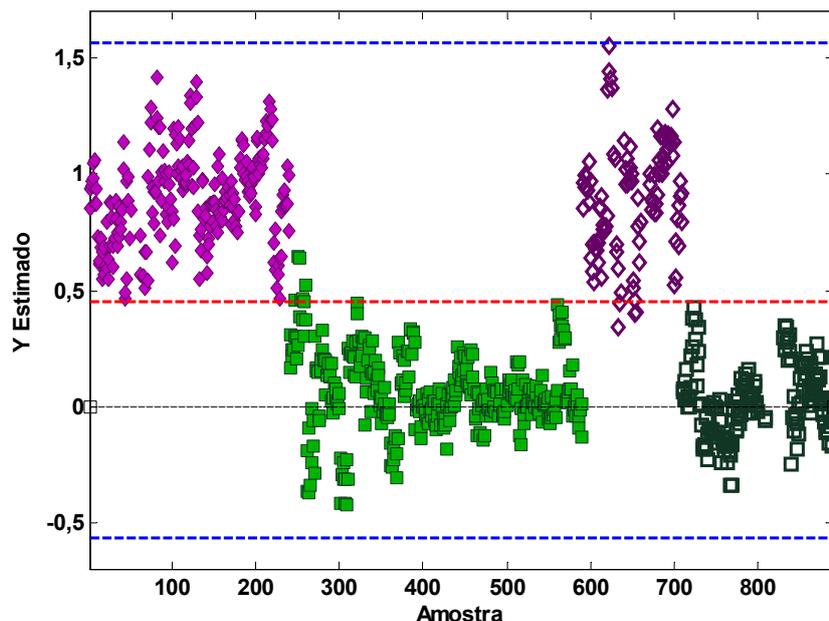
Do mesmo modo que no modelo A, na fase de validação realizou-se uma etapa de detecção de *outliers*, na qual 56 espectros foram excluídos do conjunto em questão. De forma semelhante ao modelo A, 40 das replicatas identificadas correspondem a amostras contendo 25B-NBOH que apresentavam o verso dos selos com imagens e coloração, devido aos elevados valores de influência e resíduos. Estas replicatas podem ser observadas na figura 30(a), de modo que as outras não podem ser visualizadas claramente devido à escala do gráfico. A figura 30(b) apresenta o gráfico de influência Hotelling ( $T^2$ ) e resíduos (Q) para as replicatas dos conjuntos de treinamento e validação após a exclusão dos *outliers*. Esse resultado confirma a limitação do método para a amostras com figuras no verso.



**Figura 30.** Gráfico de influência ( $T^2$ ) versus resíduos para o modelo B antes (a) e após a detecção de *outliers* na fase de validação (b). Amostras de NBOMe (◆) e NBOH (■).

Os limites inferior e superior gerados pela rotina de detecção de *outliers* para o modelo B foram de -0,54 e 1,55, respectivamente e o limite de discriminação de 0,45. Na figura 31, é apresentada a dispersão das amostras de acordo com seus valores estimados no vetor de classes ( $\mathbf{y}$ ). A linha tracejada em vermelho indica o limite de discriminação entre as classes NBOMe e NBOH. Replicatas com valores estimados que ultrapassassem o limite de discriminação foram consideradas como falsos negativos/positivos dependendo da classe.

É importante salientar que cada amostra apreendida apresenta 10 espectros como replicatas, e, como explicado anteriormente, para concluir que uma amostra é um falso positivo/negativo, é necessário que mais da metade das replicatas sejam falsos positivos/negativos, isto é, 6 ou mais replicatas. Considerando este fator, é possível observar na figura 31 que somente uma amostra foi considerada falso positivo para a classe NBOMe na fase de treinamento (falso negativo para a classe NBOH), resultando em taxas de eficiência de 97,1 e 100% para as fases de treinamento e validação, respectivamente, conforme apresentado na tabela 4.



**Figura 31.** Valores estimados de classe para a discriminação por PLS-DA de amostras de NBOMe (◆) e NBOH (■). Símbolos preenchidos indicam replicatas na fase de treinamento e símbolos vazios indicam replicatas na fase de validação).

**Tabela 4.** Figuras de mérito para o modelo B utilizando PLS-DA como ferramenta de discriminação.

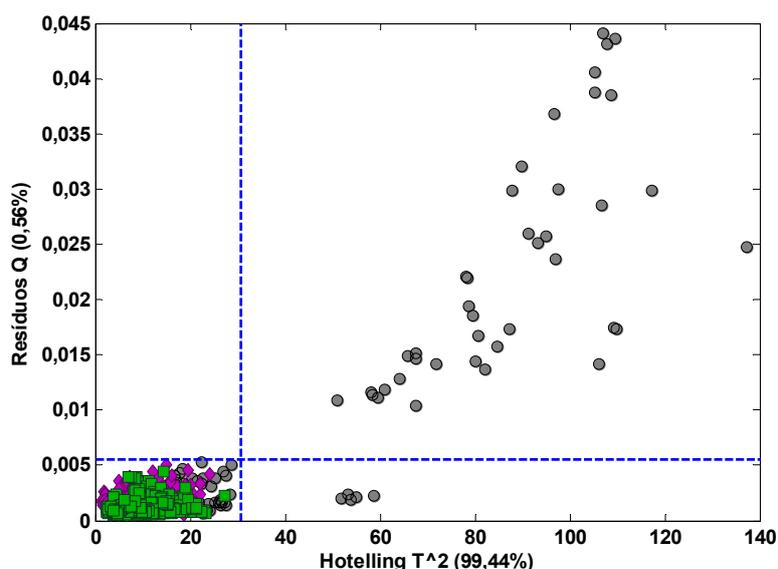
Modelo B				
Fase	Classe	TFP (%)	TFN (%)	TEF (%)
Treinamento	NBOMe	2,9	0	97,1
	NBOH	0	2,9	97,1
Validação	NBOMe	0	0	100
	NBOH	0	0	100

TFP: Taxa de falsos positivos; TFN: Taxa de falsos negativos; TEF: Taxa de eficiência.

Para avaliação da capacidade de predição do modelo B com classes não incluídas na fase de treinamento foi realizado com um conjunto de validação composto por 11 apreensões de selos com as substâncias LSD, DOC, MDMA e acetaminofeno. Ao analisar os espectros dessas amostras no modelo B observou-se que apenas a metade dos espectros deste conjunto apresentou valores de  $T^2$ , Q e valores estimados de classe acima dos seus respectivos limites. A outra metade dos espectros apresentou valores compatíveis com os observados na fase de treinamento, no primeiro quadrante do gráfico de influência e resíduos, apresentado na figura 32. Neste gráfico, é possível

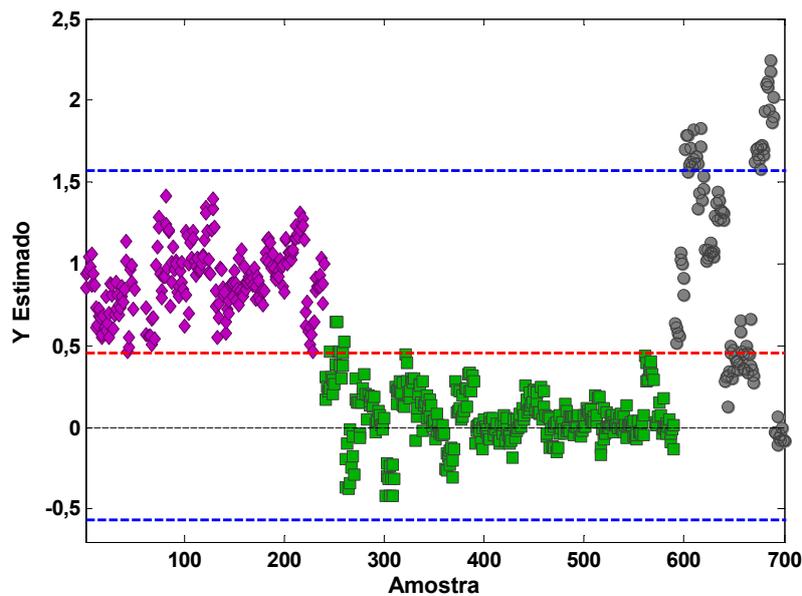
observar a distribuição de algumas das replicatas do grupo contendo LSD, DOC, MDMA e acetaminofeno (círculos cinza) na região de união dos dois parâmetros, e outra parte no primeiro quadrante. Este resultado indica que o modelo desenvolvido não foi muito eficiente em identificar essas amostras como contendo drogas diferentes da fase de treinamento, que continha apenas NBOMe e NBOH.

Uma explicação para essa deficiência pode estar na baixa concentração dessas drogas no selo e na sensibilidade do NIR. Outro fator relevante se refere à grande sobreposição entre as bandas do papel e das NPS. Como o limite de detecção no NIR é de cerca de 0,1% m/m,<sup>47</sup> é possível que os espectros que apresentaram baixos valores de  $T^2$  e  $Q$  tenham sido obtidos em locais que continham drogas em concentrações menores do que o limite da técnica, fato corroborado pela dispersão das replicatas de uma mesma amostra. Este fato será confirmado posteriormente através do estudo da distribuição dos NPS nos selos, na seção 4.3. De acordo com Oliveri e colaboradores, a ocorrência de erros de discriminação para classes não modeladas é esperada em modelos PLS-DA.<sup>58</sup> Contudo, esperava-se que essa ocorrência pudesse ser minimizada pela detecção de *outliers*.



**Figura 32.** Gráfico de influência ( $T^2$ ) versus resíduos para o modelo B usando espectros de outras drogas como grupo de validação (●). Amostras de NBOMe (◆) e NBOH (■).

Na figura 33 é apresentado o gráfico de valores estimados para o modelo B, quando o grupo de validação é composto por espectros das 11 amostras de outras drogas (110 espectros). É possível observar que a dispersão dessas amostras é grande e algumas replicatas ultrapassam o limite superior, indicando que esses limites também contribuem para a identificação desses *outliers*.



**Figura 33.** Valores estimados de classe para a discriminação por PLS-DA de amostras NBOMe (◆) e NBOH (■) como conjunto de validação (●) composto por outras drogas.

#### 4.1.4. Classificação por SIMCA

Os modelos A e B desenvolvidos por SIMCA foram baseados no mesmo pré-processamento e validação cruzada utilizados para os modelos desenvolvidos por PLS-DA, de modo que a avaliação de eficiência entre estas duas ferramentas de classificação possa ter parâmetros de comparação semelhantes. Já a exclusão de *outliers* foi realizada para as replicatas que apresentaram valores de resíduos Q e Hotelling  $T^2$  maiores que o limite de confiança de 99% de forma manual, por meio de modelos PCA também em três etapas de exclusão. Para tanto, antes do desenvolvimento dos modelos SIMCA, os conjuntos de treinamento e validação dos modelos A e B foram submetidos a modelos PCA, com os números de CPs já estabelecidos (aplicados nos modelos SIMCA) e as replicatas que estivessem além dos limites de Q e  $T^2$  foram excluídas manualmente.

##### 4.1.4.1. Modelo A: Selos sem drogas versus selos com drogas

A ferramenta de classificação SIMCA, como abordado na seção 2.4.4, gera, para cada classe, um modelo PCA. Deste modo, durante o desenvolvimento do modelo A por SIMCA, o número de componentes principais escolhido para cada uma das duas classes foi baseado nos menores percentuais de falsos positivos/negativos obtidos após a consolidação do modelo na fase de treinamento. A classe contendo amostras sem drogas (classe 1) apresentou menores valores de falsos positivos/negativos com 8

componentes principais, enquanto a classe contendo amostras com drogas (classe 2) apresentou menores percentuais para estes parâmetros com 4 componentes principais. Em todas as etapas de detecção de *outliers* na fase de treinamento foram excluídas 99 replicatas (87 de NPS e 12 de branco). Dessas 99 replicatas excluídas, 77 coincidem com as replicatas excluídas no modelo A desenvolvido por PLS-DA. Isto é, de modo geral os modelos por PLS-DA e SIMCA conseguem identificar como *outliers* praticamente as mesmas amostras.

Quando se utiliza SIMCA para a classificação de amostras, devido ao fato desta ferramenta ser baseada na probabilidade de uma amostra pertencer a certa classe, existem diferentes parâmetros para a classificação. No caso do modelo A, o parâmetro escolhido para a classificação das amostras em diferentes classes foi o parâmetro classe mais provável, que fornece somente uma opção de classe para cada amostra. Este parâmetro foi definido para o modelo A devido ao contexto de apreensão de selos. Quando estes materiais são apreendidos, só há duas possibilidades: ou eles contêm drogas ou não. Deste modo, o modelo A, desenvolvido por SIMCA, apresentará menos erros de classificação, uma vez que suas possibilidades para classificar amostras são bem definidas, não havendo, por exemplo, possibilidade de uma mesma amostra ser classificada em mais de uma classe, gerando um resultado inconclusivo.

Tomando isto em consideração, o modelo A apresentou uma taxa de eficiência de 91,7%, na fase de treinamento, uma vez que 30 replicatas do branco foram classificadas como NPS, gerando uma taxa de falso positivo de 8,3%. Já na fase de validação, 52 replicatas foram excluídas como *outliers* (34 de NPS e 18 do branco). Nesta fase a eficiência foi de 91,1%, uma vez que uma amostra de NPS foi classificada como branco, fornecendo taxa de falso negativo de 2,9%, e 14 replicatas do branco foram classificadas como NPS, gerando uma taxa de falso positivo de 6,0%. Considerando esse resultado, a ferramenta SIMCA apresentou uma eficiência semelhante ao PLS-DA para o modelo A, uma vez que houve uma diminuição na taxa de eficiência em ambas as fases menor que 10%. Na tabela 5 são apresentadas as figuras de mérito para o modelo A, bem como o número de componentes principais de cada classe, que geraram os menores valores de falso positivo/negativo.

**Tabela 5.** Figuras de mérito para o modelo A utilizando SIMCA como ferramenta de classificação.

Modelo A					
Fase	Classe	CP	TFP (%)	TFN (%)	TEF (%)
Treinamento	Sem NPS	8	8,3	0	91,7
	NPS	4	0	8,3	91,7
Validação	Sem NPS	-	2,9	6,0	91,1
	NPS	-	6,0	2,9	91,1

CP: Componentes principais; TFP: Taxa de falsos positivos; TFN: Taxa de falsos negativos; TEF: Taxa de eficiência.

#### 4.1.4.2. Modelo B: Selos com NBOMe versus selos com NBOH

Para o modelo B construído por SIMCA, um diferente parâmetro de classificação foi escolhido devido ao contexto forense. Quando amostras de selos são apreendidas pela polícia e contêm alguma substância controlada, não há somente duas possibilidades de composição (NBOMe e NBOH). Os selos podem apresentar variadas substâncias e o parâmetro “classe mais provável” não é o mais coerente. Para esse modelo, o parâmetro escolhido foi a probabilidade de pertencer a determinada classe. Quando a probabilidade de mais da metade das replicatas de uma amostra for maior que 50% para uma certa classe, esta amostra é classificada nesta classe. Se essa probabilidade for menor que 50% para mais que cinco das replicatas, a amostra não é classificada nesta referida classe, podendo ser classificada na outra ou em nenhuma classe.

O número de componentes principais do modelo SIMCA de cada classe foi escolhido como aquele que fornecia a menor taxa de falso positivo/negativo na fase de treinamento. A classe dos NBOMe (classe 1) foi desenvolvida com 11 componentes principais, enquanto que a classe dos NBOH (classe 2) com 7 componentes principais. Na fase de treinamento, 52 replicatas foram consideradas outliers (29 de NBOMe e 23 de NBOH). Nesta fase, a eficiência do modelo foi de 36,4% e 87,9% para a classe NBOMe e NBOH, respectivamente, considerando as probabilidades de seis ou mais replicatas de uma mesma amostra. Para as classes NBOMe e NBOH as taxas de falsos negativos foram de 63,6% e 12,1%, respectivamente, isto é, as amostras referentes a essas porcentagens não foram atribuídas em nenhuma classe.

Estes resultados indicam que os conjuntos de amostras apreendidas de NBOMe e NBOH apresentam diferenças significativas dentre eles, que é mais pronunciada para

as amostras de NBOMe. A heterogeneidade entre as amostras de uma mesma classe torna a definição da região de classe pelas CPs uma tarefa difícil, fato que explica a elevada taxa de falso negativo para a classe NBOMe. Portanto, o modelo SIMCA não foi considerado eficiente para classificar amostras de NBOMe, apesar de classificar bem as amostras de NBOH.

Na fase de validação, foram excluídas 22 replicatas (9 de NBOMe e 13 de NBOH) em três etapas de exclusão de *outliers* e não foram detectados falsos positivos e nem amostras inconclusivas, porém as taxas de falsos negativos foram maiores em comparação com a fase de treinamento, sendo de 63,6% para a classe NBOMe e 38,9% para a classe NBOH. Essas taxas são bastante elevadas, e resultam em taxas de eficiência baixas, fato que indicaria que o modelo desenvolvido não é capaz de classificar futuras amostras corretamente.

Na tabela 6 estão descritos, de acordo com a atribuição de classe, os números das amostras em cada uma das classes (NBOMe e NBOH) em ambas as fases do modelo. Números de 1 a 24 e 1 a 12 se referem às amostras de NBOMe nos conjuntos de treinamento e validação, respectivamente. Números de 25 a 59 e 13 a 31 se referem às amostras de NBOH nos conjuntos de treinamento e validação, respectivamente. Na tabela 7 são apresentadas as figuras de mérito do modelo B por SIMCA, bem como as componentes principais utilizadas para as duas classes na fase de treinamento.

**Tabela 6.** Atribuição de classes e número das amostras no modelo B por SIMCA nas fases de treinamento e validação.

	Classe atual (treinamento)		Classe atual (validação)	
	NBOMe	NBOH	NBOMe	NBOH
Preditas como NBOMe	8 amostras (3, 4, 12, 17, 18, 20, 21 e 22)	-	4 amostras (2, 6, 10 e 11)	-
Preditas como NBOH	-	29 amostras (27, 28, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57 e 59)	-	11 amostras (14, 17, 18, 19, 20, 21, 26, 27, 28, 29 e 30)
Preditas em nenhuma classe	14 amostras (1, 2, 5, 6, 7, 9, 10, 11, 13, 14, 15, 16, 19 e 23)	4 amostras (25, 29, 55 e 58)	7 amostras (1, 3, 4, 5, 7, 9 e 12)	7 amostras (13, 15, 16, 23, 24, 25 e 31)
<i>Outliers</i>	2 amostras (8 e 24)	2 amostras (26 e 32)	1 amostra (8)	1 amostra (22)

**Tabela 7.** Figuras de mérito para o modelo B utilizando SIMCA como ferramenta de classificação.

Modelo B						
Fase	Classe	CP	TFP (%)	TFN (%)	TIN(%)	TEF (%)
Treinamento	NBOMe	11	0	63,6	0	36,4
	NBOH	7	0	12,1	0	87,9
Validação	NBOMe	-	0	63,6	0	36,4
	NBOH	-	0	38,9	0	61,1

Pela tabela 6 é possível perceber que, tanto na fase de treinamento quanto na de validação, as amostras de NBOMe, em sua maioria, não foram corretamente classificadas dentro da classe atual, fornecendo uma elevada taxa de falso negativo (tabela 7). Já para as amostras de NBOH, a quantidade não classificada foi menor,

quando comparada aos NBOMe, principalmente na fase de treinamento. Estes resultados indicam que o conjunto de dados da classe NBOMe apresenta uma maior variabilidade, ou seja, é um conjunto mais heterogêneo que o conjunto de dados da classe NBOH, fator já observado na dispersão dos valores de classe estimados para estas amostras no modelo B desenvolvido por PLS-DA (figura 31).

Para a avaliação da capacidade de predição, todas as amostras que contêm outras drogas foram classificadas fora das classes NBOMe e NBOH. Este resultado confirma que o SIMCA é bem mais eficiente que o PLS-DA na identificação de amostras que não apresentam nenhuma similaridade com o grupo de treinamento.

#### **4.1.5. Comparação de resultados por PLS-DA e SIMCA**

Os resultados para os modelos A e B por PLS-DA indicam que esta ferramenta fornece uma discriminação especial para as amostras, não havendo possibilidade de classificação ambígua, devido ao limite de classificação gerado. Os resultados de eficiência utilizando o PLS-DA como ferramenta de classificação são maiores, comparados ao SIMCA, e levam inicialmente a crer que esta ferramenta é mais adequada para a aplicação forense pretendida. É importante lembrar que tanto os modelos por PLS-DA como os modelos SIMCA foram desenvolvidos com a exclusão das amostras anômalas, por meio de uma rotina de detecção de *outliers* ou exclusão manual.

Mesmo com os limites mais flexíveis do modelo SIMCA, cujos resultados indicaram uma elevada taxa de classificação ambígua e amostras não classificadas, este também se mostrou útil em relação a identificação de amostras de classes que não foram incluídas na fase de treinamento. No modelo B, o conjunto de outras drogas não foi classificado como pertencente a nenhuma das classes definidas na fase de treinamento, comprovando a maior eficiência do SIMCA nesse aspecto.

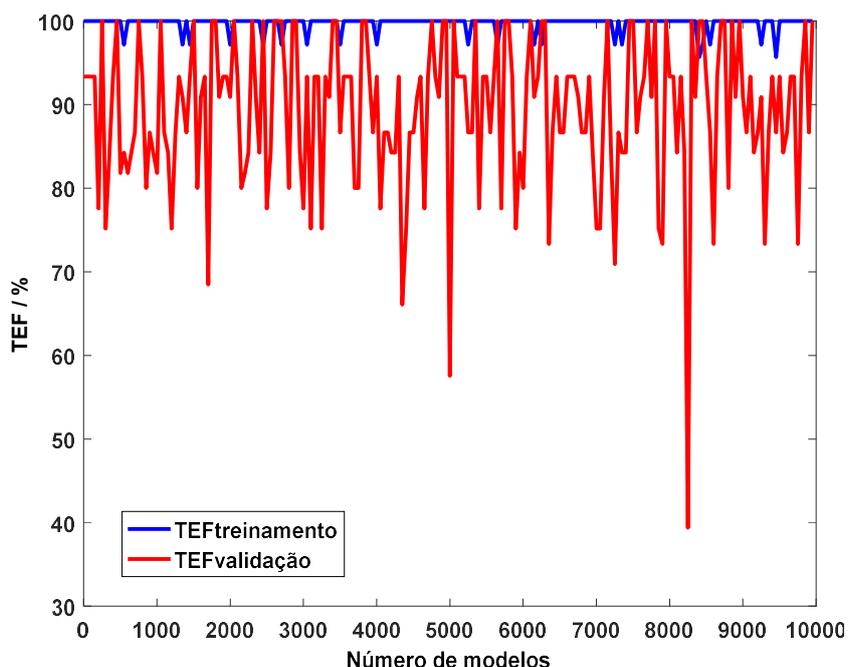
Em relação às taxas de eficiência dos modelos desenvolvidos por PLS-DA e SIMCA, nota-se que aqueles desenvolvidos por PLS-DA apresentam maiores taxas de eficiência quando comparados com os modelos desenvolvidos por SIMCA. Isto se deve, principalmente, pela forma com a qual a ferramenta SIMCA realiza a classificação de amostras. Uma vez que são definidas regiões de confiança para cada classe no espaço das CPs, existe a possibilidade de sobreposição entre elas. Adicionalmente, quando o conjunto de amostras é pouco homogêneo, no que diz respeito a suas características físico-químicas, as regiões delimitadas pelas CPs no SIMCA, com base nas amostras de treinamento, podem não ser representativas o suficiente para abranger todas as

situações das novas amostras que serão analisadas em rotina, aumentando os erros de classificação e diminuindo a eficiência do modelo.

#### 4.1.6. Avaliação da robustez do modelo B por PLS-DA com permutação de amostras

Ponderando as vantagens e desvantagens dos modelos SIMCA e PLS-DA, o modelo PLS-DA foi considerado mais apropriado, sendo escolhido para a continuidade do desenvolvimento do método e avaliação de sua robustez pelo teste de permutação descrito na seção 3.1.8.

Após o cálculo de 10000 modelos B por PLS-DA, valores de eficiência para as fases de treinamento e validação relativos a cada um desses modelos foram gerados. A figura 34 apresenta 200 destes valores (um a cada 50 modelos), para facilitar a visualização, no decorrer dos modelos desenvolvidos.



**Figura 34.** Valores de taxa de eficiência das as fases de treinamento e validação de 10000 modelos B por PLS-DA (Em azul, taxas de eficiência da fase de treinamento e em vermelho da fase de validação).

É possível observar nesta figura que as eficiências da fase de treinamento variam menos que na fase de validação. A média de eficiência obtida para a fase de treinamento foi de 99,7% e o desvio padrão para estas taxas de eficiência foi de 0,9%. Já para a fase de validação, observou-se uma variação maior. A média da taxa para

esta fase foi de 89,3% e desvio padrão de 9,1%. Estes valores estão indicados na tabela 8. É importante considerar que, apesar da taxa de eficiência média para o conjunto validação ter sido elevada, esta foi inferior à obtida na seção 4.1.3.2 e apresentou um desvio padrão consideravelmente elevado. Portanto, a forma com a qual são divididos os conjuntos de treinamento e validação influencia na eficiência do modelo, principalmente no que diz respeito à fase de validação, uma vez que esta conta com um menor número de amostras. Nesse sentido o modelo B ainda não pode ser considerado robusto quanto à divisão do conjunto validação. Esse resultado pode ser explicado pela grande heterogeneidade das amostras de selos e tende a ser minimizado com o aumento do número de amostras do conjunto de treinamento.

**Tabela 8.** Médias e desvio padrão das eficiências dos 10000 modelos B por PLS-DA.

Fase	Média da eficiência (%)	Desvio padrão da eficiência
Treinamento	99,7	0,9
Validação	89,3	9,1

Apesar de modelo não poder ser considerado robusto pelos resultados da tabela 8, deve-se observar que dentre os 10000 modelos construídos apenas 670 apresentaram taxa de eficiência menores que 75% na fase de validação, o que representa apenas 6,7% dos modelos. Portanto, apesar da influência na divisão dos conjuntos, a probabilidade de se obter um modelo com eficiência menor que 75% foi de aproximadamente 7% utilizando as amostras disponíveis, o que pode ser considerada aceitável e que pode ser minimizado à medida que se acrescentam novas amostras ao modelo.

É importante considerar que, a ocorrência de falsos positivos referentes à classe dos NBOH (ou seja, que são originalmente da classe NBOMe), não é um fator muito problemático nos laboratórios de perícia, uma vez que uma amostra que originalmente contém NBOMe e que é classificada como NBOH por meio desse método de triagem, segue para a análise confirmatória por LC-QTOF-MS, que é capaz de detectar o erro, gerando um laudo pericial definitivo correto. Por outro lado, se houvesse falsos positivos para a classe NBOMe (ou seja, que são originalmente da classe NBOH), essas amostras representariam um erro mais sério, uma vez que o protocolo atual da PCDF estabelece que essas amostras seriam posteriormente analisadas por GC-MS, que não é capaz de detectar o erro, devido à fragmentação da molécula. Neste caso, essa mesma amostra deveria ser novamente analisada, pelo método correto (LC-QTOF-MS), gerando mais gastos e elevando o tempo para a conclusão do laudo pericial. Com os 10000 modelos desenvolvidos no teste de permutação, obteve-se, nas fases de

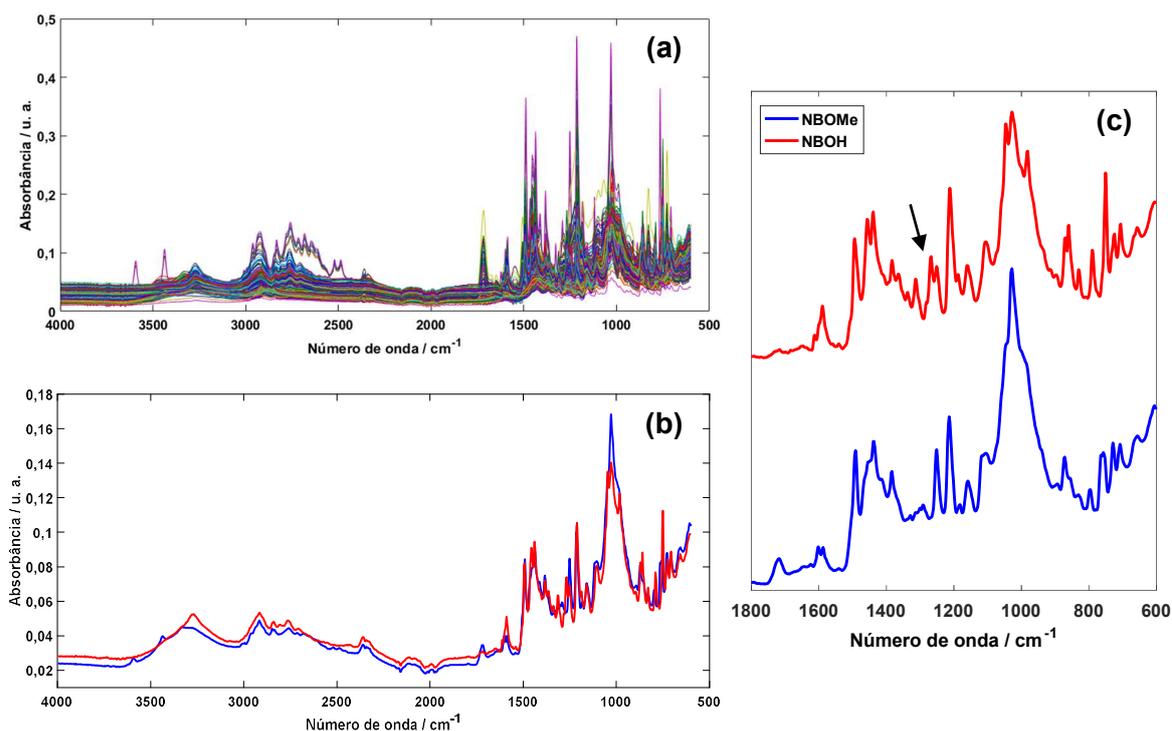
treinamento e validação, 0,3 e 6,9% de taxa de falsos positivos para a classe NBOMe, respectivamente, o que representa o percentual de amostras de NBOH identificadas como NBOMe. Apesar desse percentual de erros ser um problema, o fato de terem sido menores que 10% foi considerado aceitável.

Todos os resultados referentes ao método NIR para identificação de drogas sintéticas em selos apresentados nessa seção (seção 4.1) foram publicados em artigo científico na revista *Microchemical Journal*.<sup>79</sup>

## 4.2. Discriminação de drogas sintéticas em selos por espectroscopia no infravermelho médio (MIR)

### 4.2.1. Conjunto de dados e pré-processamento

Na figura 35(a) é apresentado o conjunto espectral de todas as amostras de NBOMe e NBOH apreendidas e utilizadas nos conjuntos de treinamento e validação dos modelos A e B por infravermelho médio. Pode se observar pela figura 35(a) que os sinais espectrais apresentavam uma variação de linha de base significativa, fato que justifica a utilização de segunda derivada como pré-processamento de sinal. As figuras 35(b) e 35(c) indicam os espectros médios para os NBOMe (linha azul) e para os NBOH (linha vermelha), evidenciando-se, em  $1250\text{ cm}^{-1}$ , uma banda pouco intensa para o NBOMe, referente às vibrações assimétricas do grupamento éter na molécula. Para o NBOH, esta banda se modifica, uma vez que há a substituição do grupamento éter pelo grupamento hidroxila, de modo que uma única banda se torna um pico duplo.



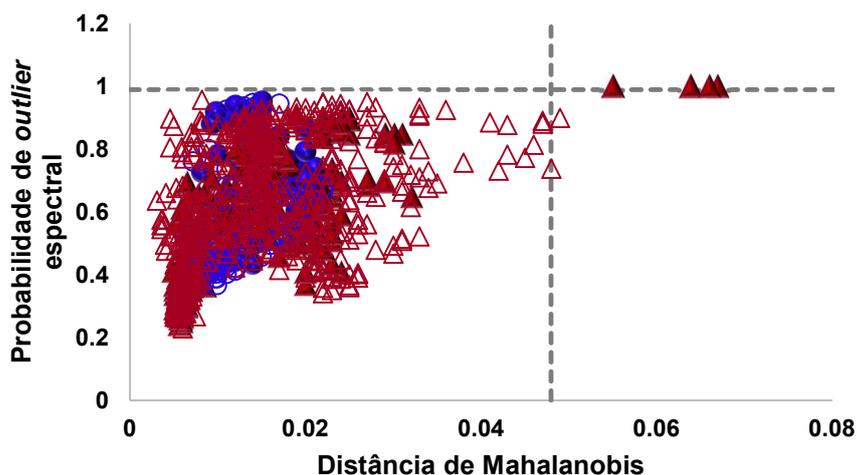
**Figura 35.** (a) Espectros ATR-FTIR das amostras apreendidas de NBOMe e NBOH usadas no conjunto de treinamento. (b) Espectros médios do NBOMe (—) e NBOH (—). (c) Visão aproximada dos espectros médios do NBOMe e NBOH de  $600\text{ cm}^{-1}$  a  $1800\text{ cm}^{-1}$ .

Além da segunda derivada, utilizou-se também os dados centrados na média como pré-processamento. A escolha desses pré-processamentos foi realizada pelo

*software* OPUS, após o desenvolvimento de vários modelos de classificação e sua organização por ordem crescente dos valores de RMSECV. Adicionalmente, o melhor modelo proposto pelo OPUS foi desenvolvido com seleção de variáveis. A região espectral utilizada foi de 2299,7 a 1277,4  $\text{cm}^{-1}$  para o modelo A e as regiões de 3997,4 a 2638,4  $\text{cm}^{-1}$  e 2299,7 a 1618,1  $\text{cm}^{-1}$  para o modelo B. É importante destacar que a banda característica de diferenciação de NBOMe e NBOH em 1250  $\text{cm}^{-1}$  não está incluída no intervalo espectral do modelo B. Isto ocorreu porque a otimização dos modelos é realizada de forma automática pelo *software*, de modo que o melhor modelo selecionado realizou a discriminação com outras regiões espectrais.

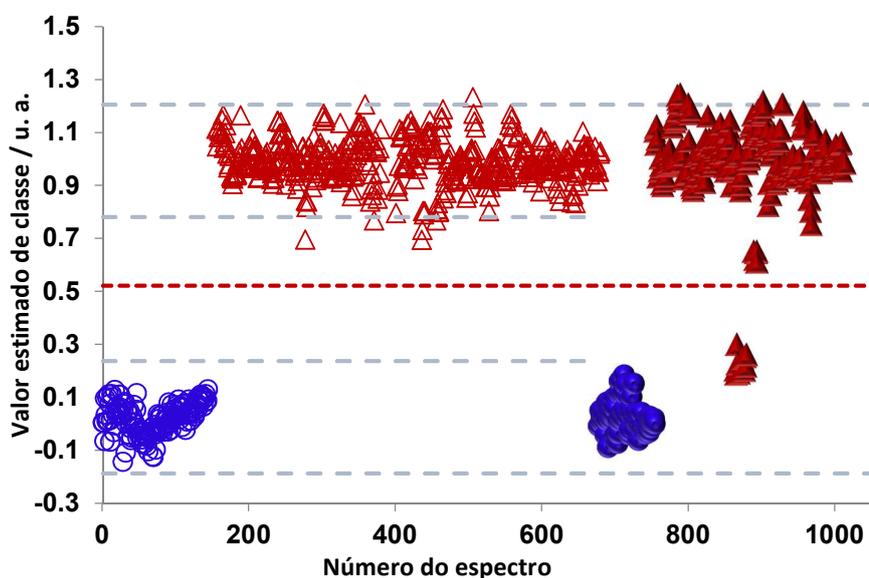
#### 4.2.2. Desenvolvimento dos modelos A e B

O modelo A foi desenvolvido com 13 variáveis latentes, valor que apresentou as menores taxas de erros de classificação pelos modelos desenvolvidos na extensão QUANT do *software* OPUS. Na fase de treinamento, 32 espectros foram considerados *outliers* por meio dos elevados valores de probabilidade de *outlier* espectral e distância de Mahalanobis considerando um limite de confiança de 99,9%. Os espectros excluídos na fase de treinamento são referentes a amostras contendo LSD e 25I-NBOH. Na fase de validação, 10 espectros foram excluídos após uma rodada de exclusão de *outliers*. O gráfico de probabilidade de um espectro ser *outlier* em função do resíduo espectral e pela distância de Mahalanobis dos espectros, após a exclusão de *outliers* na fase de treinamento e observação dos espectros das amostras de validação, é apresentado na figura 36. É observado que apenas poucos espectros de validação foram identificados como *outliers*, indicando que os critérios do *software* OPUS são menos restritivos que os utilizados no método NIR desenvolvido no PLS\_toolbox em MATLAB.



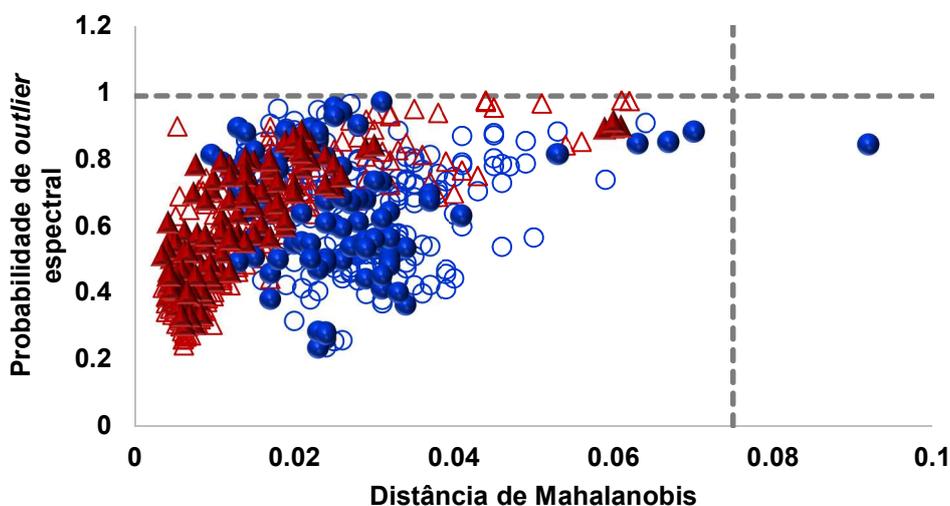
**Figura 36.** Probabilidade de *outlier* espectral pela distância de Mahalanobis das amostras usadas no Modelo A por ATR-FTIR.

Como se pode observar na figura 37, que apresenta os valores de classe estimados de cada amostra, na fase de treinamento há uma eficiência de discriminação de 100%, com ausência de falsos positivos/negativos. Já na fase de validação é possível se observar um grupo de espectros, equivalente a duas amostras, da classe dos NPS, que foram erroneamente classificadas como branco, fornecendo uma taxa de falso positivo para a classe branco de 3,6% (ou 3,6 % de falso negativo em relação à classe NPS).



**Figura 37.** Valores estimados de classe para discriminação entre amostras contendo NPS ( $\Delta$  – fase de treinamento;  $\blacktriangle$  – fase de validação) e amostras que não contêm NPS ( $\circ$  – fase de treinamento;  $\bullet$  – fase de validação). (---) Limite de discriminação: 0,52; (---) Limite superior (1,20) e inferior (-0,18).

O modelo B, por sua vez, foi desenvolvido com 9 variáveis latentes, valor que forneceu menores erros de discriminação. A exclusão de *outliers* na fase de treinamento eliminou 31 espectros, referentes a amostras contendo 25I-NBOH, em um limite de confiança de 99,9%, tal como realizado no modelo A. Durante a fase de validação, em uma única rodada de exclusão de *outliers*, somente 1 espectro foi eliminado do conjunto de dados. A figura 38 apresenta o gráfico de probabilidade de *outlier* espectral pela distância de Mahalanobis das amostras utilizadas no desenvolvimento das fases de treinamento e validação do modelo B, após as duas rodadas de exclusão de *outliers*. A tabela 9 resume a quantidade de *outliers* excluídos em ambas as fases de desenvolvimento dos modelos A e B por diferentes critérios de exclusão: distância de Mahalanobis, resíduos espectrais e resíduos de estimativa de classe. É importante salientar que, para cada um desses critérios, existem valores limites que definem se uma amostra será ou não excluída do conjunto de dados. Também deve-se observar que a soma de *outliers* identificados nos três critérios não corresponde ao total de *outliers* excluídos em uma rodada, uma vez que um mesmo *outlier* pode ser eliminado por mais de um critério, como pode ser observado para a rodada de exclusão na validação do modelo A.



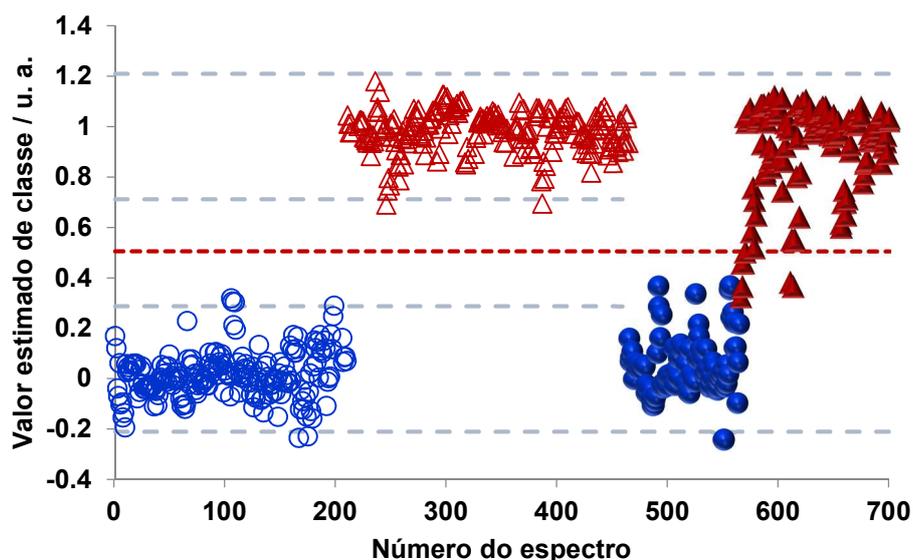
**Figura 38.** Probabilidade de *outlier* espectral pela distância de Mahalanobis das amostras usadas no Modelo B por ATR-FTIR.

**Tabela 9.** Número de *outliers* detectados por distância de Mahalanobis, resíduos espectrais e resíduos na estimativa de classe nos modelos A e B.

Modelo	Nº de <i>outliers</i> detectados em cada teste				
	Nº de espectros	Distância de Mahalanobis	Resíduos Espectrais	Resíduos na estimativa de classe	Total
MA <sub>Trein</sub>	680	0	0	24	24
MA <sub>out1</sub>	656	0	0	8	8
MA <sub>ot</sub>	648	---	---	---	---
MA <sub>Val</sub>	335	5	5	5	10
MB <sub>Trein</sub>	465	3	1	16	20
MB <sub>out1</sub>	445	0	0	11	11
MB <sub>ot</sub>	434	---	---	---	---
MB <sub>Val</sub>	240	1	0	0	1

MA: modelo A; MB: modelo B; Trein: treinamento; Val: validação; Out1: modelo com a primeira etapa de exclusão de *outliers*; Ot: modelo otimizado, com a segunda etapa de exclusão de *outliers*.

O gráfico de dispersão das amostras do modelo B é apresentado na figura 39. Nele é possível observar que, na fase de treinamento, não há falsos positivos ou negativos, de modo que a eficiência de discriminação nesta fase é de 100%. Nota-se ainda nesse gráfico que, mesmo depois das etapas de exclusão de *outliers*, alguns poucos espectros da classe NBOMe na fase de treinamento ainda apresentaram valores estimados de classe maiores que o limite superior dessa classe (o que também é observado para dois espectros da classe NBOH, além do limite inferior desta classe). Esses espectros não foram excluídos segundo a recomendação da norma ASTM E1655-12, que indica que após três etapas de exclusão o modelo já pode ser considerado otimizado.<sup>63</sup>



**Figura 39.** Valores estimados de classe para discriminação entre amostras contendo NBOH ( $\Delta$  – fase de treinamento;  $\blacktriangle$  – fase de validação) e amostras que não contêm NBOMe ( $\circ$  – fase de treinamento;  $\bullet$  – fase de validação). (---) Limite de discriminação: 0,50; (---) Limite superior (1,21) e inferior (-0,21).

A tabela 10 resume as taxas de falso positivo/negativos e eficiência para os modelos A e B em ambas as fases de desenvolvimento. Essas taxas são consideradas adequadas para a classificação em questão. Além disso, elas tendem a ser diminuídas ao se aumentar a representatividade do método, com adição de novas amostras, já analisadas pelos modelos, na fase de treinamento e validação, a fim de que estas ampliem representatividade do conjunto de dados.

**Tabela 10.** Figuras de mérito dos modelos A e B por ATR-FTIR.

	Modelo A			Modelo B				
	Classe	TFP (%)	TFN (%)	TEF (%)	Classe	TFP (%)	TFN (%)	TEF (%)
<b>Treinamento</b>	Sem NPS	0	0	100	NBOMe	0	0	100
	NPS	0	0	100	NBOH	0	0	100
<b>Validação</b>	Sem NPS	3,8	0	<b>96,2</b>	NBOMe	3,7	0	<b>96,3</b>
	NPS	0	3,8	<b>96,2</b>	NBOH	0	3,7	<b>96,3</b>

TFP: Taxa de falso positivo; TFN: Taxa de falso negativo; TEF: Taxa de eficiência.

### 4.2.3. Análise de amostras desconhecidas

A aplicação em rotina dos modelos A e B desenvolvidos no MIR foi realizada com 19 e 17 amostras apreendidas, respectivamente. Seus espectros foram testados nos modelos, os quais forneceram os valores de classe estimados para estas amostras, bem como os critérios para considerar a amostra como um *outlier*. Na tabela 11 são apresentados os resultados da aplicação do método em rotina. É possível observar que, das 19 amostras analisadas no modelo A, uma não continha NPS conforme resultados do método de referência, o que foi concordante com o resultado do método proposto. Somente uma amostra forneceu um resultado inconclusivo (amostra n.º 19), de modo que a taxa de predição correta obtida foi de 94,7%. As 17 amostras identificadas pelo modelo A como contendo drogas e que tiveram resultado conclusivo foram analisadas no modelo B, sendo observados três erros de identificação e duas com resultados inconclusivos (*outliers*), apesar de não serem. Considerando estes fatores, a taxa de predição correta obtida pelo modelo B foi de 82,4%.

**Tabela 11.** Resultados por GC-MS/LC-MS e ATR-FTIR para os Modelos A e B.

Amostra	Conteúdo	Modelo A			Modelo B		
		CA	CP	Resultado	CA	CP	Resultado
1	25B-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
2	25B-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
3	25B-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
4	25B-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
5	25B-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
6	25B-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
7	25B-NBOH	NPS	NPS	Ok	NBOH	NBOMe	X
8	25B-NBOH	NPS	NPS	Ok	NBOH	<i>Outlier</i>	I
9	25B-NBOH	NPS	NPS	Ok	NBOH	<i>Outlier</i>	I
10	25E-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
11	25E-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
12	25I-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
13	25I-NBOH	NPS	NPS	Ok	NBOH	NBOH	Ok
14	MDA	NPS	NPS	Ok	<i>Outlier</i>	<i>Outlier</i>	Ok
15	MPHP	NPS	NPS	Ok	<i>Outlier</i>	NBOMe	X
16	MPHP	NPS	NPS	Ok	<i>Outlier</i>	NBOMe	X
17	MDPHP	NPS	NPS	Ok	<i>Outlier</i>	<i>Outlier</i>	Ok
18	ND	B	B	Ok	---	---	---
19	ND	B	<i>Outlier</i>	I	---	---	---
Predição correta (%)				<b>94,7</b>	<b>82,4</b>		

B: Branco; CA: Classe atual; CP: Classe predita; I: Inconclusivo; ND: Não detectado;

OK: Resultado correto; X: Resultado errado; Outras drogas: MDA, MPHP e MDPHP.

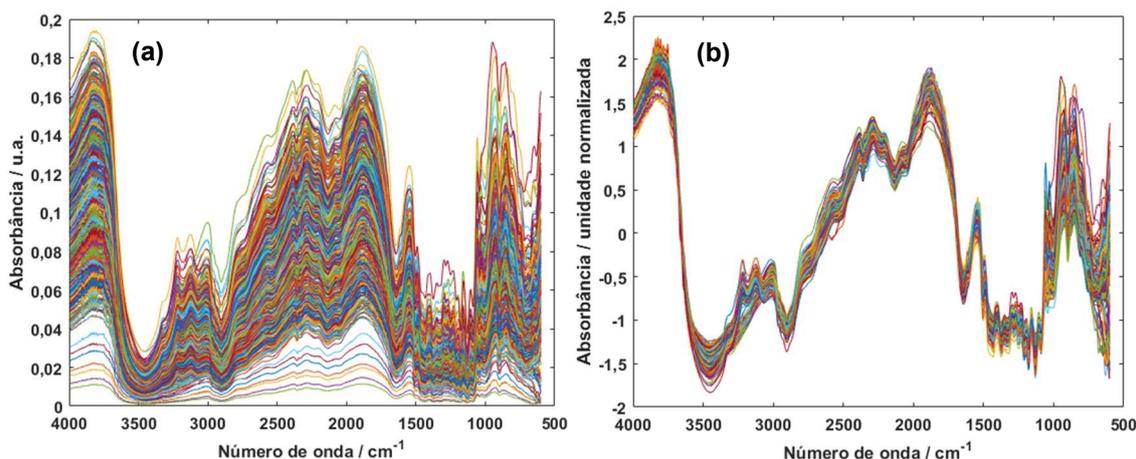
Quando a taxa de predição é calculada considerando-se somente as amostras de NBOMe e NBOH (que são as que definem as classes no desenvolvimento do modelo) e desconsiderando as amostras consideradas *outliers* pelo modelo B, seu valor se eleva para 92,3%, que é uma taxa de predição elevada.

Os resultados obtidos para os modelos A e B por MIR indicam que esta ferramenta pode ser aplicada em análises de rotina nos laboratórios forense, uma vez que, associando-se um espectrômetro como o Alpha II com um software de análise multivariada, o resultado para triagem da maioria das amostras é relativamente rápido e não destrutivo. Além disso, os modelos desenvolvidos podem ser continuamente atualizados com novas amostras apreendidas, de modo que se tornam mais robustos para futuras análises.

### **4.3. Avaliação da distribuição de NPS por imagem hiperespectral e MCR-ALS**

#### **4.3.1. Amostras simuladas**

Para a avaliação das imagens hiperespectrais, o pré-processamento que ofereceu melhores resultados foi a SNV, que foi aplicada para todas as imagens. Na figura 40(a) e 40(b) são apresentados os espectros brutos e espectros após a realização do pré-processamento por SNV de uma amostra simulada em que o 25E-NBOMe foi impregnado por gotejamento, respectivamente. É possível perceber que após o pré-processamento houve uma normalização dos espectros com uma redução significativa da variação entre eles. Todas as oito imagens de selos simulados foram avaliadas, havendo uma melhora no comportamento dos espectros e nos resultados obtidos em todos os casos.



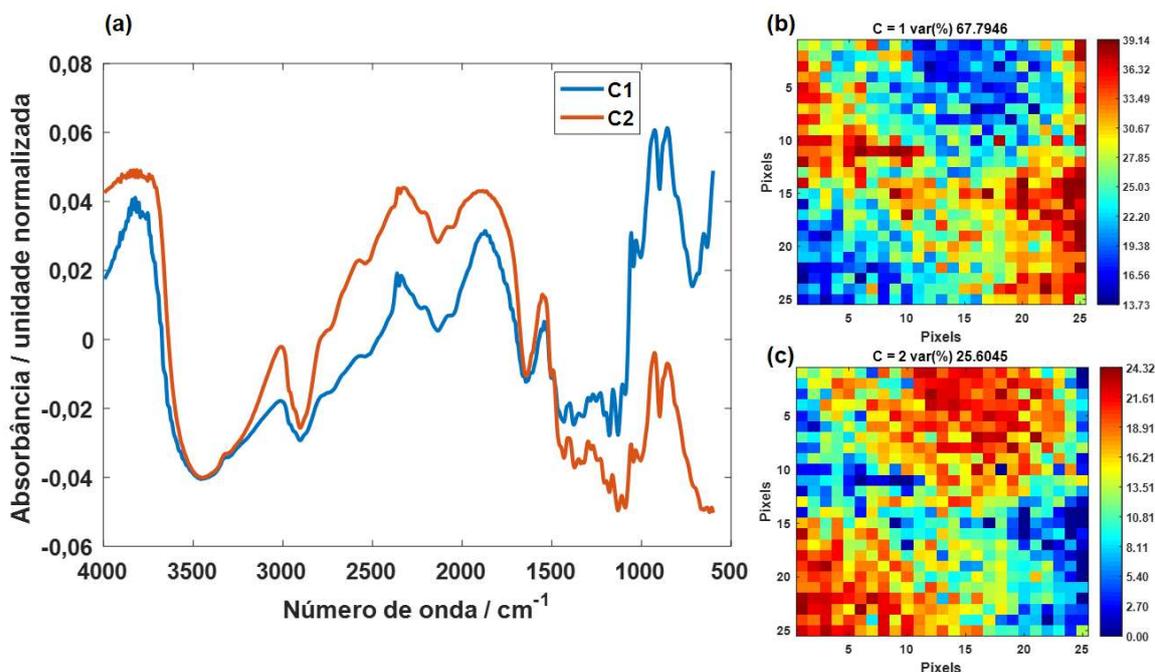
**Figura 40.** (a) Espectros brutos da amostra impregnada por gotejamento. (b) Espectros pré-processados por SNV da amostra impregnada por gotejamento.

As imagens que abrangeram toda a área mapeada, contendo 25 microsselos (25×25 *pixels*), apesar da quantidade de espectros, não apresentaram boa resolução espacial e resultados da decomposição por MCR. Como a amostra contendo 25 microsselos tem um tamanho 25 vezes maior que um único microsselo, mesmo aumentando o número de espectros de 400 (20×20) para 625 (25×25) a resolução espacial da imagem é muito prejudicada. Além disso, os dados contendo a informação química necessária para diferenciar 25E-NBOMe do papel também apresentaram redução da sua qualidade, devido principalmente à dificuldade em ajustar o foco para a aquisição dos espectros em toda a área mapeada. Outro fator que também pode ter influenciado os resultados nas amostras com mapeamento mais amplo foi a perda de material durante a impregnação por imersão e aspersão.

Na figura 41 são apresentados os espectros puros recuperados para a amostra simulada no modo imersão constituída por 25 microsselos e os mapas de concentração obtidos. Comparando esses resultados com os obtidos com a análise de um único microsselo (resultados apresentados em seguida), nesses espectros puros não foi possível observar diferenças significativas entre as bandas, que permitiam identificar os componentes um como sendo referente ao papel e outro referente ao 25E-NBOMe. As maiores diferenças observadas nos espectros puros recuperados estão relacionadas à diferença de intensidade de sinal, que podem estar relacionadas principalmente à variação do foco do equipamento por toda a amostra. Como a área de 25 microsselos é relativamente grande e a superfície dos selos não é totalmente plana, o caminho óptico variava ligeiramente no decorrer da área mapeada. Esta variação foi suficiente para influenciar a intensidade de sinal. Além das diferenças de intensidade já comentadas, também foram observadas diferenças nos espectros nas regiões de 3800-4000 cm⁻¹ e

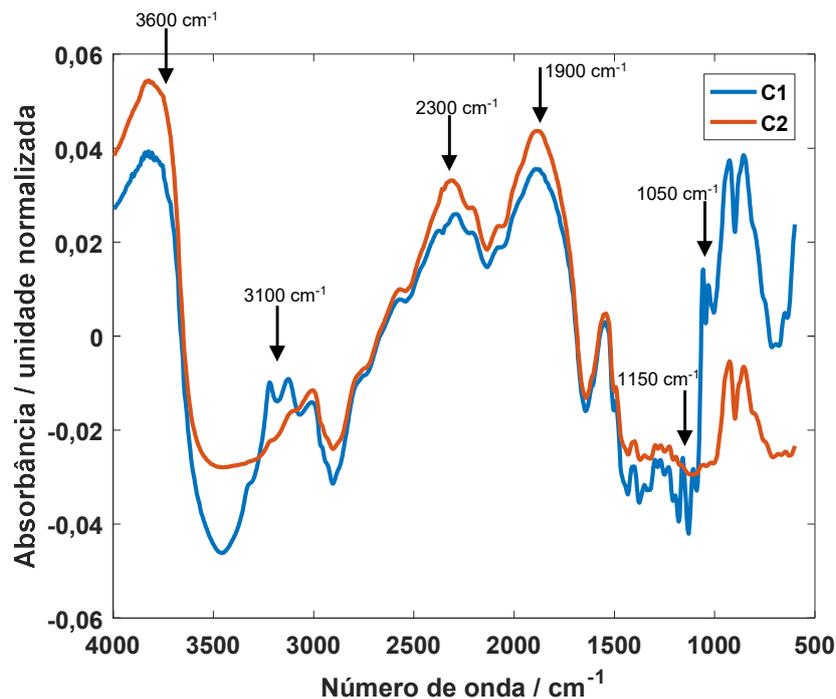
700-550  $\text{cm}^{-1}$ . Contudo, não foi possível realizar a atribuição dessas bandas a grupos orgânicos específicos ou encontrar uma explicação para essas diferenças.

Como este resultado foi observado para a maioria das amostras constituídas por 25 microsselos, decidiu-se avaliar os resultados somente para as imagens de microsselos que, apesar de terem sido obtidas com um número menor de espectros por imagem ( $20 \times 20$  pixels), eram imagens com um tamanho em torno de 25 vezes menor que as do selo completo, apresentando, portanto, maior resolução espacial.



**Figura 41.** (a) Espectros puros recuperados para o selo impregnado por imersão. (b) Perfil de concentração da componente 1 e (c) perfil de concentração da componente 2.

A imagem de microsselo que apresentou melhores resultados foi aquela obtida pela impregnação por gotejamento. A figura 42 apresenta os espectros puros recuperados para este caso, na qual foram observadas algumas bandas com diferença significativa que permitiu a atribuição de um espectro ao branco (suporte basicamente constituído de celulose) e outro ao 25E-NBOMe. É possível observar duas bandas próximas na região de 3200 a 3100  $\text{cm}^{-1}$  para um espectro da componente 1 (C1) que está ausente na componente 2 (C2). Estas bandas referem-se ao estiramento de amina primária, um grupamento que está presente na molécula de 25E-NBOMe e ausente na estrutura química da celulose, indicando que o componente 1 do MCR é o 25E-NBOMe. Estas bandas foram identificadas como principal “marcador” para diferenciar o espectro puro recuperado para o papel e para o 25E-NBOMe. Já o espectro referente a componente 2, com a ausência dessas bandas, apesar de pequenas proeminências, foi atribuído ao papel.



**Figura 42.** Espectros puros recuperados para o microselo impregnado por gotejamento. C1 é referente ao 25E-NBOMe e C2 é referente ao papel.

Além das bandas supracitadas, é observado que a grande maioria das bandas referentes à celulose estão presentes no espectro atribuído ao 25E-NBOMe. Uma possível justificativa para esse resultado é que a deconvolução por MCR não está conseguindo separar por completo o espectro do papel do espectro do 25E-NBOMe, de modo que bandas relativas à celulose ainda aparecem no espectro puro recuperado para o 25E-NBOMe. Isso ocorre devido a concentração de 25E-NBOMe ser de apenas 3,3% em média, de forma que o espectro recuperado e atribuído à droga é, na verdade, um espectro que também tem bandas referentes ao suporte.

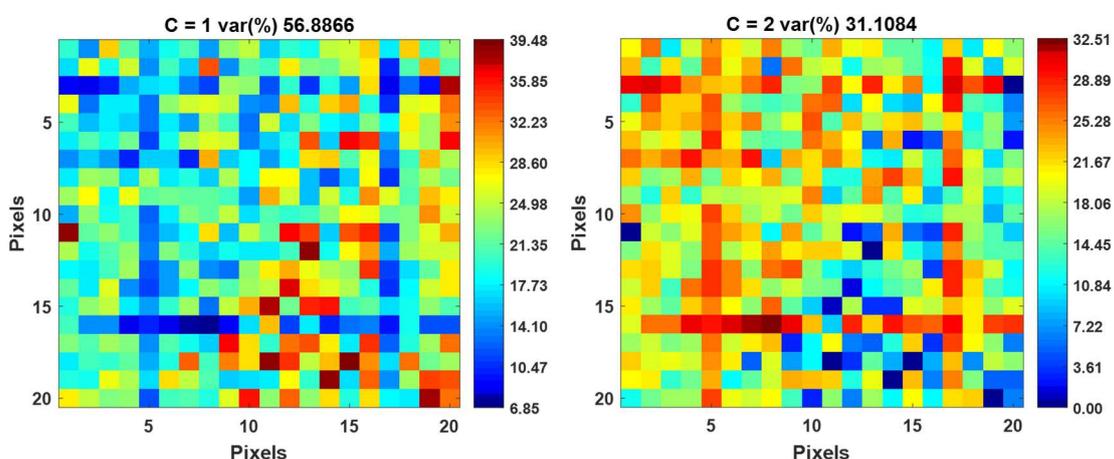
Existem duas bandas bem acentuadas entre 2300 e 1900  $\text{cm}^{-1}$ . Essas bandas aparecem em todos os espectros, independentemente de ser um espectro referente ao papel ou ao 25E-NBOMe. A banda em torno de 2300  $\text{cm}^{-1}$  é resultado da deformação axial simétrica da molécula do gás carbônico. Esse sinal foi observado em todas as amostras, mas teve sua intensidade suprimida em algumas delas, possivelmente devido à mistura de sinais. Entretanto, nas amostras sem impregnação, cujos resultados serão apresentados nas próximas seções, observou-se a banda de  $\text{CO}_2$  com maior destaque.

A banda de estiramento O-H, em torno de 3600  $\text{cm}^{-1}$ , sempre aparece em ambos os espectros puros recuperados. Há também uma banda em torno de 1150  $\text{cm}^{-1}$ , que está relacionada à deformação angular da ligação C-O de álcoois primários, presentes na celulose do papel, e uma banda em torno de 1050  $\text{cm}^{-1}$ , referente à vibração da

ligação C-O-C no anel da piranose presente na celulose. Nesta mesma região (de 1350 a 1000  $\text{cm}^{-1}$ ) também é possível fazer atribuição de banda para o estiramento C-N de amina, mas como as bandas nesta região são mais sutis nos espectros recuperados, não existe uma certeza na atribuição.

Adicionalmente, em alguns espectros puros recuperados que serão apresentados no decorrer desta seção, é perceptível uma pequena banda em torno de 1500  $\text{cm}^{-1}$  que pode estar relacionada ao estiramento C=C em anel aromático ou deformação angular da ligação N-H, presentes somente no 25E-NBOMe. Considerando essas informações, utilizou-se unicamente a banda de estiramento de amina (região de 3200 a 3100  $\text{cm}^{-1}$ ) como principal banda de diferenciação dos espectros puros recuperados.

O perfil de intensidade relativa obtido para cada uma das componentes na amostra simulada por gotejamento apresenta um perfil espelhado, uma vez que são definidas somente duas componentes: papel e droga. Isto significa que *pixels* que apresentam elevado valor de intensidade relativa de uma componente terão baixos valores de intensidade no perfil de outra componente. Os perfis de intensidade relativa obtidos são relacionados com os seus espectros puros. Cada mapa de concentração possui uma escala que vai das cores azul ao vermelho, indicando baixos e altos valores de intensidade relativa da componente, respectivamente. Quando, no mapa de intensidades relativas, existem áreas com abundância (cores mais próximas do vermelho), identifica-se ali a maior presença da componente em questão e menor presença da outra componente. Quando as áreas são de menor abundância (cores mais próximas do azul), o inverso ocorre. Os mapas de intensidade relativa para as componentes 1 (25E-NBOMe) e 2 (papel) para o microsselo por gotejamento são apresentados na figura 43.

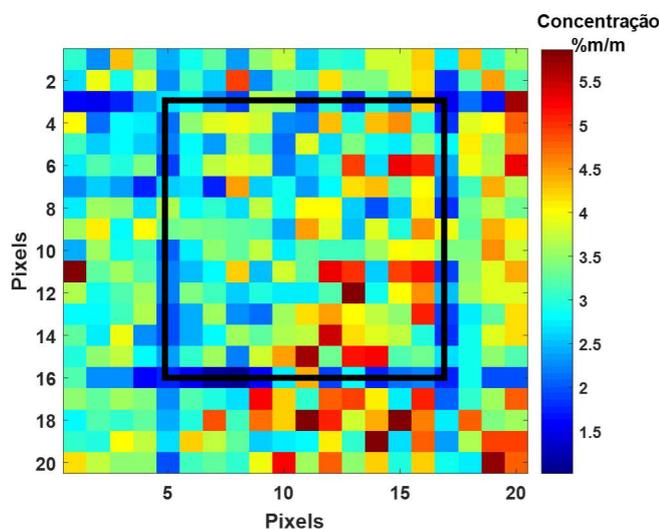


**Figura 43.** Perfis de concentração para o microsselo impregnado por gotejamento. C1 é referente ao 25E-NBOMe e C2 referente ao papel.

Os valores de intensidade relativa foram obtidos a partir da deconvolução dos sinais por MCR na interface Hyper-Tools. Partindo da suposição que a deconvolução conseguiu separar razoavelmente os compostos, as intensidades relativas são diretamente proporcionais à concentração do composto identificado em cada componente. Para tornar essa escala de melhor interpretação, os valores dos perfis de intensidade relativa da componente referente ao 25E-NBOMe foram convertidos para concentração em %m/m do composto nos selos. Essa conversão foi realizada por meio da equação 15.

$$CR_i = \frac{\overline{CR} \times IP_i}{\overline{IP}} \quad (\text{Equação 15})$$

Onde  $CR_i$  é a estimativa da concentração na unidade de %(m/m) obtida na conversão para cada pixel  $i$ ,  $\overline{CR}$  é a concentração média de 25E-NBOMe esperada com base no preparo das amostras simuladas (equivalente a 3,3% m/m),  $IP_i$  é a intensidade relativa do pixel  $i$  obtida a partir da deconvolução dos sinais e  $\overline{IP}$  é a intensidade relativa média obtida pela média aritmética das intensidade relativas de todos os pixels da imagem relativa ao 25E-NBOMe. Após a conversão da escala para concentração, o mapa para a componente 1, relativa ao 25E-NBOMe, é apresentado na figura 44.



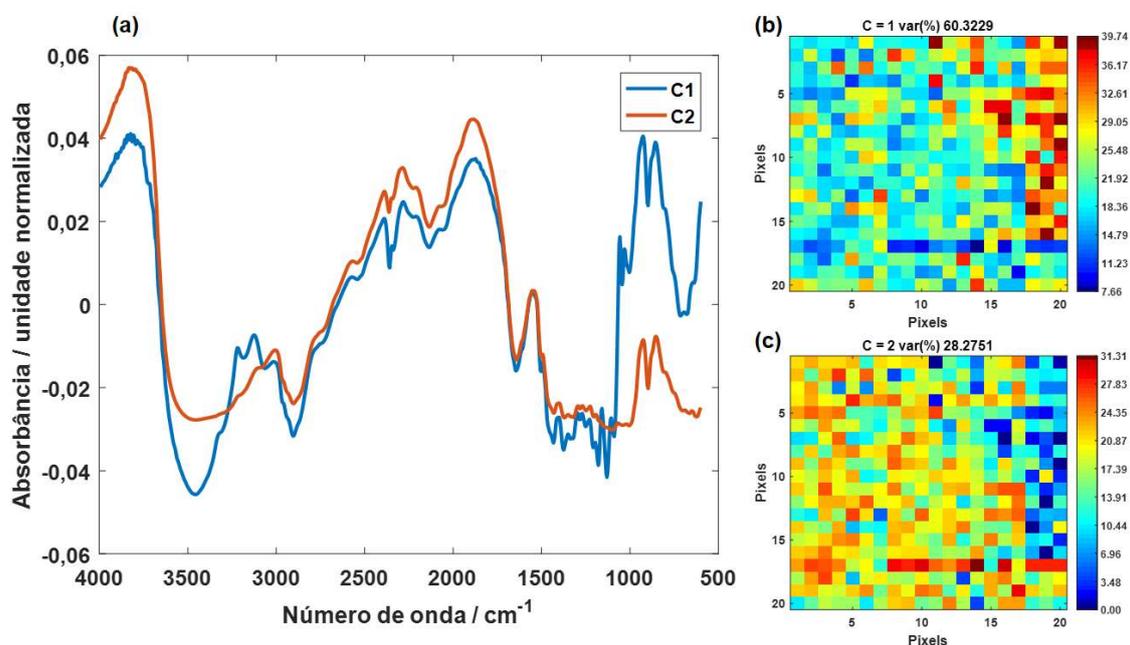
**Figura 44.** Perfil de concentração do 25E-NBOMe para o selo impregnado por gotejamento com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo. Marcação em preto indica a região da parte picotada do microselo.

Na figura 44, é possível observar uma marcação quadrada de coloração predominantemente azul (indicada na imagem). Essa marcação é exatamente a delimitação do microselo, isto é, a parte picotada, na qual o microselo é destacado dos outros. Esta colocação azul que indica baixas concentrações relativas de 25E-NBOMe podem também ter sido obtidas devido à diferença de foco nesta região do

papel, que é diferente da parte central do microsselo. Na mesma figura observam-se áreas na parte inferior direita com concentrações relativas de 25E-NBOMe mais elevadas (coloração vermelha). Este resultado indica que a maior parte da substância impregnada neste microsselo se depositou nesta área. A partir dessas observações é possível afirmar que a distribuição da substância no microsselo, neste caso, não foi homogênea, havendo um depósito maior de 25E-NBOMe em um dos lados do microsselo.

A partir dos resultados obtidos por gotejamento, foi realizada a deconvolução selecionando as variáveis referentes à região das bandas de estiramento de amina primária (3200 a 3100  $\text{cm}^{-1}$ ). Contudo, foi observado que para algumas amostras o uso da região selecionada dificultava a resolução do MCR, pois foram obtidos mapas de intensidade relativa com comportamento aleatório, sem evidenciar nem a região picotada observada na figura 44. Por essa razão, escolheu-se trabalhar com toda a região espectral.

Resultado semelhante ao microsselo impregnado por gotejamento foi obtido para o microsselo impregnado por imersão, porém com uma leve diminuição na banda que diferencia 25E-NBOMe do papel (entre 3200 a 3100  $\text{cm}^{-1}$  referente à amina primária). Os espectros puros e os mapas de intensidade para as componentes 1 (C1) e 2 (C2) são apresentados na figura 45.

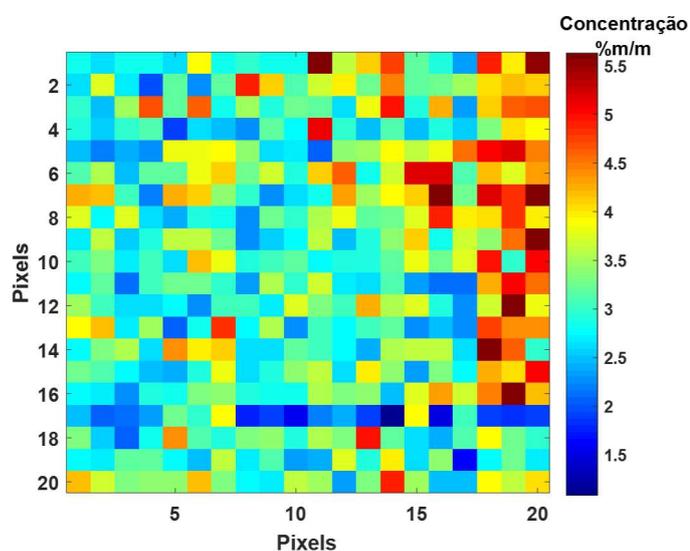


**Figura 45.** (a) Espectros puros recuperados para o microsselo impregnado por imersão. (b) Perfil de concentração da componente 1 e (c) perfil de concentração da componente 2.

Os espectros puros recuperados para a amostra impregnada por imersão são bem semelhantes aos recuperados para a amostra impregnada por gotejamento. A

diferença mais visível se encontra na região em torno de  $2300\text{ cm}^{-1}$ , na qual surge um pequeno vale para a amostra de imersão. Essa diferença se dá provavelmente pela influência de  $\text{CO}_2$  do ambiente, possivelmente pela menor supressão desta banda pelo espectro do 25E-NBOMe.

Tanto na amostra por gotejamento, como na amostra por imersão, a primeira componente foi identificada como referente ao 25E-NBOMe devido à presença da banda de amins primárias. O fato da droga sintética ter sido atribuída à componente 1 também é resultado da maior taxa de variância explicada, que está relacionada com a variação dos sinais devido a mudança nas concentrações de 25E-NBOMe ao longo da área mapeada. A figura 46 apresenta o mapa em unidades de %m/m de 25E-NBOMe obtido para a amostra de imersão.

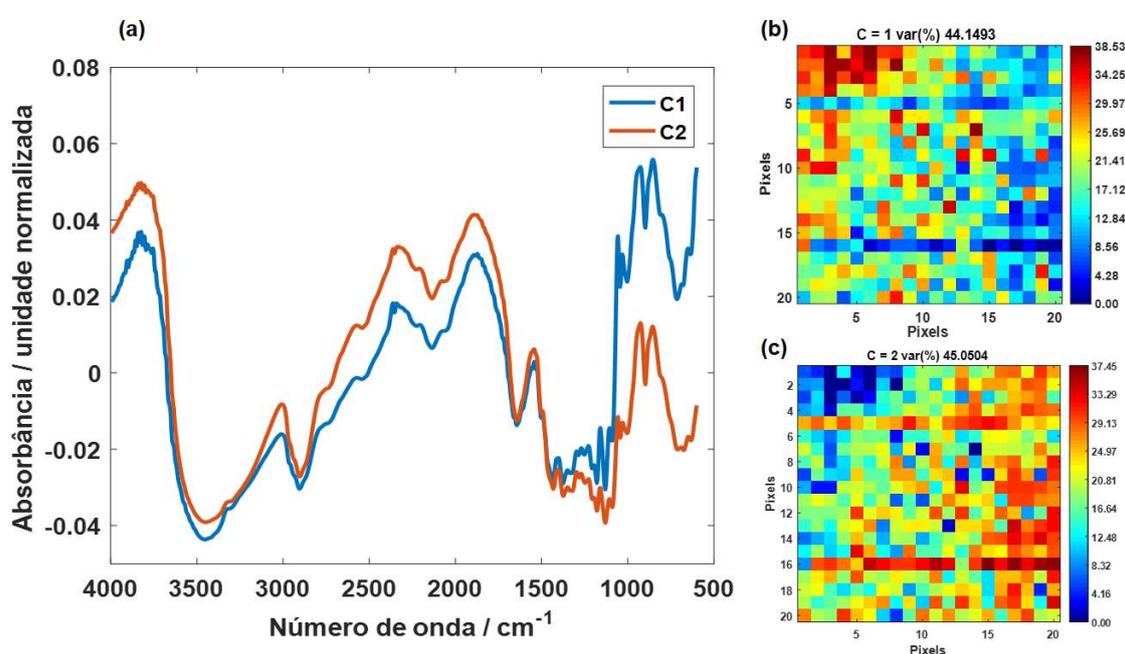


**Figura 46.** Perfil de concentração do 25E-NBOMe para o selo impregnado por imersão com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo.

Na figura 46, assim como no perfil de concentração obtido para a amostra impregnada por gotejamento (figura 44), observa-se também a demarcação da região serrilhada do microsselo, mais evidente na parte inferior (coloração azul contínua) com baixas concentrações de 25E-NBOMe. Já na parte superior direita, observa-se um acúmulo da substância, evidenciado pela coloração vermelha que indica concentrações mais elevadas de 25E-NBOMe. Como na amostra impregnada por gotejamento, esta, impregnada por imersão, apresentou um perfil heterogêneo de distribuição da substância no microsselo.

No caso do microsselo impregnando por aspersão, a avaliação dos espectros puros recuperados não foi considerada eficiente. Neste caso não foi observada a presença da banda dupla em torno de  $3100$  a  $3200\text{ cm}^{-1}$  que diferencia a celulose do

25E-NBOMe. A figura 47(a) apresenta as estimativas dos espectros puros recuperados que, apesar de apresentarem diferenças de intensidade de sinal em algumas bandas, não possibilitam fazer nenhum tipo de atribuição de bandas para vibrações de grupos orgânicos específicos que diferenciasssem a substância sintética do papel. Considerando este fato, a interpretação dos perfis de intensidade relativa para este microselo também fica comprometida, uma vez que esta depende da obtenção de uma boa deconvolução dos componentes. Não é possível se afirmar com toda a certeza que determinada área com abundância em concentração de C1 se refere a altas concentrações de 25E-NBOMe ou celulose, pois não se pode afirmar que C1 é 25E-NBOMe e C2 é celulose pelos espectros puros ou vice-versa. Os mapas de concentração para C1 e C2 são apresentados nas figuras 47(b) e 47(c).

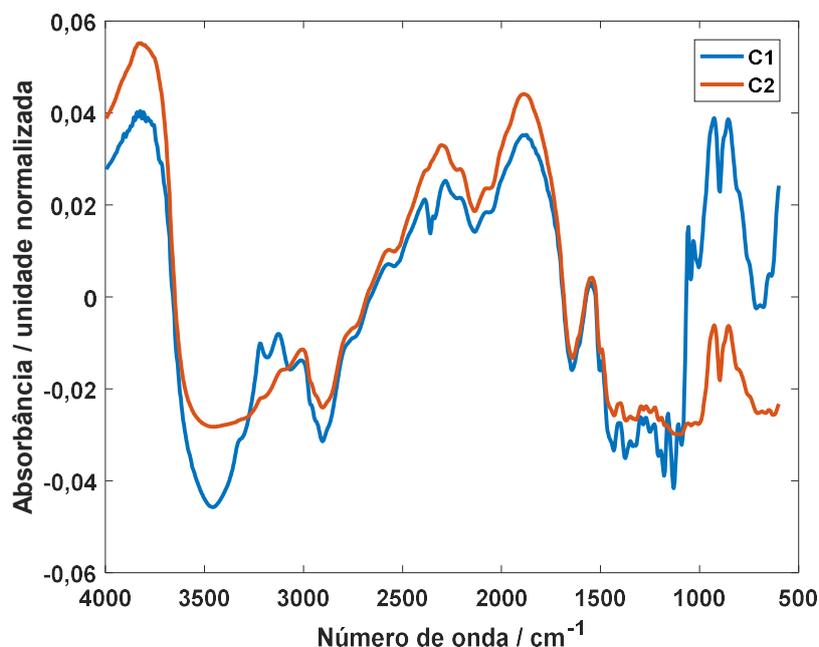


**Figura 47.** (a) Espectros puros recuperados para o microselo impregnado com 25E-NBOMe por aspersão. (b) Perfil de concentração da componente 1 e (c) perfil de concentração da componente 2.

É possível que esta ineficiência de recuperação dos espectros puros para a amostra de aspersão esteja relacionada a forma com a qual o 25E-NBOMe foi impregnado no selo. Como a aspersão é um processo no qual é inevitável a perda de material, deve-se considerar que, durante a produção da amostra, uma quantidade significativa da solução não foi impregnada no selo, de modo que a concentração da droga no microselo foi efetivamente menor, comparando-se com os microselos impregnados por gotejamento e imersão. Devido ao limite de detecção da técnica por infravermelho ser consideravelmente alto (em torno de 0,1% m/m, dependendo da absorvidade da molécula) e, associando-a a uma análise multivariada que tenta

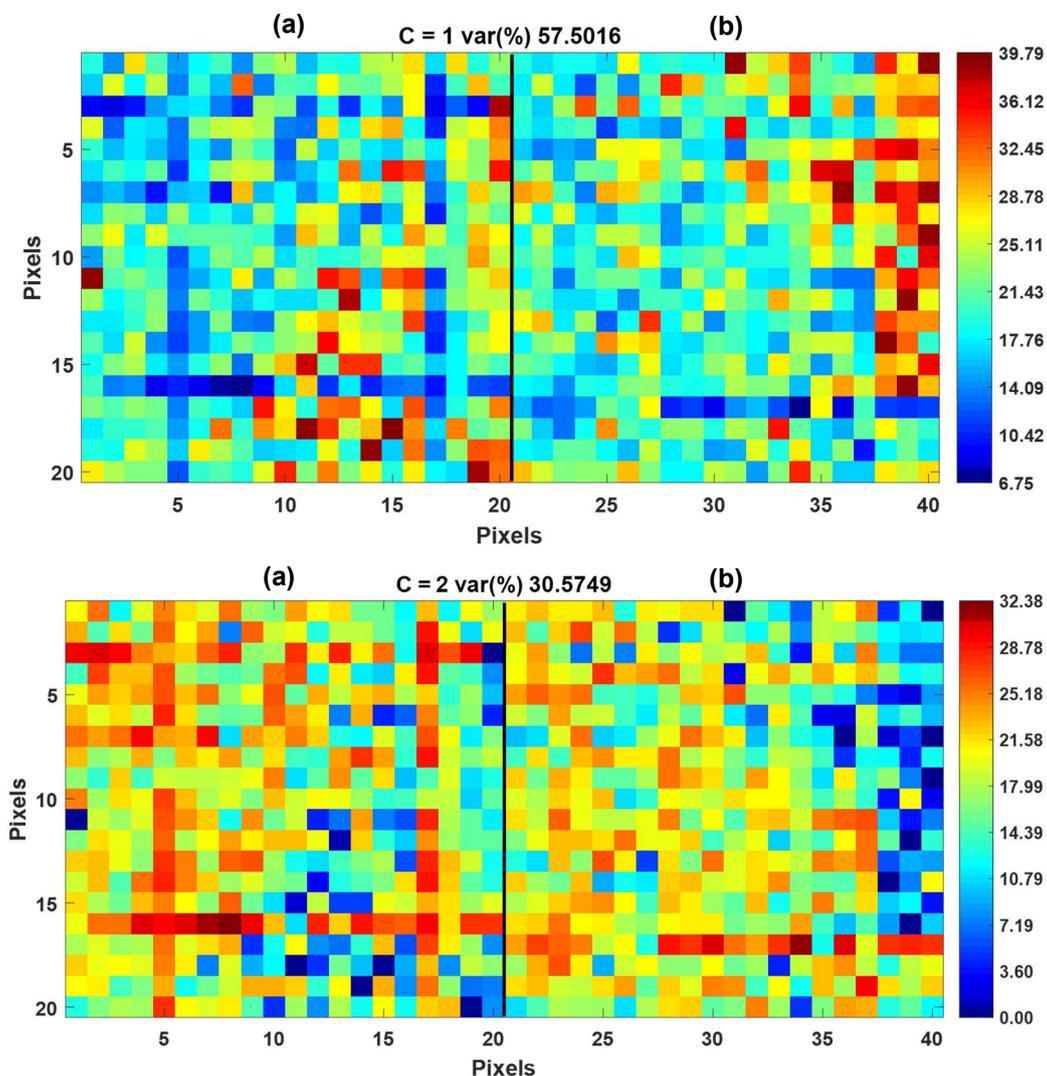
recuperar espectros puros em uma mistura, a chance de se obter um resultado que seja capaz de diferenciar com confiança os dois componentes é pequena. Como a quantidade de 25E-NBOMe disponível para novas impregnações não era suficiente, seu valor de mercado é relativamente elevado (cerca de R\$ 2000,00 por 50 mg do padrão) e a compra desse tipo de padrão analítico é dificultado por se tratar de uma substância proscrita, não foi possível realizar a produção de novas amostras para uma segunda avaliação.

Considerando que, das três formas de impregnação, somente as realizadas por gotejamento e imersão apresentaram resultados favoráveis, decidiu-se realizar a decomposição por MCR por meio de uma matriz aumentada. Neste caso, os dados obtidos para os microselos por gotejamento e imersão são dispostos lado a lado na própria interface do HYPER-Tools, gerando uma matriz com 800 linhas (20×40 pixels) e 833 colunas (números de onda). Esta matriz aumentada foi decomposta por MCR utilizando dois componentes, gerando-se dois espectros puros e perfis de intensidade relativa para os dois componentes em cada amostra (gotejamento e imersão). A princípio, as vantagens do uso de matrizes aumentadas são: (1) obtenção dos mesmos perfis espectrais para as duas amostras analisadas simultaneamente, o que permite uma melhor comparação dos resultados e (2) aumento no número de espectros utilizados na deconvolução, o que fornece mais informação para o algoritmo. Os espectros puros recuperados para a análise aumentada/conjunta dos microselos por gotejamento/imersão são apresentados na figura 48.



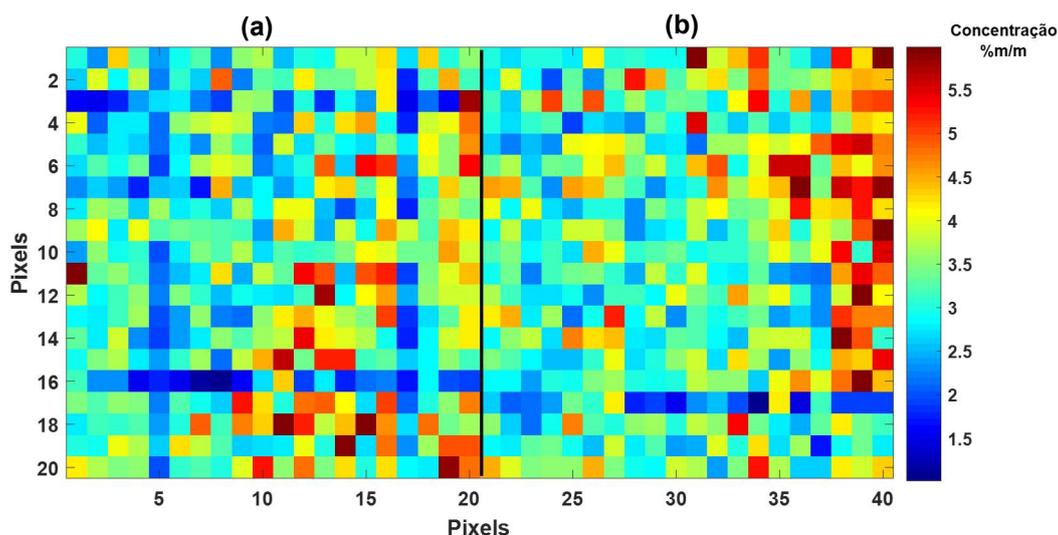
**Figura 48.** Espectros puros recuperados para a matriz de dados aumentada por gotejamento e imersão.

Os espectros puros apresentados na figura 48 são bem semelhantes àqueles apresentados nas figuras 42 e 45, referentes às amostras impregnadas por gotejamento e imersão que foram analisadas de forma separada. Os perfis de concentração das duas componentes obtidos após a deconvolução dos sinais da matriz aumentada são apresentados na figura 49.



**Figura 49.** Perfis de concentração para o microselo impregnado por (a) gotejamento e (b) imersão a partir da matriz aumentada. C1 é referente ao 25E-NBOMe e C2 referente ao papel.

Também neste caso, a escala para o perfil de concentração do 25E-NBOMe (C2) foi convertida para concentração em %m/m de 25E-NBOMe no selo e o mapa com a escala de cores convertida é apresentado na figura 50.



**Figura 50.** Perfil de concentração do 25E-NBOMe para o selo impregnado por (a) gotejamento e (b) imersão (matriz aumentada) com conversão da escala de cores para concentração em %m/m de 25E-NBOMe no selo.

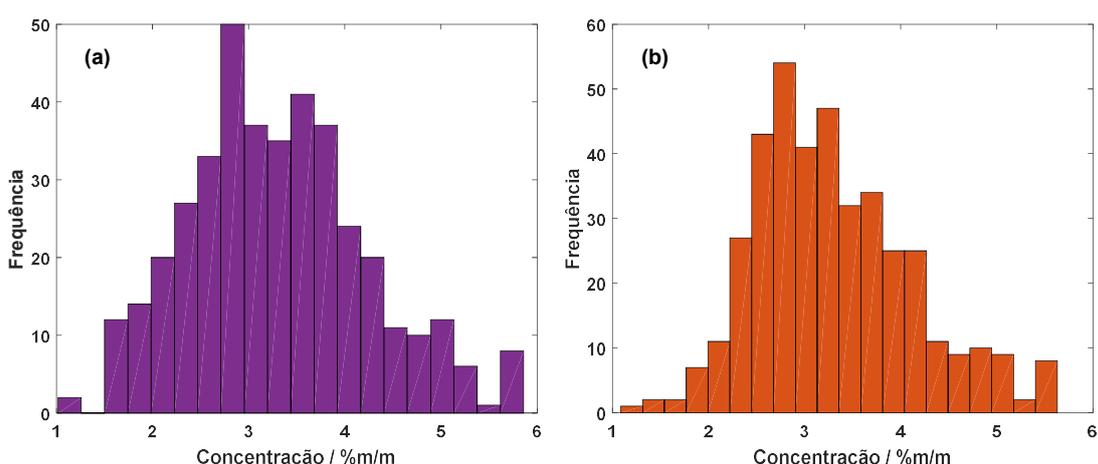
Observando o mapa de intensidade relativa para o 25E-NBOMe (C1) da matriz aumentada, é possível perceber uma semelhança visual com os mapas de concentração do 25E-NBOMe por gotejamento e imersão separadamente. O mapa da figura 50(a) apresenta maiores concentrações de 25E-NBOMe no canto inferior direito, assim como o mapa da C1 obtida por gotejamento (figura 43). O mapa da figura 50(b) apresenta maiores concentrações de 25E-NBOMe no canto superior direito, o que se assemelha ao mapa da C1 obtida por imersão (figura 45). Este resultado reforça a confiança nos resultados obtidos para os mapeamentos dos microselos de imersão e gotejamento, uma vez que a deconvolução utilizando dois procedimentos diferentes convergiu em resultados visualmente similares. Devido as vantagens destacadas anteriormente, apesar dos resultados visualmente semelhantes, sugere-se que o resultado obtido por matriz aumentada seja o que mais deve se aproximar das concentrações reais de 25E-NBOMe nos microselos, uma vez que este foi obtido pela deconvolução conjunta de um maior número de espectros.

#### 4.3.1.1. Avaliação dos histogramas dos mapas de concentração e teste de hipótese para as amostras simuladas obtidas gotejamento e imersão

A fim de se comparar a distribuição dos perfis de concentração dos mapas de 25E-NBOMe obtidos nos microselos, com um nível de confiabilidade definido, foram gerados os histogramas dos mapas obtidos por gotejamento, aspensão e na

deconvolução da matriz aumentada por gotejamento/imersão. Os histogramas apresentam a frequência com a qual um determinado valor de concentração aparece na imagem recuperada após a aplicação do MCR. A partir deste ponto, para melhor clareza no texto, os perfis de concentração obtidos pela deconvolução das matrizes individuais por gotejamento e imersão serão indicados por G1 e I1, respectivamente. Já os perfis de concentração da matriz aumentada por gotejamento e imersão serão indicados por G2 e I2, respectivamente.

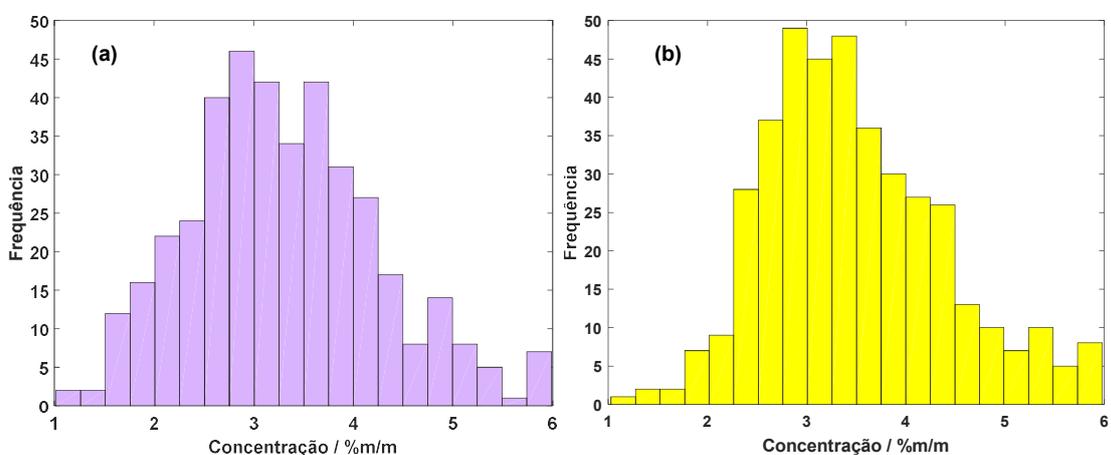
Na figura 51 são apresentados os histogramas de G1 e I1. É possível perceber que a região com maior frequência de concentração é de 2,5 a 3,0% m/m de 25E-NBOMe nos selos. Com o objetivo de se comparar estatisticamente os histogramas obtidos, a fim de se concluir se as distribuições de 25E-NBOMe nos selos apresentam padrões semelhantes ou não, aplicou-se o teste de hipótese Kolmogorov-Smirnov (*K-S test*), que é um teste não-paramétrico, isto é, aplicado para populações cujas distribuições não são características ou conhecidas.<sup>80</sup> Este teste determina se dados de amostras randômicas e independentes são provenientes de uma mesma população contínua. Se o teste fornece o valor zero ( $H = 0$ ), a hipótese nula não é rejeitada a um nível de significância de 95% ( $p = 0,05$ ), ou seja, as populações comparadas são semelhantes neste nível de confiança. Se o oposto ocorrer, e o teste fornecer valor igual a um ( $H = 1$ ), a hipótese nula é rejeitada, aceitando-se a hipótese alternativa, que considera as populações diferentes em um nível de confiança de 95%. Neste caso, o teste de hipótese de Kolmogorov-Smirnov foi aplicado comparando-se duas amostras, sendo um teste biamostrai. O teste também pode fornecer o valor de  $p$  (ou  $p$ -valor), que é a probabilidade, variando de 0 a 1, para que a hipótese nula seja aceita.<sup>81</sup>



**Figura 51.** Histogramas dos perfis de concentração obtidos a partir da deconvolução de dados por MCR das imagens dos selos impregnados por (a) gotejamento e (b) imersão por meio de matrizes simples.

A aplicação do teste K-S para as amostras por gotejamento e imersão (G1 e I1) forneceu  $H = 0$  e  $p = 0,0976$ , ou seja, a hipótese nula foi aceita e as distribuições são estatisticamente equivalentes a um nível de 95% de confiança. Este resultado indica que, pelas amostras analisadas, a forma com a qual a droga foi impregnada (gotejamento ou imersão) não difere estatisticamente. Este resultado contraria um pouco a observação experimental de que a impregnação por gotejamento garante que não haja perda de material, enquanto na impregnação por imersão, visivelmente ocorrem perdas. Outra hipótese para esse resultado é que a região mapeada deveria ser maior para se ter uma melhor representatividade para avaliar a homogeneidade.

Optou-se, portanto, por realizar o mesmo teste comparando-se os perfis de concentração obtidos para gotejamento e imersão por meio da matriz aumentada, a fim de se avaliar, se, nesta forma de deconvolução das duas amostras simultaneamente, há semelhança ou não nas distribuições a um nível de 95% de confiança. Os histogramas dos perfis de G2 e I2 são apresentados na figura 52. Neste caso, foi possível perceber visualmente que há uma diferença entre os perfis desses histogramas. Esta diferença foi confirmada pelo teste K-S, de modo que, para este caso,  $H = 1$  e  $p = 0,0341$ , ou seja, rejeitou-se a hipótese nula e distribuição dos dois conjuntos de dados diferem a um nível de 95% de confiança.



**Figura 52.** Histogramas dos perfis de concentração obtidos a partir da deconvolução de dados por MCR das imagens dos selos impregnados por (a) gotejamento e (b) imersão por meio de matriz aumentada.

Foram também avaliados se os perfis de concentração obtidos por gotejamento da matriz simples (G1) e gotejamento da matriz aumentada (G2) diferem ou não entre si, e, para tanto, também se aplicou o teste K-S. O objetivo é que o teste confirme se os resultados obtidos nas diferentes deconvoluções são ou não estatisticamente

diferentes. Entre o histograma do perfil de concentração da matriz por gotejamento (G1) e da matriz aumentada por gotejamento (G2), o teste K-S forneceu  $H = 0$  e  $p = 1$ , ou seja, a hipótese nula foi aceita e as distribuições são semelhantes a um nível de 95% de confiança, indicando que independente da matriz ser simples ou aumentada, a distribuição da droga é modelada de forma semelhante para a amostra impregnada por gotejamento. Resultado contrário foi obtido para os histogramas do perfil de concentração da matriz simples por imersão (I1) e da matriz aumentada por imersão (I2), sendo  $H = 1$  e  $p = 0,0145$ . Os testes realizados indicam que a diferença na distribuição da substância no selo, devido à forma de impregnação, só é estatisticamente indicada por meio da matriz aumentada. Mesmo assim, essa diferença ocorre em baixas probabilidades (0,0341 para G2 contra I2 e 0,0145 para I1 contra I2). A tabela 12 resume os resultados obtidos pelo teste K-S para os histogramas de perfis de concentração.

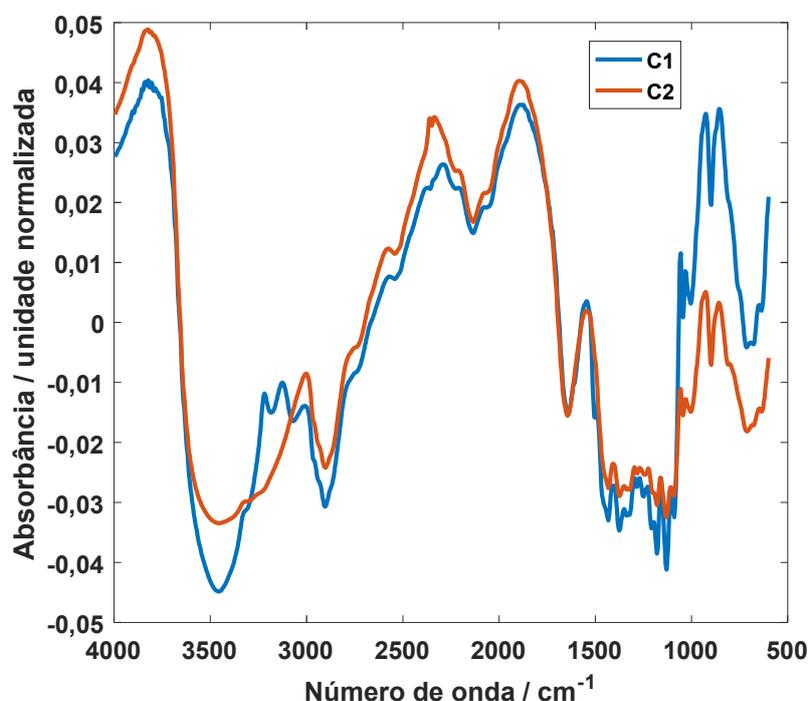
**Tabela 12.** Resumo dos resultados do teste de hipótese de Kolmogorov-Smirnov (teste K-S) para as amostras de gotejamento e imersão por matrizes simples e aumentada.

Comparação	H	Resultado
G1 × I1	0	Distribuições semelhantes
G2 × I2	1	Distribuições distintas
G1 × G2	0	Distribuições semelhantes
I1 × I2	1	Distribuições distintas

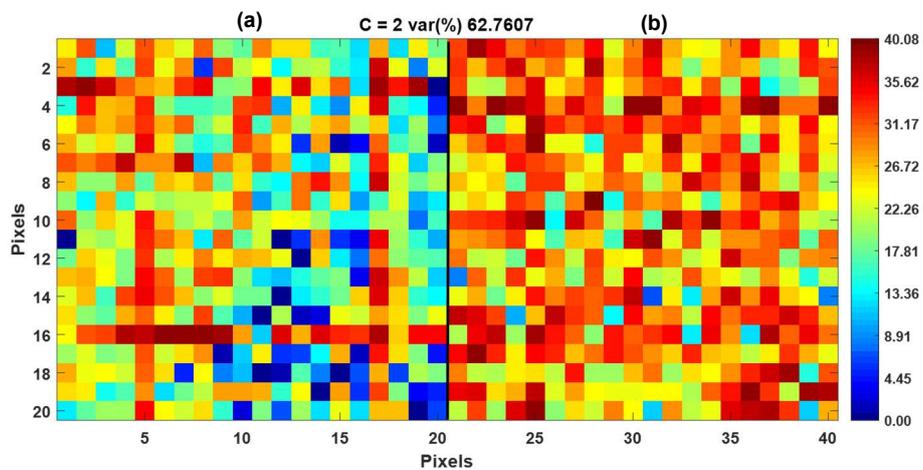
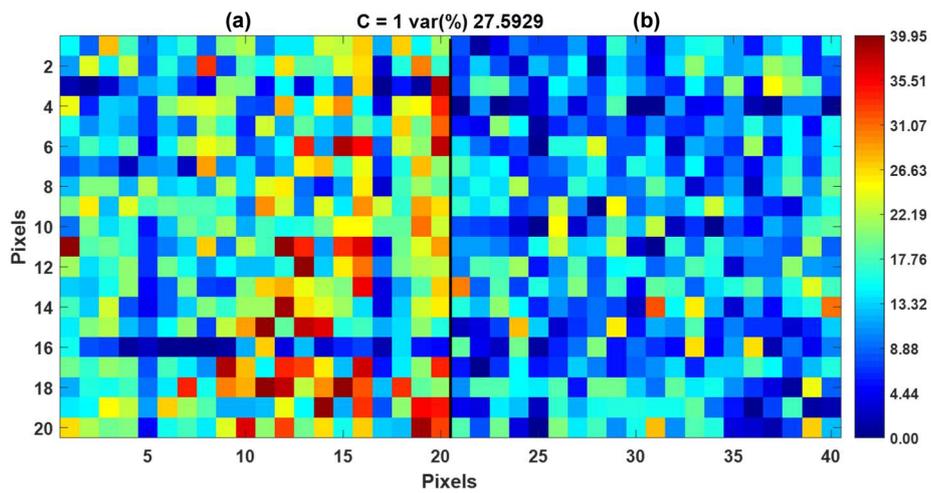
Os resultados reforçam que a análise com matriz aumentada apresenta resultados mais condizentes com o esperado experimentalmente. Houve uma melhor diferenciação entre as duas formas de impregnação (gotejamento e imersão), que não foi observada quando o conjunto de dados foi analisado separadamente. É importante destacar que estes resultados são preliminares, devido à quantidade de amostras simuladas que puderam ser produzidas. Contudo, os resultados podem ser considerados promissores, indicando que a análise hiperespectral na região do infravermelho pode estimar a distribuição espacial de 25E-NBOMe em selos e diferenciar entre o processo de impregnação.

#### 4.3.2. Avaliação das imagens de selo contendo 25E-NBOMe e selos não impregnados

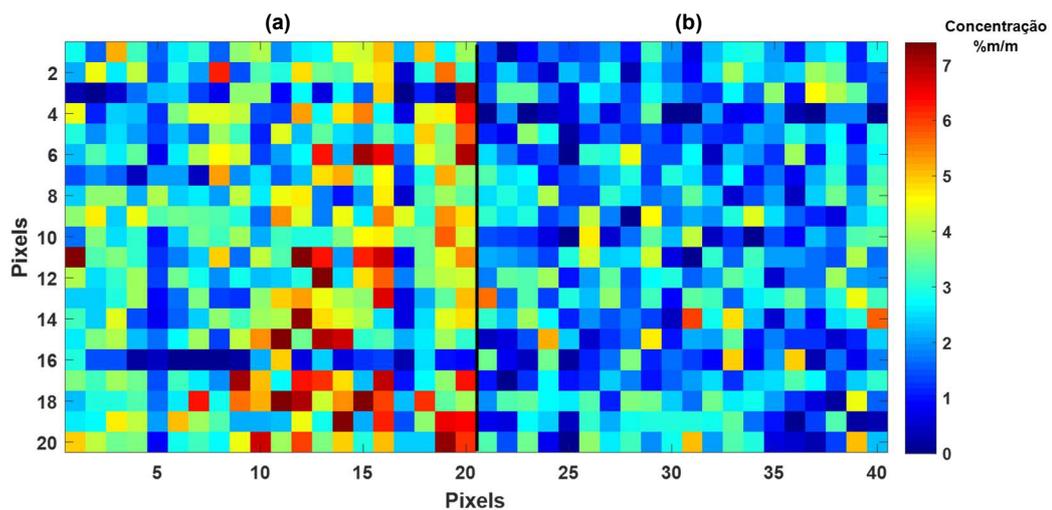
Para avaliar a possibilidade de erros falso positivos na análise de selos sem nenhuma droga impregnada, matrizes aumentadas de selos impregnados por gotejamento, imersão e selos sem impregnação de 25E-NBOMe foram analisadas por MCR. Duas matrizes aumentadas foram geradas, utilizando-se o mapeamento de microselos: gotejamento/branco e imersão/branco. É importante salientar que o microselo sem impregnação passou pelo mesmo processo que o microselo impregnado, porém sem a adição do 25E-NBOMe. Ou seja, ele foi gotejado com metanol e posteriormente o solvente foi evaporado à temperatura ambiente. Os espectros puros recuperados para os microselos por gotejamento/branco e seus mapas de intensidade relativa são apresentados nas figuras de 53, 54 e 55.



**Figura 53.** Espectros puros recuperados para a matriz aumentada de microselos impregnados por gotejamento e sem impregnação (branco). C1 corresponde ao espectro do papel e C2 ao espectro do 25E-NBOMe.



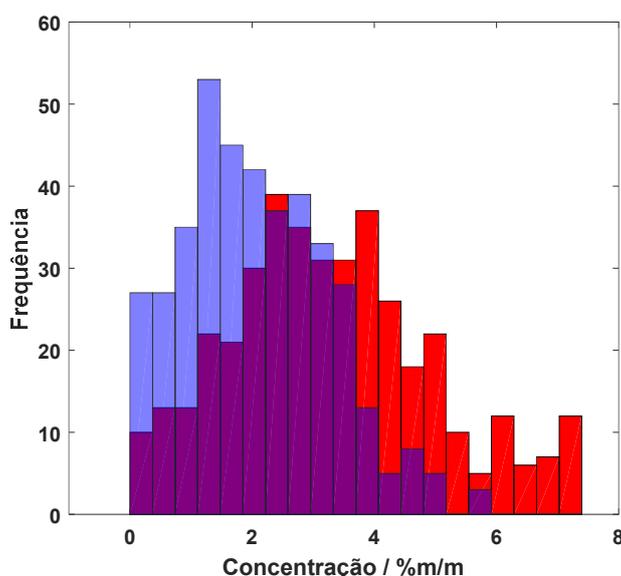
**Figura 54.** Perfis de concentração para o microsselo impregnado por (a) gotejamento e (b) sem impregnação a partir da matriz aumentada. C1 é referente ao 25E-NBOME e C2 referente ao papel.



**Figura 55.** Perfis de concentração para os microsselos impregnados por (a) gotejamento e (b) sem impregnação a partir da matriz aumentada com conversão da escala de cores para concentração em %m/m de 25E-NBOME no selo.

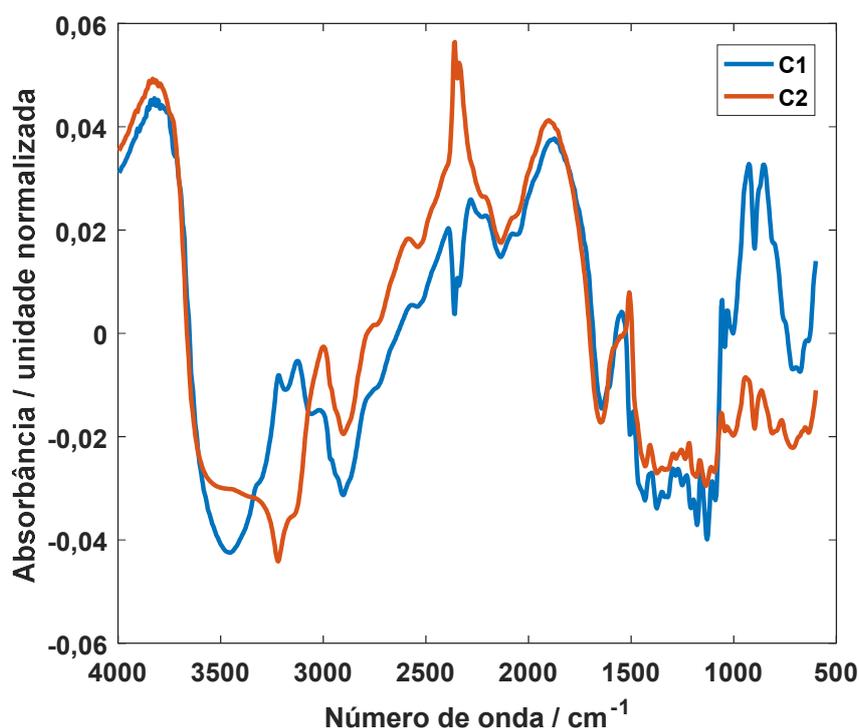
É possível observar que, para a matriz aumentada gotejamento/branco, os espectros puros estimados são bem diferentes (figura 53), principalmente na região referente ao estiramento da amina ( $3200\text{ cm}^{-1}$ ), na qual existe uma banda no espectro referente ao 25E-NBOMe que está ausente no espectro referente ao papel. Comparando-se esses espectros com aqueles obtidos para a matriz simples por gotejamento (figura 42), é possível perceber que não há as proeminências na banda que coincidem com a banda de estiramento da amina primária no espectro referente ao papel. Pode-se observar visualmente que o mapa de intensidade do microselo impregnado por gotejamento se manteve com o mesmo comportamento em relação aos outros perfis de concentração desta amostra já avaliados anteriormente. Já o perfil de concentração para o selo não impregnado (branco) se apresenta mais homogêneo, com concentrações de 25E-NBOMe menores que 3% na maioria dos pixels (79% dos pixels), o que é esperado, uma vez que não houve impregnação desta substância. A presença de alguns pixels indicando concentrações mais elevadas de 25E-NBOMe no microselo sem impregnação evidenciam um erro de deconvolução do MCR, que identifica intensidade relativas elevadas de uma componente, apesar da ausência da droga no microselo (figura 55). Contudo, como destacado, essas maiores concentrações foram obtidas apenas em poucos *pixels*.

Ao se avaliar os histogramas obtidos a partir da matriz aumentada gotejamento/branco (figura 56), é possível observar um claro perfil distinto entre eles, o que é confirmado aplicando-se o teste K-S,  $H = 1$  e  $p = 1,10 \times 10^{-19}$ .

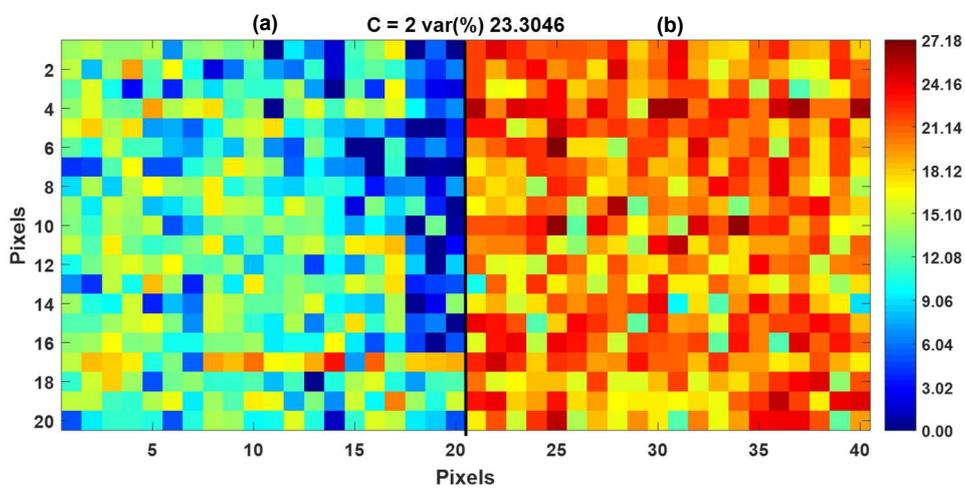
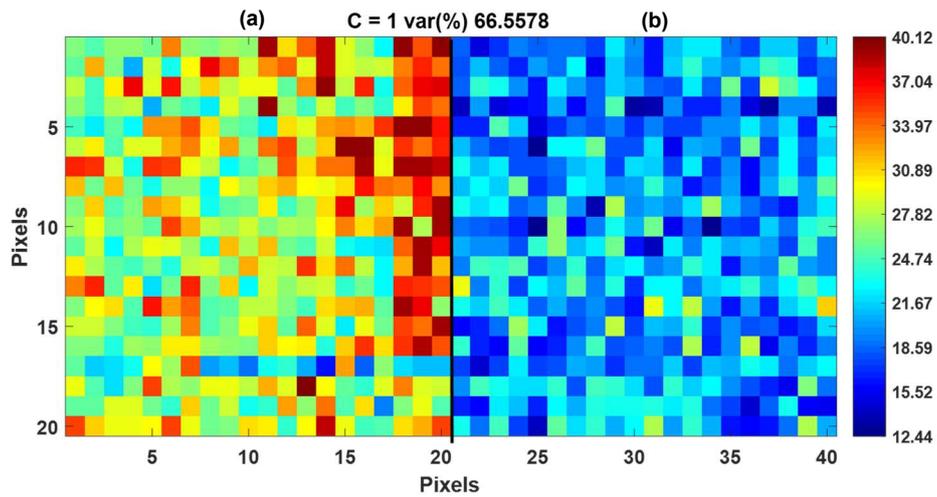


**Figura 56.** Histogramas sobrepostos dos perfis de concentração obtidos a partir da deconvolução de dados por MCR da matriz aumentada dos microselos impregnados por gotejamento (em vermelho) e sem impregnação (em azul).

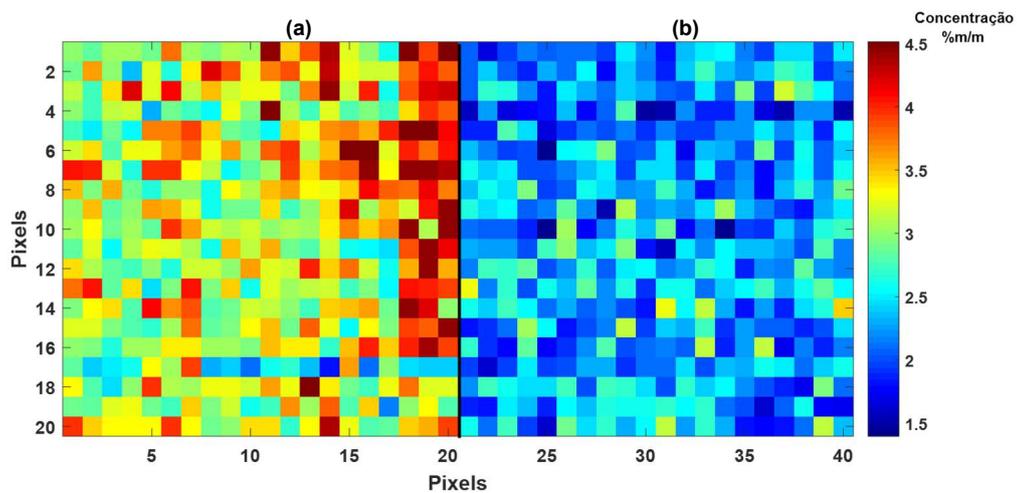
Os resultados obtidos para a matriz aumentada imersão/branco foram ligeiramente diferentes dos resultados da matriz aumentada gotejamento/branco. Isso já é observado pelos espectros puros recuperados para esta matriz aumentada (figura 57). Neste caso é evidente a banda de CO<sub>2</sub> em 2360 cm<sup>-1</sup>, que aparece na componente do papel. Já o perfil de intensidade relativa do microsselo impregnado por imersão é semelhante aos avaliados anteriormente. Observa-se que o perfil de intensidade do branco apresentou menores concentrações e distribuição mais homogênea com apenas poucos pixels (3,5%) indicando concentrações maiores que 3 % m/m, que se referem a erros de deconvolução do MCR (figura 59).



**Figura 57.** Espectros puros recuperados para a matriz aumentada de microsselos impregnados por imersão e sem impregnação (branco). C1 corresponde ao espectro do papel e C2 ao espectro do 25E-NBOMe.

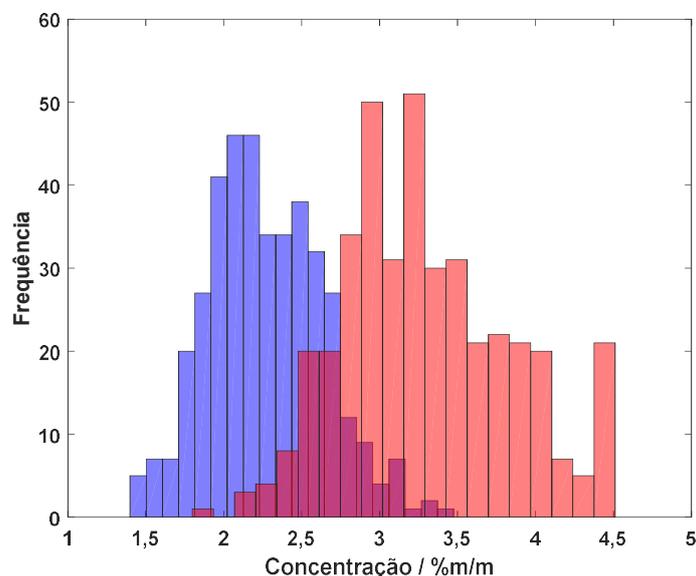


**Figura 58.** Perfis de concentração para o microselo impregnado por (a) imersão e (b) sem impregnação a partir da matriz aumentada. C1 é referente ao 25E-NBOME e C2 referente ao papel.



**Figura 59.** Perfis de concentração para os microselos impregnados por (a) imersão e (b) sem impregnação a partir da matriz aumentada com conversão da escala de cores para concentração em %m/m de 25E-NBOME no selo.

Ao se avaliar os histogramas obtidos a partir da matriz aumentada imersão/branco (figura 60), também foi possível observar um claro perfil distinto entre eles. Pelo teste K-S foram confirmadas a diferença entre as distribuições, sendo obtidos os seguintes resultados  $H = 1$  e  $p = 1,17 \times 10^{-106}$ .



**Figura 60.** Histogramas sobrepostos dos perfis de concentração obtidos a partir da deconvolução de dados por MCR da matriz aumentada dos microselos impregnados por imersão (em vermelho) e sem impregnação (em azul).

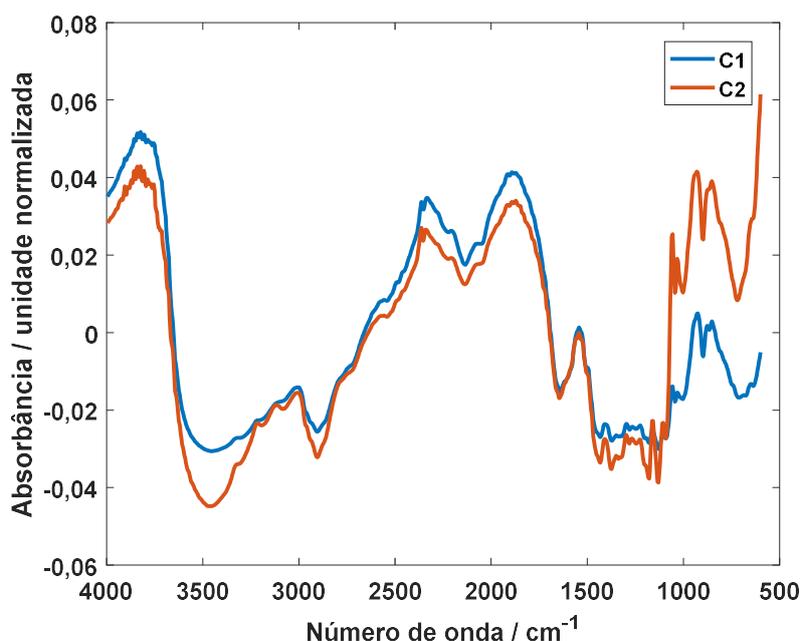
Os resultados apresentados nas figuras 53 a 60 indicam que os modelos foram capazes de diferenciar amostras que contém ou não 25E-NBOMe, sendo que nas amostras de branco os níveis de concentração foram muito menores que os observados nas amostras impregnadas com a droga. Estes resultados indicam um caminho promissor para avaliação da distribuição de substâncias ilícitas em selos, bem como a aplicação de mapeamento e imagem hiperespectral para outros tipos de materiais que são usados como suportes para substâncias ilícitas.

#### 4.3.3. Avaliação dos modelos para amostra apreendida

Quatro amostras apreendidas foram avaliadas por meio de imagens hiperespectrais e MCR, a fim de se verificar se é possível observar um padrão de distribuição de substâncias em amostras reais. Das quatro amostras analisadas, três continham 25E-NBOH e uma continha 25I-NBOH (conteúdos confirmados por LC-QTOF-MS), porém apenas uma (contendo 25E-NBOH) apresentou resultados satisfatórios após a deconvolução da matriz de dados. Como no caso do microselo impregnado por aspensão, três dessas amostras não forneceram espectros puros

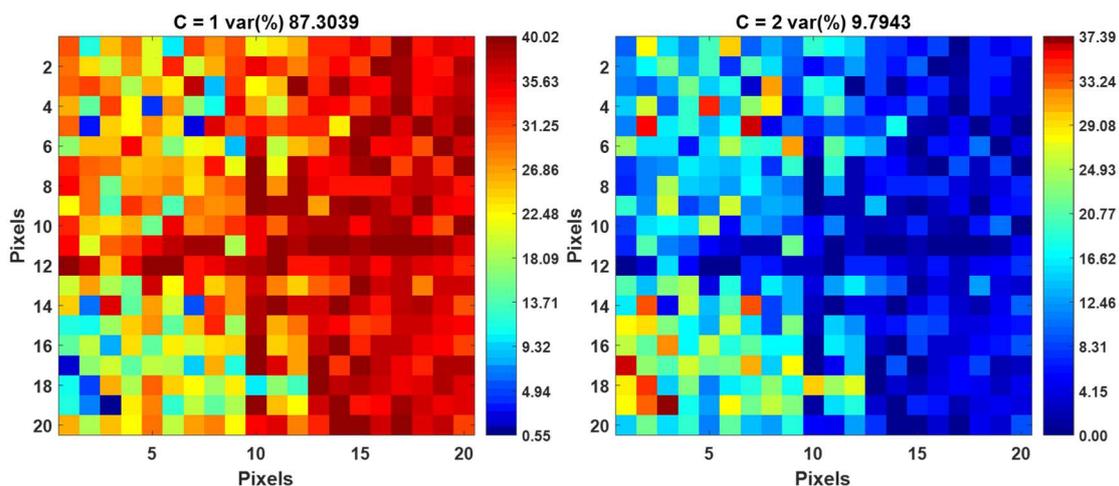
diferentes o suficiente (com a banda dupla relativa à amina) para que algum dos componentes pudesse ser identificado com segurança como contendo uma das feniletilaminas analisadas.

A única amostra que apresentou resultado aceitável forneceu uma diferença bem discreta entre os espectros puros recuperados, na região espectral do estiramento de amina primária, como pode ser observado na figura 61.



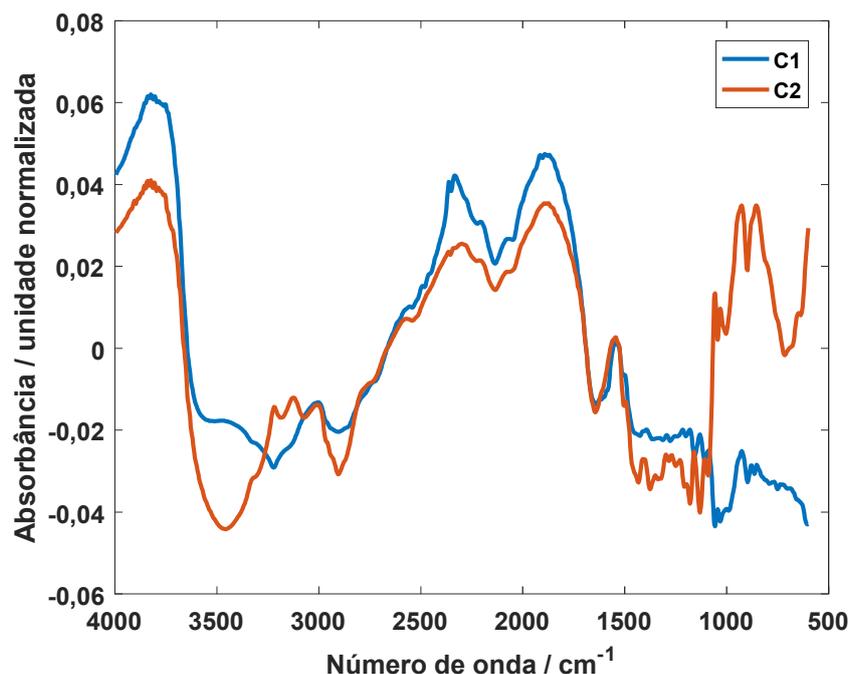
**Figura 61.** Espectros puros recuperados para a amostra apreendida contendo 25E-NBOH. C1 é referente ao papel e C2 é referente ao 25E-NBOH.

É possível observar que a intensidade do pico duplo obtidos no espectro puro recuperado para o 25E-NBOH é baixa, mas mesmo assim gera uma diferença visual significativa, quando comparada à estimativa do espectro puro do papel. A área mapeada da amostra apreendida foi uma região de cruzamento de quatro microsselos com tamanho semelhante à área mapeada de apenas um microsselo. Ao observar os perfis de intensidade relativa obtidos para a amostra apreendida, é possível perceber que a distribuição da substância na área mapeada não é homogênea, havendo um acúmulo de pixels de maior intensidade do lado esquerdo da região mapeada (figura 62, C2).



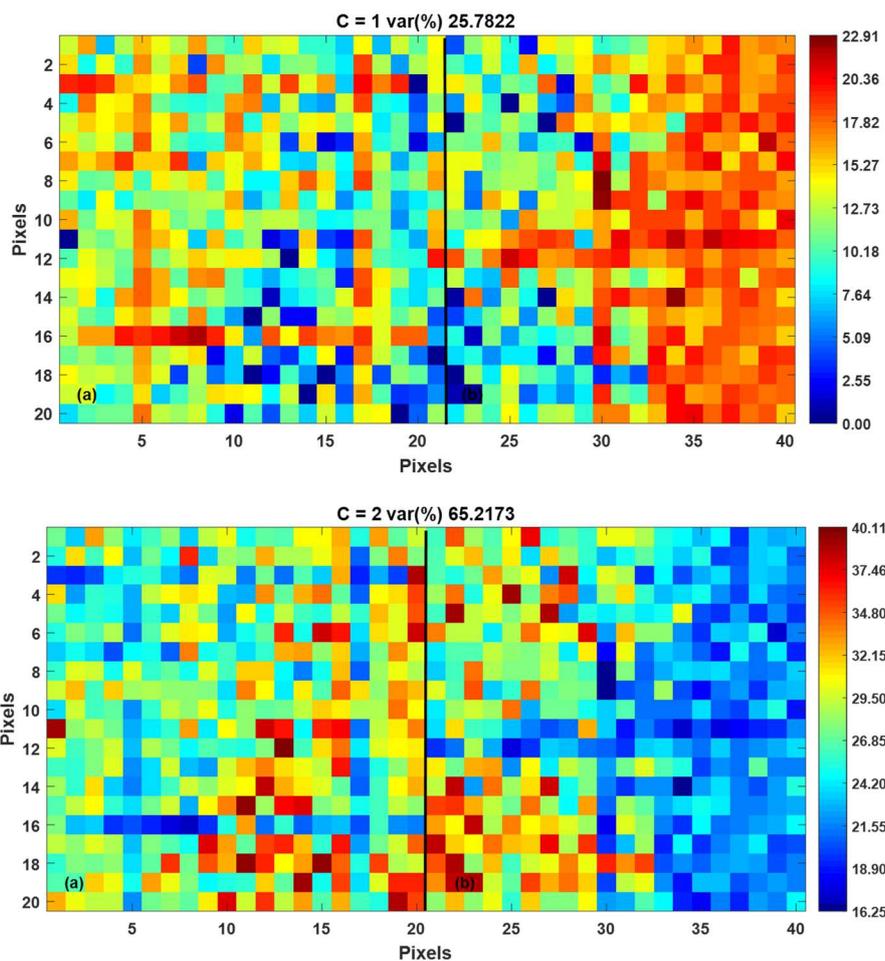
**Figura 62.** Perfis de concentração para a amostra apreendida. C1 é referente ao papel e C2 referente ao 25E-NBOH.

Como realizado com as matrizes de microsselos impregnados por gotejamento e imersão, foi avaliado para a amostra apreendida se, realizando-se o aumento da matriz, os resultados obtidos para os perfis de concentração e espectros puros recuperados desta amostra apresentariam resultados melhores em comparação com a deconvolução individual. Deste modo, a matriz de dados para a amostra apreendida foi deconvolvida juntamente com a matriz de dados do microsselo impregnado por gotejamento. Isto pôde ser realizado, apesar das substâncias impregnadas serem diferentes, porque a banda que diferencia os espectros puros é a de estiramento da amina primária, presente em ambas as substâncias. Os espectros puros recuperados são apresentados na figura 63. É possível perceber que, com o aumento da matriz e deconvolução conjunta com a matriz de gotejamento, a diferença dos espectros é bem melhor notada na região de 3000 a 3500  $\text{cm}^{-1}$ , comprovando, mais uma vez, a eficiência da aplicação de matriz aumentada. Além disso, o espectro obtido para a componente 2 na figura 63 (25E-NBOMe) é concordante em praticamente todas as regiões, quando comparado com os espectros puros recuperados para essa substância na matriz aumentada por gotejamento/imersão.

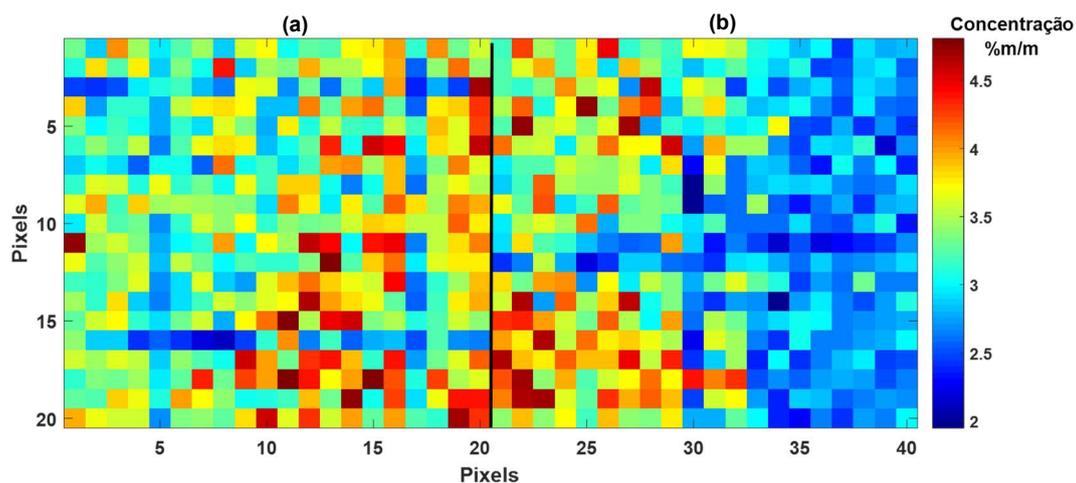


**Figura 63.** Espectros puros recuperados para a matriz aumentada amostra apreendida contendo 25E-NBOH e microselo impregnado com 25E-NBOMe por gotejamento. C1 é referente ao papel e C2 é referente ao 25E-NBOH.

Os perfis de concentração para ambas as componentes também indicam a eficiência no uso da matriz aumentada. É possível perceber um aumento de destaque nos *pixels* nos quais a substância impregnada está presente no selo apreendido (figura 64b, C2), quando comparado ao mapa de concentração da matriz simples para a amostra apreendida. Os valores da escala do mapa de concentração também foram convertidos para concentração em %m/m, com base na amostra simulada por gotejamento. O perfil de concentração da C2, com conversão da escala para concentração em % m/m é apresentado na figura 65. Nesta imagem é possível observar um acúmulo maior de 25E-NBOH no lado esquerdo da área mapeada, de modo que, do lado direito, a concentração da substância é mais baixa, próxima de 2,0% m/m, conforme a conversão da escala de cores.



**Figura 64.** Perfis de concentração para o microsselo impregnado por (a) gotejamento e (b) para a amostra apreendida da matriz aumentada. C1 é referente ao papel e C2 referente ao 25E-NBOMe/25E-NBOH.



**Figura 65.** Perfil de concentração do 25E-NBOMe para o microsselo impregnado por (a) gotejamento com conversão da escala de cores para concentração em %m/m e (b) perfil de concentração de 25E-NBOH no microsselo apreendido com conversão da escala de cores para concentração em %m/m.

Os resultados para a amostra apreendida indicam que pode haver, realmente, uma heterogeneidade na dispersão de substâncias nos selos e o mapeamento utilizando uma técnica de infravermelho associada à Quimiometria pode ser interessante para avaliação do perfil de distribuição de substâncias impregnadas em selos a depender da forma de impregnação. Estes resultados se mostraram promissores, mas é necessária aplicação em mais amostras simuladas. Além disso, as áreas mapeadas foram muito restritas, de forma que resultados mais representativos dos diferentes tipos de impregnação de drogas sintéticas em selos poderiam ser obtidos com o mapeamento de áreas maiores e com maior resolução espacial. Contudo, para que o tempo de análise seja factível de uso na prática é necessário um equipamento com maior velocidade de aquisição.

## 5. CONCLUSÕES E PERSPECTIVAS

A técnicas de análise aplicadas neste trabalho se basearam no uso da espectroscopia no infravermelho (NIR ou MIR) para caracterização e classificação de amostras de selos contendo drogas sintéticas impregnadas, normalmente as chamadas NSP, sugerem que os modelos de classificação, discriminação e caracterização dessas amostras podem ser aplicados como métodos alternativos para a triagem de amostras em laboratórios forenses.

A utilização de um espectrômetro NIR portátil, associado a modelos PLS-DA, se mostrou uma alternativa viável para a aplicação de um método para triagem de amostras de selos apreendidas, que pode até mesmo ser aplicado *in loco*. Os resultados obtidos indicaram a eficiência do PLS-DA em discriminar selos que contêm dos que não contêm drogas (modelo A), bem como discriminar selos contendo NBOMe dos contendo NBOH (modelo B). Os modelos desenvolvidos por PLS-DA apresentaram resultados coerentes com seu princípio de discriminação, baseado em uma classificação binária (0 ou 1). Apesar dos resultados de robustez terem apontado que o modelo B não pode ser considerado robusto com o número de amostras atual, esse problema tende a ser contornado com a incorporação de mais amostras no banco de dados, cujo número dependerá da variação observada nas apreensões dessas drogas.

Apesar dos resultados de eficiência obtidos com o SIMCA para o modelo B terem sido inferiores, quando comparados com ao PLS-DA, a avaliação da capacidade de identificar que uma amostra pertence a uma classe de droga não incluída na fase de treinamento foi mais eficiente, o que se deve à sua capacidade modelar uma região de confiança na classe definida na fase de treinamento.

Os resultados obtidos para os mesmos modelos (A e B) desenvolvidos na região do MID indicam elevada similaridade com os modelos desenvolvidos por NIR em relação, principalmente, às taxas de eficiência. Além disso, estes modelos apresentam a vantagem de terem sido desenvolvidos no mesmo *software* do equipamento de análise, possibilitando a realização de discriminação de amostras por um analista treinado apenas com um procedimento operacional padrão. De modo geral, para modelos de classificação e discriminação, as diferentes regiões espectrais de trabalho (NIR e MIR) forneceram resultados promissores e concordantes entre si, de modo que as duas regiões espectrais podem ser consideradas aptas para o desenvolvimento de métodos de identificação das drogas analisadas.

Adicionalmente, a análise das imagens hiperespectrais por MCR-ALS indicou a possibilidade da avaliação da homogeneidade de 25E-NBOMe incorporado em selos, por meio da avaliação de amostras simuladas por duas técnicas de impregnação,

gotejamento e imersão, uma vez que os resultados obtidos para a amostra impregnada por aspersão não foram satisfatórios. Foi possível, por meio da deconvolução das imagens hiperespectrais, separar espectros, determinar a distribuição da substância no microselo e estimar sua concentração pontual por meio das intensidades relativas fornecidas pelos modelos.

A avaliação estatística das distribuições de 25E-NBOMe nos microselos (teste K-S) sugere que existe diferença das distribuições da droga no microselo, dependendo da forma de impregnação. A fim de se garantir, com maior confiança, a existência de diferença na distribuição conforme a técnica de impregnação, mais amostras devem ser simuladas e feita uma amostragem de uma área maior e com maior resolução para aumentar a representatividade da análise e a confiabilidade dos resultados.

Considerando esses aspectos, o objetivo específico proposto inicialmente para verificar se a informação da homogeneidade da distribuição seria viável para avaliar a similaridade ou correlação de apreensões com a mesma substância pode ser considerado apenas parcialmente cumprido. Com os resultados obtidos, não foi possível correlacionar apreensões pela forma de impregnação. Portanto, os resultados obtidos para as imagens hiperespectrais de selos são promissores, mas carecem de mais análises com novas amostras.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

1. Kyriakou, C.; Marinelli, E.; Frati, P.; Santurro, A.; Afxentiou, M.; Zaami, S.; Busardo, F. P.; *Eur. Rev. Med. Pharmacol. Sci.* **2015**, *19*, 3270.
2. Lobo Vicente, J.; Chassaingne, H.; Holland, M. V.; Reniero, F.; Kolář, K.; Tirendi, S.; Vandecasteele, I.; Vinckier, I.; Guillou, C.; *Forensic Sci. Int.* **2016**, *265*, 107.
3. ANVISA, Portaria N.º 344/98, **1998**.
4. Leth-Petersen, S.; Gabel-Jensen, C.; Gillings, N.; Lehel, S.; Hansen, H. D.; Knudsen, G. M.; Kristensen, J. L.; *Chem. Res. Toxicol.* **2016**, *29*, 96.
5. Coelho Neto, J.; Andrade, A. F. B.; Lordeiro, R. A.; Machado, Y.; Elie, M.; Ferrari Júnior, E.; Arantes, L. C.; *Forensic Toxicol.* **2017**, *35*, 415.
6. Arantes, L. C.; Júnior, E. F.; de Souza, L. F.; Cardoso, A. C.; Alcântara, T. L. F.; Lião, L. M.; Machado, Y.; Lordeiro, R. A.; Neto, J. C.; Andrade, A. F. B.; *Forensic Toxicol.* **2017**, *35*, 408.
7. Bell, S.; *Annu. Rev. Anal. Chem.* **2009**, *2*, 297.
8. Romão, W.; Schwab, N. V.; Bueno, M. I. M. S.; Sparrapan, R.; Eberlin, M. N.; Martiny, A.; Sabino, B. D.; Maldaner, A. O.; *Quim. Nova* **2011**, *34*, 1717.
9. Wu, A. H. B.; French, D.; *Clin. Chim. Acta* **2013**, *420*, 4.
10. Iqbal, M. A.; Nizio, K. D.; Ueland, M.; Forbes, S. L.; *TrAC - Trends Anal. Chem.* **2017**, *91*, 112.
11. Martins, A. R.; Talhavini, M.; Vieira, M. L.; Zacca, J. J.; Braga, J. W. B.; *Food Chem.* **2017**, *229*, 142.
12. e Brito, L. R.; Martins, A. R.; Braz, A.; Chaves, A. B.; Braga, J. W.; Pimentel, M. F.; *TrAC - Trends Anal. Chem.* **2017**, *94*, 54.
13. Thanasoulas, N. C.; Piliouris, E. T.; Kotti, M. S. E.; Evmiridis, N. P.; *Forensic Sci. Int.* **2002**, *130*, 73.
14. Materazzi, S.; Risoluti, R.; Pinci, S.; Saverio Romolo, F.; *Talanta* **2017**, *174*, 673.
15. Egan, W. J.; Er, W. I. E. B.; Organ, S. L. M.; *Appl. Spectrosc.* **1999**, *53*, 219.
16. Schmitt-Kopplin, P.; Englmann, M.; *Electrophoresis* **2005**, *26*, 1209.

17. Szafarska, M.; Wietecha-Postuszny, R.; Woźniakiewicz, M.; Kościelniak, P.; *Forensic Sci. Int.* **2011**, *212*, 78.
18. Álvarez, Á.; Yáñez, J.; Contreras, D.; Saavedra, R.; Sáez, P.; Amarasiriwardena, D.; *Forensic Sci. Int.* **2017**, *280*, 169.
19. Risoluti, R.; Materazzi, S.; Gregori, A.; Ripani, L.; *Talanta* **2016**, *153*, 407.
20. Correa, D. N.; Santos, J. M.; Eberlin, L. S.; Eberlin, M. N.; Teunissen, S. F.; *Anal. Chem.* **2016**, *88*, 2515.
21. U. S. D. J. *SCIENTIFIC WORKING GROUP FOR THE ANALYSIS OF SEIZED DRUGS RECOMMENDATIONS*; **2014**; Vol. Version 7.
22. Medeiros, P. M.; Simoneit, B. R. T.; *J. Sep. Sci.* **2007**, *30*, 1516.
23. Stefanuto, P. H.; Perrault, K. A.; Stadler, S.; Pesesse, R.; Leblanc, H. N.; Forbes, S. L.; Focant, J. F.; *Anal. Bioanal. Chem.* **2015**, *407*, 4767.
24. Skoog, D. A.; Holler, F. J.; Nieman, T. A. *Princípios de Análise Instrumental*; 5ª edição.; Porto Alegre, **2002**.
25. Shimamoto, G. G.; Terra, J.; Bueno, M. I. M. S.; *J. Braz. Chem. Soc.* **2013**, *24*, 731.
26. Melucci, D.; Monti, D.; D'Elia, M.; Luciano, G.; *J. Forensic Sci.* **2012**, *57*, 86.
27. Materazzi, S.; Gregori, A.; Ripani, L.; Apriceno, A.; Risoluti, R.; *Talanta* **2017**, *166*, 328.
28. Pereira, L. S. A.; Lisboa, F. L. C.; Neto, J. C.; Valladão, F.; Sena, M.; *Microchem. J.* **2017**, *133*, 96.
29. Favretto, D.; Pascali, J. P.; Tagliaro, F.; *J. Chromatogr. A* **2013**, *1287*, 84.
30. Meyer, M. R.; *Arch. Toxicol.* **2016**, *90*, 2421.
31. Carvalho, T. C.; Oliveira, I. F.; Tose, L. V.; Vanini, G.; Kill, J. B.; Neto, A. C.; Machado, L. F.; Ambrosio, J. C. L.; Lacerda, V.; Vaz, B. G.; Romão, W.; *Anal. Methods* **2016**, *8*, 614.
32. González-Mariño, I.; Gracia-Lor, E.; Bagnati, R.; Martins, C. P. B.; Zuccato, E.; Castiglioni, S.; *Anal. Bioanal. Chem.* **2016**, *408*, 4297.
33. Ibáñez, M.; Sancho, J. V.; Bijlsma, L.; Van Nuijs, A. L. N.; Covaci, A.; Hernández, F.; *TrAC - Trends Anal. Chem.* **2014**, *57*, 107.

34. Cawley, A.; Pasin, D.; Ganbat, N.; Ennis, L.; Smart, C.; Greer, C.; Keledjian, J.; Fu, S.; Chen, A.; *Anal. Methods* **2016**, *8*, 1789.
35. ANVISA RDC N.º 143 **2017**.
36. ANVISA RDC nº 76 **2016**.
37. Zuba, D.; Sekuła, K.; Buczek, A.; *Forensic Sci. Int.* **2013**, *227*, 7.
38. Morini, L.; Bernini, M.; Vezzoli, S.; Restori, M.; Moretti, M.; Crenna, S.; Papa, P.; Locatelli, C.; Osculati, A. M. M.; Vignali, C.; Groppi, A.; *Forensic Sci. Int.* **2017**, *279*, e1.
39. Poklis, J. L.; Devers, K. G.; Arbefeville, E. F.; Pearson, J. M.; Houston, E.; Poklis, A.; *Forensic Sci. Int.* **2014**, *234*, e14.
40. Tyrkkö E, Andersson M, K. R.; *Ther. Drug Monit.* **2015**, 190.
41. Wayhs, C. A. Y.; Reis, M. Dos; Mariotti, K. D. C.; Romão, W.; Vaz, B. G.; Ortiz, R. S.; Limberger, R. P.; *Rev. Bras. Crim.* **2016**, *5*, 22.
42. Brunton, L. L.; Chabner, B. A.; C., K. B. *Goodman & Gilman - As Bases Farmacológicas da Terapêutica*; The McGraw-Hill, Ed.; 12nd ed.; New York, **2012**.
43. Poklis, J. L.; Raso, S. A.; Alford, K. N.; Poklis, A.; Peace, M. R.; *J. Anal. Toxicol.* **2015**, *39*, 617.
44. ANVISA RDC N.º 6 **2014**.
45. McGonigal, M. K.; Wilhide, J. A.; Smith, P. B.; Elliott, N. M.; Dorman, F. L.; *Forensic Sci. Int.* **2017**, *275*, 83.
46. Pavia, D. L.; Lampman, G. M.; Kriz, G. S.; Vyvyan, J. R. *Introdução à Espectroscopia*; Cengage Learning, Ed.; Tradução da 4ª edição norte-americana; São Paulo.
47. Pasquini, C.; *J. Braz. Chem. Soc.* **2003**, *14*, 198.
48. Porfire, A.; Rus, L.; Vonica, L. A.; Tomuta, I.; *J. Pharm. Biomed. Anal.* **2012**, *70*, 301.
49. Luo, W.; Wu, J.; Wang, X.; Lin, X.; Li, H.; *Anal. Methods* **2013**, *5*, 1337.
50. Soares, L. F.; Da Silva, D. C.; Bergo, M. C. J.; Coradin, V. T. R.; Braga, J. W. B.; Pastore, T. C. M.; *Quim. Nova* **2017**, *40*, 418.

51. Workman Jr, J. J.; *Appl. Spectrosc. Rev.* **2006**, *31*, 251.
52. Luo, W.; Wu, J.; Wang, X.; Lin, X.; Li, H.; *Anal. Methods* **2013**, *5*, 1337.
53. Tibola, C. S.; Medeiros, E. P.; Simeone, M. L. F.; Oliveira, M. A. *Espectroscopia no Infravermelho Próximo para Avaliar Indicadores de Qualidade Tecnológica e Contaminantes em Grãos*; **2018**.
54. Ferreira, M. M. C. *Quimiometria - Conceitos, Métodos e Aplicações*; Editora Unicamp, 1ª Ed.; São Paulo, **2015**.
55. Chau, F. T.; Liang, Y. Z.; Gao, J.; Shao, X. G. *Chemometrics: From Basics to Wavelets Transform*; Sons, J. W. & Ed.; Hoboken, New Jersey, **2004**.
56. Wise, B. M.; Gallagher, N. B.; Bro, R.; Shaver, J. M.; Windig, W.; Koch, R. S. *PLS\_Toolbox Version 4.0 for use with MATLAB™*; **2006**.
57. Ballabio, D.; Consonni, V.; *Anal. Methods* **2013**, *5*, 3790.
58. Oliveri, P.; Downey, G.; *TrAC - Trends Anal. Chem.* **2012**, *35*, 74.
59. Brereton, R. G.; *J. Chemom.* **2011**, *25*, 225.
60. Rodionova, O. Y.; Titova, A. V.; Pomerantsev, A. L.; *TrAC - Trends Anal. Chem.* **2016**, *78*, 17.
61. Oliveri, P.; López, M. I.; Casolino, M. C.; Ruisánchez, I.; Callao, M. P.; Medini, L.; Lanteri, S.; *Anal. Chim. Acta* **2014**, *851*, 30.
62. Brereton, R. G.; Lloyd, G. R.; *J. Chemom.* **2014**, *28*, 213.
63. American Society for Test and Materials. *Standart Pratices for Infrared Multivariate Quantitative Analysis*; **2015**, 1.
64. Juan, Anna; Jaurnot, Joaquim; Tauler, R.; *Anal. Methods* **2014**, *6*, 4964.
65. Março, Paulo Henrique; Valderrama, Patrícia; Alexandrino, Guilherme Lionello; Poppi, Ronei Jesus; Tauler, R.; *Quim. Nova* **2014**, *37*, 1525.
66. Felten, J.; Hall, H.; Jaumot, J.; Tauler, R.; Juan, A. De; Gorzsás, A.; *Nat. Protoc.* **2015**, *10*, 217.
67. Amigo, M.; Babamoradi, H.; Elcoroaristizabal, S.; *Anal. Chim. Acta* **2015**, *896*, 34.
68. Wilczynski, S.; Koprowski, R.; Marmion, M.; Duda, P.; Blosska-Fajfrowska, B.; *Talanta* **2016**, *160*, 1.

69. Farkas, A.; Nagy, B.; Marosi, G.; *Period. Polytech. Chem. Eng.* **2018**, *62* (1), 1.
70. Carvalho, M. A. De; Talhavini, M.; Pimentel, M. F.; Amigo, J. M.; Pasquini, C.; Junior, S. A.; Weber, I. T.; *Anal. Methods* **2018**, *10*, 4711.
71. Martins, A. R.; Dourado, C. S.; Talhavini, M.; Braz, A.; Braga, J. W. B.; *Forensic Sci. Int.* **2019**, *296*, 91.
72. Morais, D. R.; Barbosa, I. L.; Cunha, K. F.; Tripodi, G. L.; Angolini, C. F. F.; Franco, M. F.; Aquino, E. M.; Eberlin, M. N.; Costa, J. L.; *Mass Spectrom.* **2017**, *52*, 701.
73. Friedrich, D. M.; Hulse, C. A.; von Gunten, M.; Williamson, E. P.; Pederson, C. G.; O'Brien, N. A.; *Photonic Instrumentation Engineering* **2014**, *8992*, 1.
74. Kyzar, E. J.; Nichols, C. D.; Gainetdinov, R. R.; Nichols, D. E.; Kalueff, A. V.; *Trends Pharmacol. Sci.* **2017**, *38*, 992.
75. Teófilo, R. F.; Martins, J. P. A.; Ferreira, M. M. C.; *J. Chemom.* **2009**, *23*, 32.
76. Xiaobo, Z.; Jiewen, Z.; Povey, M. J. W.; Holmes, M.; Hanpin, M.; *Anal. Chim. Acta* **2010**, *667*, 14.
77. Bruker OPUS User Manual **2006**.
78. Mobaraki, N.; Amigo, J. M. ; *Chemom. Intell. Lab. Syst.* **2018**, *172*, 174.
79. Magalhães, L. de O.; Arantes, L. C.; Braga, J. W. B.; *Microchem. J.* **2019**, *144*, 151.
80. Henderson, A. R.; *Clin. Chim. Acta* **2006**, *366*, 112.
81. Ferreira, J. C.; Patino, C. M.; *J. Bras. Pneumol.* **2015**, *41*, 3713.

## 7. ANEXOS

### ANEXO I

# PROCEDIMENTO OPERACIONAL PADRÃO PARA AQUISIÇÃO DE ESPECTROS DE FT-IR E ANÁLISE DA PRESENÇA OU AUSÊNCIA DE DROGAS (MODELO A) E DISCRIMINAÇÃO ENTRE NBOME E NBOH.

## 1. AQUISIÇÃO DOS ESPECTROS

1.1. Posicione o selo sobre o cristal de ATR com a face contendo imagem voltada para cima, feche o acessório e adquira o espectro.

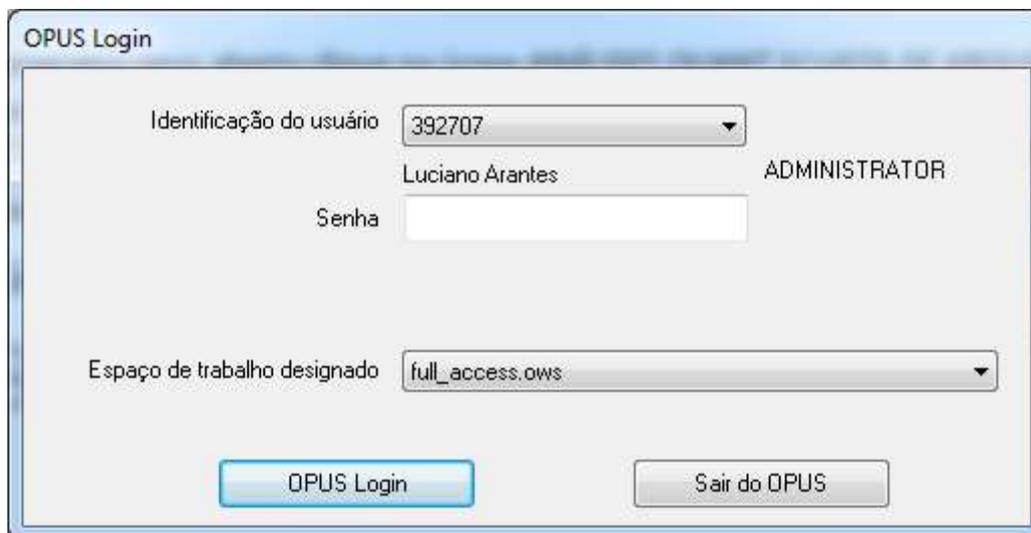
**OBS: caso a amostra apresente gravuras ou regiões com tinta mesmo no seu verso, a aquisição dos espectros deve ser realizada apenas em posições sem a presença de tinta ou gravuras.**

1.2. Converta o espectro para a escala de absorbância e salve o arquivo.

## 2. Análise da presença ou ausência de drogas em selos (MODELO A)

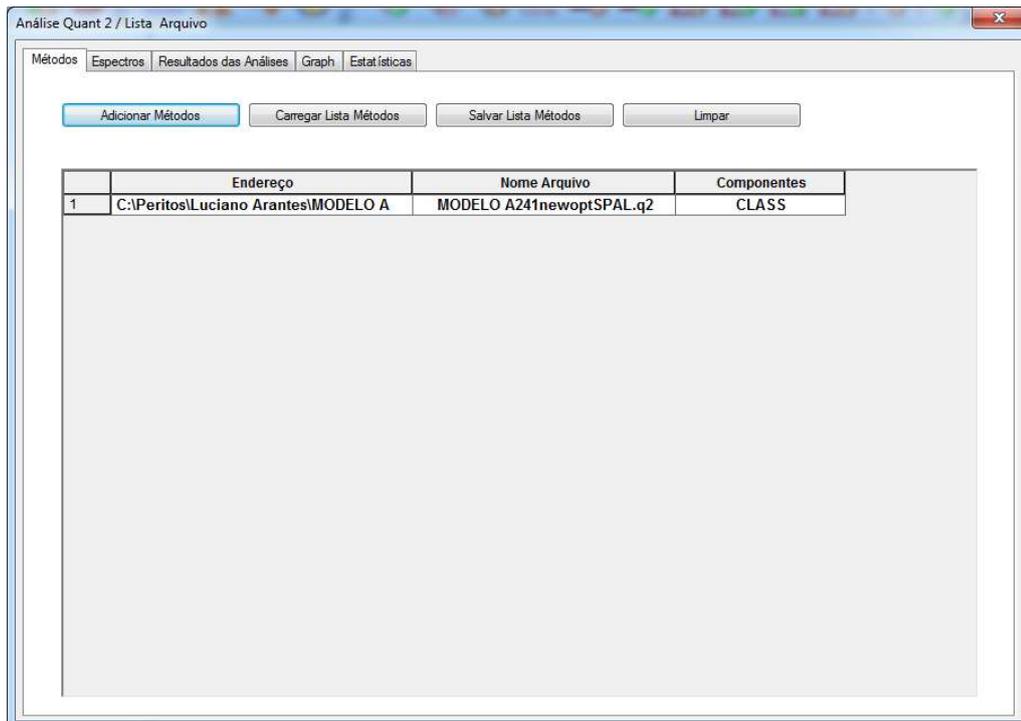
2.1. Clique no ícone OPUS 7.5 na área de trabalho.

2.2. A janela a seguir será exibida, Altere a identificação do usuário, caso necessário. Selecione 'full\_access.ows' no espaço de trabalho designado. Insira a senha e clique em OPUS login.



2.3. Com o programa opus aberto clique no ícone *ANÁLISES QUANT 2/ LISTA DE*

*ARQUIVOS* . A seguinte janela será exibida:



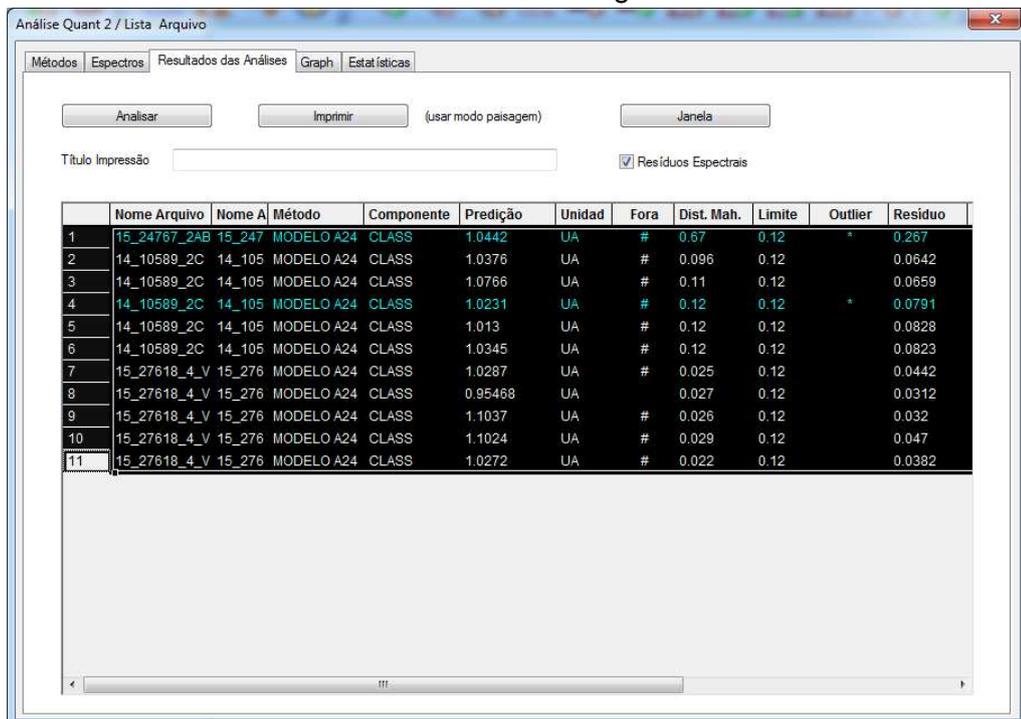
2.4. Na aba **MÉTODOS** clique em adicionar método e selecione o método desejado. Para o modelo A, navegue para o endereço *C:\Peritos\Luciano Arantes\MODELO A* E selecione o arquivo: “MODELO A241newoptSPAL.q2v”.

2.5. Siga para aba **ESPECTROS**, clique em **ADICIONAR ESPECTROS** e selecione os espectros que serão analisados.

**OBS: os espectros devem ser adquiridos conforme descrito no item 1 deste POP.**

Continue para a aba **RESULTADOS DAS ANÁLISES** e clique em **ANALISAR**.

2.6. O resultado da análise será exibido como na figura abaixo:



- 2.7. Marque o ícone *RESÍDUOS ESPECTRAIS* e depois selecione os resultados das amostras analisadas clicando no número das mesmas. Em seguida copie os resultados clicando *Ctrl+C*.
- 2.8. Abra a planilha Excel *Modelo\_A.xlsx*, localizada no caminho: *C:\Peritos\Luciano Arantes\MODELO A* e cole os resultados na aba ROTINA.
- 2.9. As colunas OUTLIERS (coluna P) e CLASSE PREDITA (coluna Q) resumem o resultado da análise.
- 2.10. A tabela abaixo mostra as quatro possibilidades de resultado que podem ser obtidas:

OUTLIER	CLASSE PREDITA	CONCLUSÃO DA ANÁLISE
SIM	BRANCO	A amostra foi identificada como OUTLIER. Logo, apresenta características diferentes das amostras utilizadas no treinamento do modelo. Portanto, não se deve confiar na classe predita.
SIM	NPS	A amostra foi identificada como OUTLIER. Logo, apresenta características diferentes das amostras utilizadas no treinamento do modelo. Portanto, não se deve confiar na classe predita.
NÃO	BRANCO	A amostra não foi identificada como OUTLIER. Logo, pode ser analisada pelo modelo. A classe predita é compatível com selos sem drogas (BRANCO).
NÃO	NPS	A amostra não foi identificada como OUTLIER. Logo, pode ser analisada pelo modelo. A classe predita é compatível com selos contendo alguma das drogas sintéticas incluídas no treinamento do modelo (DROGAS)*.

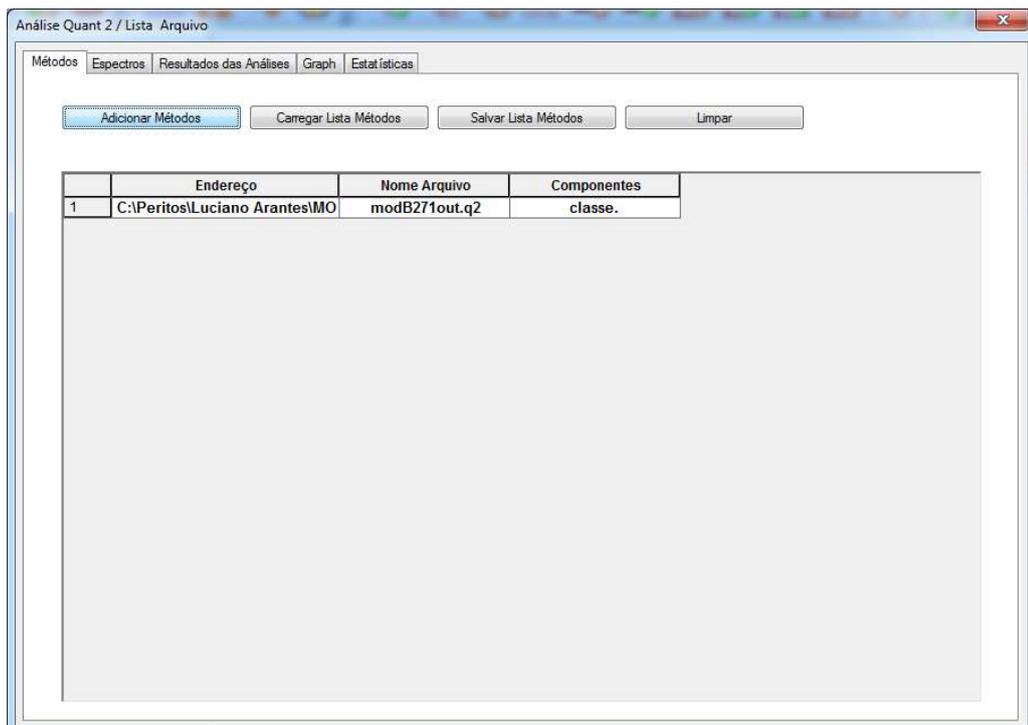
\* NBOMe, NBOH, LSD, MDMA, DOC, 2C-I

### 3. DISCRIMINAÇÃO DE SELOS CONTENDO NBOMe E NBOH (MODELO B)

- 3.1. Com o programa opus aberto clique no ícone *ANÁLISES QUANT 2/ LISTA DE*

*ARQUIVOS* 

A seguinte janela será exibida:



- 3.2. Na aba **MÉTODOS** clique em adicionar método e selecione o método desejado. Para o MODELO B, navegue para o endereço C:\Peritos\Luciano Arantes\MODELO B\MODELO B e selecione o arquivo: “modB271out.q2”. Siga para aba espectros, clique em **ADICIONAR ESPECTROS** e selecione os espectros que serão analisados.  
OBS: os espectros devem ser adquiridos conforme descrito no item 1 deste POP.
- 3.3. Continue para a aba **RESULTADOS DAS ANÁLISES** e clique em **ANALISAR**.
- 3.4. O resultado da análise será exibido como na figura abaixo:

	Nome Arquivo	Nome A	Método	Componente	Predição	Unidad	Fora	Dist. Mah.	Limite	Outlier	Residuo
1	15_24767_2AB	15_247	modB271out.	classe.	-0.12078	u.a	#	0.041	0.09		0.00185
2	14_10589_2C	14_105	modB271out.	classe.	0.78144	u.a		0.027	0.09		0.000373
3	14_10589_2C	14_105	modB271out.	classe.	0.78149	u.a		0.026	0.09		0.000357
4	14_10589_2C	14_105	modB271out.	classe.	0.89369	u.a		0.032	0.09		0.000337
5	14_10589_2C	14_105	modB271out.	classe.	0.93677	u.a		0.032	0.09		0.000345
6	14_10589_2C	14_105	modB271out.	classe.	0.78342	u.a		0.035	0.09		0.000334
7	15_27618_4_V	15_276	modB271out.	classe.	0.34362	u.a		0.0094	0.09		0.00026
8	15_27618_4_V	15_276	modB271out.	classe.	0.25106	u.a		0.0091	0.09		0.00024
9	15_27618_4_V	15_276	modB271out.	classe.	0.4315	u.a		0.0072	0.09		0.000284
10	15_27618_4_V	15_276	modB271out.	classe.	0.088181	u.a		0.012	0.09		0.000267
11	15_27618_4_V	15_276	modB271out.	classe.	0.34413	u.a		0.0093	0.09		0.000267

- 3.5. Marque o ícone *RESÍDUOS ESPECTRAIS* e depois selecione os resultados das amostras analisadas clicando no número das mesmas. Em seguida copie os resultados clicando *Ctrl+C*.
- 3.6. Abra a planilha Excel *Modelo\_B.xlsx*, localizada no caminho: C:\Peritos\Luciano Arantes\MODELO B e cole os resultados na aba ROTINA.
- 3.7. As colunas OUTLIERS (coluna P) e CLASSE PREDITA (coluna Q) resumem o resultado da análise.
- 3.8. A tabela abaixo mostra as quatro possibilidades de resultado que podem ser obtidas:

<i>OUTLIER</i>	<i>CLASSE PREDITA</i>	<i>CONCLUSÃO DA ANÁLISE</i>
<i>SIM</i>	<i>NBOMe</i>	<i>A amostra foi identificada como OUTLIER. Logo, apresenta características diferentes das amostras utilizadas no treinamento do modelo. Portanto, não se deve confiar na classe predita.</i>
<i>SIM</i>	<i>NBOH</i>	<i>A amostra foi identificada como OUTLIER. Logo, apresenta características diferentes das amostras utilizadas no treinamento do modelo. Portanto, não se deve confiar na classe predita.</i>
<i>NÃO</i>	<i>NBOMe</i>	<i>A amostra não foi identificada como OUTLIER. Logo, pode ser analisada pelo modelo. A classe predita é compatível com selos contendo NBOMe.</i>
<i>NÃO</i>	<i>NBOH</i>	<i>A amostra não foi identificada como OUTLIER. Logo, pode ser analisada pelo modelo. A classe predita é compatível com selos contendo NBOH.</i>