

LAURENCE SOUZA VIEIRA

**VERIFICAÇÃO DA AUTENTICIDADE DE AZEITE DE OLIVA POR APLICAÇÃO
DE MÉTODOS ESPECTROSCÓPICOS, CROMATOGRÁFICOS E
QUIMIOMÉTRICOS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Agroquímica, para obtenção do título de *Doctor Scientiae*.

Orientadora: Maria Eliana L. R. de Queiroz

Coorientadores: Antônio Augusto Neves
André Fernando de Oliveira

**VIÇOSA - MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

V658v
2021

Vieira, Laurence Souza, 1991-
Verificação da autenticidade de azeite de oliva por
aplicação de métodos espectroscópicos, cromatográficos e
quimiométricos / Laurence Souza Vieira. – Viçosa, MG, 2021.
134 f. : il. (algumas color.) ; 29 cm.

Inclui anexos.

Orientador: Maria Eliana Lopes Ribeiro de Queiroz.

Tese (doutorado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 119-128.

1. Azeite. 2. Alimentos - Adulteração e inspeção.
3. Espectroscopia de infravermelho. 4. Cromatografia a gás.
5. Quimiometria. I. Universidade Federal de Viçosa.
Departamento de Química. Programa de Pós-Graduação em
Agroquímica. II. Título.

CDD 22. ed. 641.3463

Bibliotecário(a) responsável: Alice Regina Pinto Pires CRB6 2523


LAURENCE SOUZA VIEIRA

VERIFICAÇÃO DA AUTENTICIDADE DE AZEITE DE OLIVA POR APLICAÇÃO
DE MÉTODOS ESPECTROSCÓPICOS, CROMATOGRÁFICOS E
QUIMIOMÉTRICOS

Tese apresentada à Universidade Federal de
Viçosa, como parte das exigências do
Programa de Pós-Graduação em
Agroquímica, para obtenção do título de
Doctor Scientiae.

APROVADA: 14 de maio de 2021.

Assentimento:


Laurence Souza Vieira
Autora


Maria Eliana Lopes Ribeiro de Queiroz
Orientadora

*A Deus e à Nossa Senhora.
Aos meus pais Galvâni e Socorro.
Aos meus irmãos Giovane e Lílian.*

Com amor, DEDICO.

AGRADECIMENTOS INSTITUCIONAIS

À Universidade Federal de Viçosa e ao Departamento de Química, pela oportunidade de realização do curso.

Ao Laboratório de Identidade e Qualidade de Alimentos (IQA)/Laboratório Federal de Defesa Agropecuária (LFDA) - MG, pela oportunidade de realização do trabalho em suas dependências.

À professora Maria Eliana Lopes Ribeiro de Queiroz, por sua substancial participação no meu processo de aprendizagem, por todo incentivo e, principalmente, por sua amizade.

Aos professores Antônio Augusto Neves e André Fernando de Oliveira, pelos aconselhamentos, ensinamentos e pela dedicação para que o trabalho pudesse ser realizado.

Aos professores Jane Selia dos Reis Coimbra, Fabio Junior Moreira Novaes e Anizio Marcio de Faria, pelos conselhos, pelas brilhantes observações e por ter aceitado o convite para participar da Banca Examinadora.

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

AGRADECIMENTOS PESSOAIS

A Deus e à Nossa Senhora, por serem minha força e meus guias em todos os passos dados, pela infinita bondade e pela proteção.

Aos meus pais, Galvâni e Socorro, por não medirem esforços para me verem feliz e realizada, pelo amor incondicional e por nunca permitirem que eu desistisse; aos meus irmãos, Giovane e Lílian, pelo carinho e pelo companheirismo.

À Camila Assis, por sua fundamental e incondicional ajuda neste trabalho.

Aos amigos do LAQUA, por toda a força, apesar da distância.

E a todos aqueles que, direta ou indiretamente, estiveram ao meu lado dando-me todo o apoio durante a caminhada.

Muito obrigada!

“A mente que se abre a uma nova ideia jamais volta ao seu tamanho original” (Albert Einstein).

BIOGRAFIA

LAURENCE SOUZA VIEIRA, filha de Galvâni Martins Vieira e Maria Socorro de Fátima Souza, nasceu na cidade de Araçuaí, Minas Gerais, em 30 de janeiro de 1991.

Em março de 2009, iniciou o curso de Química, pela Universidade Federal de Viçosa (UFV), em Viçosa-MG, diplomando-se como Bacharel e Licenciada em 08 de agosto de 2014.

Em agosto de 2014, ingressou no Programa de Pós-Graduação em Agroquímica, da Universidade Federal de Viçosa (UFV), em Viçosa-MG, em nível de Mestrado, obtendo o título em 14 de outubro de 2016.

Ingressou no Programa de Pós-Graduação em Agroquímica, da Universidade Federal de Viçosa (UFV), em Viçosa-MG, em agosto de 2016, em nível de Doutorado, submetendo-se à defesa de tese em 14 de maio de 2021.

RESUMO

VIEIRA, Laurence Souza, D.Sc., Universidade Federal de Viçosa, maio de 2021. **Verificação da autenticidade de azeite de oliva por aplicação de métodos espectroscópicos, cromatográficos e quimiométricos.** Orientadora: Maria Eliana Lopes Ribeiro de Queiroz. Coorientadores: André Fernando de Oliveira e Antônio Augusto Neves.

O azeite de oliva é a fração lipídica obtida das azeitonas, frutos da oliveira (*Olea europaea* L.), por meio de processos mecânicos ou outros processos físicos. Por ser um produto com alto valor de mercado, o azeite sofre processos de adulteração. Como os métodos oficiais para identificação de adulterações em azeite de oliva são limitados e demandam tempo e grande quantidade de insumos, fazem-se necessários estudos que tornem a verificação da autenticidade de azeite de oliva mais eficiente. O objetivo deste trabalho foi verificar a autenticidade de azeite de oliva por aplicação de métodos espectroscópicos, cromatográficos e quimiométricos, fusão de dados e seleção de variáveis. Para isso, foram utilizadas amostras de azeites de oliva extra virgem genuínos, adulteradas com sete diferentes adulterantes (óleo de soja, óleo de milho, óleo de canola, óleo de girassol e de azeite de oliva virgem, lampante e tipo único), em diferentes concentrações (0 a 20% m/m). Essas amostras foram analisadas por Espectroscopia no Infravermelho Próximo com Transformada de Fourier (FT-NIR) e por Cromatografia Gasosa acoplada à Espectrometria de Massas sequencial (GC-MS/MS) e os dados foram usados para construção de modelos utilizando regressão por Mínimos Quadrados Parciais (PLS) e Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA) com utilização dos métodos Seleção de Preditores Ordenados (OPS) e Algoritmo Genérico (GA), permitindo a classificação e calibração multivariada de amostras de azeite de oliva adulteradas com altos valores de especificidade e sensibilidade, baixos valores de Raiz Quadrada do Erro Médio Quadrático de Previsão (RMSEP), Raiz Quadrada do Erro Médio Quadrático de Calibração (RMSEC) e Raiz Quadrada do Erro Médio Quadrático de Validação Cruzada (RMSECV) e boa linearidade, caracterizada por coeficientes de correlação maiores que 0,95. Um método utilizando Extração em Fase Sólida (SPE) e análise por GC-MS/MS para determinação de esqualeno em azeite de oliva foi validado, com bom desempenho nos parâmetros de validação e

com resultados encontrados para amostras comerciais de azeite de oliva, que corroboraram com as classificações realizadas pelo órgão de fiscalização. Ainda, a fusão de dados da espectroscopia e cromatografia foi utilizada para construção de modelos PLS-OPS para calibração de amostras de azeite de oliva extra virgem adulteradas, com baixos valores de RMSEC, RMSEP e RMSECV. Todos os valores dos coeficientes de correlação foram superiores a 0,99, demonstrando bom desempenho.

Palavras-chave: Azeite de oliva. Adulteração. FT-NIR. GC-MS/MS. Quimiometria.

ABSTRACT

VIEIRA, Laurence Souza, D.Sc., Universidade Federal de Viçosa, May, 2021. **Verification of the authenticity of olive oil by application of spectroscopic, chromatographic and chemiometric methods.** Adviser: Maria Eliana Lopes Ribeiro de Queiroz. Co-advisers: André Fernando de Oliveira and Antônio Augusto Neves.

Olive oil is the lipid fraction obtained from olives, fruits of the olive tree (*Olea europaea L.*), through mechanical or other physical processes. As a product with high market value, olive oil undergoes adulteration processes. As official methods for identification of adulteration in olive oil are invited and demand time and a large amount of inputs, studies are necessary to make the verification of the authenticity of olive oil more efficient. The objective of this work was to verify the authenticity of olive oil for application of spectroscopic, chromatographic and chemometric methods, data fusion and selection of variables. Therefore, genuine extra virgin olive oils adulterated with seven different adulterants (soybean oil, corn oil, canola oil, sunflower oil and virgin olive oil, lampante olive oil and olive oil) was used under in different concentrations (0 - 20% w/w). These samples were analysed by Fourier Transform Near-infrared Spectroscopy (FT-NIR) and Gas Chromatography coupled to tandem Mass Spectroscopy (GC-MS/MS) and the data was used to built models using Partial Least Squares (PLS) regression and Partial Least Squares Discriminant Analysis (PLS-DA) using Ordered Predictors Selection (OPS) and Genetic Algorithm (GA) method, allowing the classification and multivariate calibration of adulterated olive oil with high values of specificity and sensitivity, low values of Root Mean Squares Error of Prediction (RMSEP), Root Mean Squares Error of Calibration (RMSEC) and Root Mean Squares Error of Cross Validation (RMSECV), and good linearity, characterized by correlation coefficients greater than 0.95. A method using solid phase extraction and gas chromatography coupled to tandem mass spectroscopy to determine squalene in olive oil was validated, with acceptable performance in the validation parameters and results found for commercials olive oils which corroborated with the classifications used by the fiscalization agency. Furthermore, a data fusion of spectroscopy and chromatography was used to build

PLS-OPS models for the calibration of adulterated extra virgin olive oils, with low values for RMSEC, RMSEP and RMSECV, and all correlation coefficient values above 0.99, demonstrating good performance.

Keywords: Olive oil. Adulteration. FT-NIR. GC-MS/MS. Chemometry.

LISTA DE FIGURAS

Figura 1 - Processo geral de obtenção do azeite de oliva.....	26
Figura 2 - Estrutura química do esqualeno.....	29
Figura 3 - Representação gráfica do modelo PLS-DA.....	36
Figura 4 - Tabela representativa da atribuição dos valores binários para a matriz Y no modelo PLS-DA.....	37
Figura 5 - Discriminação das amostras de duas classes de acordo com as classes definidas pela função discriminante determinada pelo método PLS-DA.	37
Figura 6 - Representação gráfica do método OPS. (A): Construção do vetor informativo indicando as variáveis mais preditivas; (B): Diferenciação das variáveis segundo os valores absolutos dos elementos do vetor informativo; (C): Organização das variáveis diferenciadas em ordem decrescente de valor absoluto; (D): Definição dos subconjuntos formados pela janela de variáveis e seus incrementos; (E): Avaliações dos modelos para selecionar aquele que produz menor valor de RMSECV.....	39
Figura 7 - Esquema representativo do processo de fusão de dados: A: Fusão de dados de nível baixo; B: Fusão de dados de nível médio; C: Fusão de dados de nível alto.	42
Figura 8 - Espectros obtidos por FT-NIR para amostras de azeite de oliva extra virgem varietais e <i>blend</i> puros, para os adulterantes lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja puros e para todas as amostras adulteradas.....	53
Figura 9 - <i>Score plot</i> dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Koroneiki puro e para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após a aplicação do método PCA.....	55
Figura 10 - <i>Score plot</i> dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Koroneiki puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.....	56
Figura 11 - <i>Score plot</i> dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Arbequina e para os adulterantes lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após aplicação do método PCA.....	57
Figura 12 - <i>Score plot</i> dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Arbequina puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.....	58

Figura 13 - <i>Score plot</i> dos dados de FT-NIR para a amostra de azeite de oliva extra virgem do <i>blend</i> dos azeites puro e para os adulterantes lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após aplicação do do método PCA.	60
Figura 14 - <i>Score plot</i> dos dados de FT-NIR para a amostra de azeite de oliva extra virgem do <i>blend</i> puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.	61
Figura 15 - Variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para classificação das variedades Arbequina, Koroneiki e para o <i>blend</i> , utilizando os dados de FT-NIR.	64
Figura 16 – Discriminação das amostras de cada classe (Arbequina, Koroneiki e <i>blend</i>) pelos valores preditos obtidos por meio da atribuição dos valores binários do método PLS-DA-OPS para cada amostra, utilizando os dados de FT-NIR. Em vermelho, variedade Arbequina (Classe 1); em verde, variedade Koroneiki (Classe 2); em azul, <i>Blend</i> (Classe 3).	65
Figura 17 -.Variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para a variedade Koroneiki com os sete adulterantes.	67
Figura 18 - Variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para classificação das variedades Arbequina, utilizando os dados de FT-NIR.	69
Figura 19 - Variáveis selecionadas para a construção do modelo pelo método PLS-DA-GA para classificação do <i>blend</i> , utilizando os dados de FT-NIR.....	71
Figura 20 - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante azeite lampante. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.....	73
Figura 21 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite lampante. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.....	74
Figura 22 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite virgem. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.....	76
Figura 23 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite tipo único. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.	78
Figura 24 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de canola. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.....	80
Figura 25 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de girassol. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.....	82

Figura 26 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de milho. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.....	84
Figura 27 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de soja. A: Arbequina; B: Koroneiki; C: <i>Blend</i> ; D: Robusto.....	86
Figura 28 - Espectro de massas do esqualeno obtido pelo modo de aquisição <i>Scan</i>	96
Figura 29 - Cromatograma para a amostra do <i>blend</i> dos azeites de oliva contendo esqualeno, obtido pelo modo de aquisição MRM, com as transições 367,2 > 299,2 e 367,2 > 149,1, com tempo de retenção de 7,2 minutos.	97
Figura 30 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite lampante.	108
Figura 31 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite virgem.	109
Figura 32 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite tipo único.....	109
Figura 33 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de canola.	110
Figura 34 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de girassol.	111
Figura 35 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de milho.	111
Figura 36 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de soja.	112
Figura 37 - Valores preditos <i>versus</i> valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o modelo robusto.	113

LISTA DE TABELAS

Tabela 1 - Exemplo de tabela de contingência para avaliação do desempenho de métodos qualitativos.....	44
Tabela 2 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA, completo e pelo método OPS, para classificação das variedades Arbequina, Koroniewski e para o blend, utilizando os dados de FT-NIR.	63
Tabela 3 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas da variedade Koroniewski, com sete adulterantes, utilizando os dados de FT-NIR.	66
Tabela 4 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas da variedade Arbequina, com sete adulterantes, utilizando os dados de FT-NIR.....	68
Tabela 5 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas do blend, com sete adulterantes, utilizando os dados de FT-NIR.	70
Tabela 6 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroniewski, <i>blend</i> e robusto para o adulterante azeite lampante.	72
Tabela 7 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroniewski, <i>blend</i> e robusto para o adulterante azeite virgem.	75
Tabela 8 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroniewski, <i>blend</i> e robusto para o adulterante azeite tipo único.....	77
Tabela 9 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroniewski, <i>blend</i> e robusto para o adulterante óleo de canola.	79
Tabela 10 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroniewski, <i>blend</i> e robusto para o adulterante óleo de girassol.	81
Tabela 11 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroniewski, <i>blend</i> e robusto para o adulterante óleo de milho.	83
Tabela 12 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroniewski, <i>blend</i> e robusto para o adulterante óleo de soja.	85
Tabela 13 - Teores médios de esqualeno das amostras do <i>blend</i> dos azeites de oliva extra virgem adulteradas em cada porcentagem de adulteração, para os sete adulterantes.....	99
Tabela 14 - Teores de esqualeno de amostras de azeites de oliva coletadas para fiscalização realizada pelo IQA/LFDA-MG.	101

Tabela 15 - Classificação de amostras de azeites de oliva coletadas para fiscalização realizada pelo IQA/LFDA-MG, após análises para verificação de autenticidade de azeite de oliva.	102
Tabela 16 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para os adulterantes lampante, virgem, tipo único e óleo de canola.....	113
Tabela 17 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para os adulterantes óleo de girassol, óleo de milho, óleo de soja e para o modelo robusto.	114

LISTA DE ABREVIATURAS E SIGLAS

Con	Confiabilidade
CVC	Calibração, Validação, Calibração
Esp	Especificidade
EU	<i>European Union</i> (União Europeia)
Exa	Exatidão
FT-IR	<i>Fourier Transform Infrared Spectroscopy</i> (Espectroscopia no Infravermelho com Transformada de Fourier)
FT-NIR	<i>Fourier Transform Near-infrared Spectroscopy</i> (Espectroscopia no Infravermelho Próximo com Transformada de Fourier)
GA	<i>Genetic Algorithm</i> (Algoritmo Genético)
GC-MS/MS	<i>Gas Chromatography coupled to tandem Mass Spectrometry</i> (Cromatografia Gasosa acoplada à Espectrometria de Massas sequencial)
IOC	<i>International Olive Council</i> (Conselho Oleícola Internacional)
IQA	Laboratório de Identidade e Qualidade de Alimentos
LD	Limite de Detecção
LFDA-MG	Laboratório Federal de Defesa Agropecuária – Minas Gerais
LQ	Limite de Quantificação
<i>m/z</i>	Razão massa/carga
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
MMI	<i>Multimode Inlet</i>
MRM	<i>Multiple Reaction Monitoring</i> (Monitoramento de Múltiplas Reações)
MSC	<i>Multiplicative Signal Correction</i> (Correção de Espalhamento Multiplicativo)
NIR	<i>Near-infrared Spectroscopy</i> (Espectroscopia no Infravermelho Próximo)
nLv	<i>Number of Latent Variables</i> (Número de Variáveis Latentes)

nvars	<i>Number of Variables</i> (Número de Variáveis)
OPS	<i>Ordered Predictors Selection</i> (Seleção de Preditores Ordenados)
PC	<i>Principal Components</i> (Componentes Principais)
PCA	<i>Principal Component Analysis</i> (Análise de Componentes Principais)
PLS	<i>Partial Least Squares</i> (Mínimos Quadrados Parciais)
PLS-DA	<i>Partial Least Squares Discriminant Analysis</i> (Análise Discriminante por Quadrados Mínimos Parciais)
R	Coeficiente de Correlação
R ²	Coeficiente de Determinação
RMSEC	<i>Root Mean Squares Error of Calibration</i> (Raiz Quadrada do Erro Médio Quadrático de Calibração)
RMSECV	<i>Root Mean Squares Error of Cross Validation</i> (Raiz Quadrada do Erro Médio Quadrático de Validação Cruzada)
RMSEP	<i>Root Mean Squares Error of Prediction</i> (Raiz Quadrada do Erro Médio Quadrático de Previsão)
SEN	Sensibilidade
SNV	<i>Standard Normal Variate</i> (Variação Normal Padrão)
SPE	<i>Solid Phase Extraction</i> (Extração em Fase Sólida)
TFN	Taxa de Falsos Negativos
TFP	Taxa de Falsos Positivos
VN	Verdadeiras Negativas
VP	Verdadeiras Positivas
FN	Falsas Negativas
FP	Falsas Positivas

SUMÁRIO

INTRODUÇÃO GERAL	22
<i>CAPÍTULO 1 – REVISÃO BIBLIOGRÁFICA</i>	24
1.1. Azeite de oliva	25
1.2. Adulterações em azeite de oliva	27
1.2.1. Determinação de esqualeno em azeite de oliva	28
1.3. Espectroscopia no Infravermelho Próximo (NIR) para análise de azeite de oliva.....	30
1.4. Cromatografia gasosa para análise de azeite de oliva.....	31
1.5. Quimiometria.....	31
1.5.1. Análise de componentes principais (PCA).....	32
1.5.2. Regressão por mínimos quadrados parciais (PLS)	33
1.5.3. Análise discriminante por mínimos quadrados parciais (PLS-DA).....	35
1.5.4. Seleção de variáveis.....	38
1.5.4.1. Seleção de preditores ordenados (OPS)	38
1.5.4.2. Algoritmo genético (GA)	40
1.5.5. Fusão de dados	41
1.6. Validação de métodos qualitativos	43
<i>CAPÍTULO 2 – CONSTRUÇÃO DE MODELOS POR FT-NIR PARA VERIFICAÇÃO DA AUTENTICIDADE DE AZEITE DE OLIVA</i>	46
2.1. INTRODUÇÃO	47
2.2. OBJETIVOS	48
2.2.1. Objetivo geral.....	48
2.2.2. Objetivos específicos.....	48
2.3. MATERIAL E MÉTODOS.....	48
2.3.1. Amostras	48
2.3.2. Análise exploratória dos dados de FT-NIR.....	49
2.3.3. Construção dos modelos de classificação e calibração.....	50
2.3.3.1. Desenvolvimento e validação de modelos para classificação supervisionada e calibração multivariada das amostras	50
2.3.4. Análises por FT-NIR	51
2.4. RESULTADOS E DISCUSSÃO	52
2.4.1. Análise exploratória dos dados de FT-NIR para as amostras	52

2.4.2.	Construção dos modelos de classificação e calibração.....	62
2.4.2.1.	Desenvolvimento e validação de modelos para classificação supervisionada dos azeites de oliva varietais e blend	62
2.4.2.1.1.	Variedade Koroneiki	66
2.4.2.1.2.	Variedade Arbequina.....	68
2.4.2.1.3.	<i>Blend</i>	70
2.4.2.2.	Desenvolvimento e validação de modelos para calibração multivariada das amostras adulteradas	71
2.4.2.2.1.	Azeite lampante.....	72
2.4.2.2.2.	Azeite virgem.....	74
2.4.2.2.3.	Azeite tipo único	76
2.4.2.2.4.	Óleo de canola	78
2.4.2.2.5.	Óleo de girassol	80
2.4.2.2.6.	Óleo de milho	82
2.4.2.2.7.	Óleo de soja	84
2.4.2.3.	Discussão dos resultados	86
2.5.	CONCLUSÕES	88
<i>CAPÍTULO 3 – VALIDAÇÃO DE MÉTODO PARA DETERMINAÇÃO DO TEOR DE ESQUALENO EM AZEITE DE OLIVA POR GC/MS-MS</i>		90
3.1.	INTRODUÇÃO	91
3.2.	OBJETIVOS.....	92
3.2.1.	Objetivo geral.....	92
3.2.2.	Objetivos específicos	92
3.3.	MATERIAL E MÉTODOS.....	92
3.3.1.	Amostras	92
3.3.2.	Padrões	93
3.3.3.	Validação do método	93
3.3.3.1.	Linearidade	93
3.3.3.2.	Exatidão/Recuperação	93
3.3.3.3.	Precisão (repetitividade e precisão intermediária)	94
3.3.3.4.	Limite de Detecção (LD) e Limite de Quantificação (LQ)	94
3.3.4.	Análise de amostras de azeite de oliva adulteradas.....	94
3.3.5.	Análise de amostras de azeite de oliva coletadas para fiscalização.....	94
3.3.6.	Equipamentos.....	95
3.4.	RESULTADOS E DISCUSSÃO	95

3.4.1. Análises iniciais	95
3.4.2. Validação do método	97
3.4.3. Análise de amostras de azeite de oliva adulteradas.....	98
3.4.4. Análise de amostras de azeite de oliva coletadas para fiscalização....	100
3.5. CONCLUSÕES	102
<i>CAPÍTULO 4 – CONSTRUÇÃO DE MODELOS POR FUSÃO DE DADOS OBTIDOS POR FT-NIR E GC-MS/MS PARA VERIFICAÇÃO DA AUTENTICIDADE DE AZEITE DE OLIVA</i>	
4.1. INTRODUÇÃO	105
4.2. OBJETIVOS	106
4.2.1. Objetivo geral.....	106
4.2.2. Objetivos específicos.....	106
4.3. MATERIAL E MÉTODOS.....	106
4.3.1. Amostras	106
4.3.2. Construção dos modelos de calibração	106
4.4. RESULTADOS E DISCUSSÃO	107
4.4.1. Construção dos modelos de calibração.....	107
4.4.1.1. Azeite lampante	108
4.4.1.2. Azeite virgem	108
4.4.1.3. Azeite tipo único.....	109
4.4.1.4. Óleo de canola.....	110
4.4.1.5. Óleo de girassol	110
4.4.1.6. Óleo de milho	111
4.4.1.7. Óleo de soja.....	112
4.4.1.8. Modelo robusto	112
4.4.1.9. Parâmetros estatísticos para os modelos de calibração	113
4.4.2. Discussão dos resultados.....	114
4.5. CONCLUSÕES	116
CONCLUSÕES GERAIS.....	118
REFERÊNCIAS BIBLIOGRÁFICAS	119
ANEXO I.....	129
ANEXO II.....	134

INTRODUÇÃO GERAL

O azeite de oliva é um importante produto, cujas comercialização e produção têm crescido nos últimos anos, devido, principalmente, ao efeito nutricional da sua composição. Por ser um produto com alto valor de mercado, o azeite sofre processos de adulteração com adição de outros óleos, como girassol e soja, e de outras categorias de azeites (Hodaifa *et al.*, 2012).

Os métodos oficiais utilizados pelo Conselho Oleícola Internacional (IOC) para verificação da autenticidade de azeite de oliva, como a determinação de ceras e de Δ ECN 42, são demorados, pois incluem etapas de fracionamento e purificação dos constituintes do azeite em colunas de sílica, demandando grande quantidade de solvente e fase estacionária por amostra (IOC, 2003; IOC 2010). Além disso, métodos que utilizam outras técnicas, como FT-NIR, são limitados, dificultando, assim, a rapidez e precisão na verificação da autenticidade de azeite de oliva. Em função disso, fazem-se necessários estudos que tornem essa verificação mais eficiente.

Deste modo, esse trabalho tem como objetivo verificar a autenticidade de azeite de oliva por aplicação de métodos espectroscópicos, cromatográficos e quimiométricos, fusão de dados e seleção de variáveis. Esses processos serão empregados em amostras de azeite de oliva fraudadas com diferentes tipos de óleos vegetais e azeites de outras categorias, em diferentes concentrações. Ainda, será verificada a existência de adulterações em diferentes amostras de azeite de oliva coletadas para fiscalização.

No primeiro capítulo deste trabalho é apresentada uma revisão bibliográfica sobre os principais assuntos discutidos nesta tese. No segundo capítulo são apresentados os resultados referentes à construção de modelos de classificação e de calibração utilizando FT-NIR, PLS, PLS-DA e seleção de variáveis para diferenciar e quantificar amostras adulteradas de azeite de oliva extra virgem. Os resultados obtidos para os modelos de classificação apresentados neste capítulo foram publicados na revista Food Chemistry (Anexo II). No terceiro capítulo é apresentada a validação de um método para determinação de esqualeno, utilizando SPE e GC-MS/MS. O método validado é então utilizado para determinar o teor de

esqualeno de amostras de azeites de oliva extra virgem adulteradas e de amostras coletadas para fiscalização. No quarto capítulo são apresentados os modelos construídos por fusão de dados, PLS e seleção de variáveis, associando os dados obtidos pelas técnicas de espectroscopia (Capítulo 2) e cromatografia (Capítulo 3) para calibrar as amostras de azeite de oliva extra virgem adulteradas. O trabalho se encerra com as principais conclusões obtidas ao longo da tese.

CAPÍTULO 1 – REVISÃO BIBLIOGRÁFICA

1.1. Azeite de oliva

O azeite de oliva é definido como a fração lipídica obtida das azeitonas, frutos da oliveira (*Olea europaea* L.), por meio de processos mecânicos ou outros processos físicos, em condições de temperatura controlada, sem que isso cause qualquer alteração no produto. O azeite é um líquido viscoso, que apresenta características que o diferem de outros óleos vegetais (Codex Alimentarius Commission, 2003; Brasil, 2012).

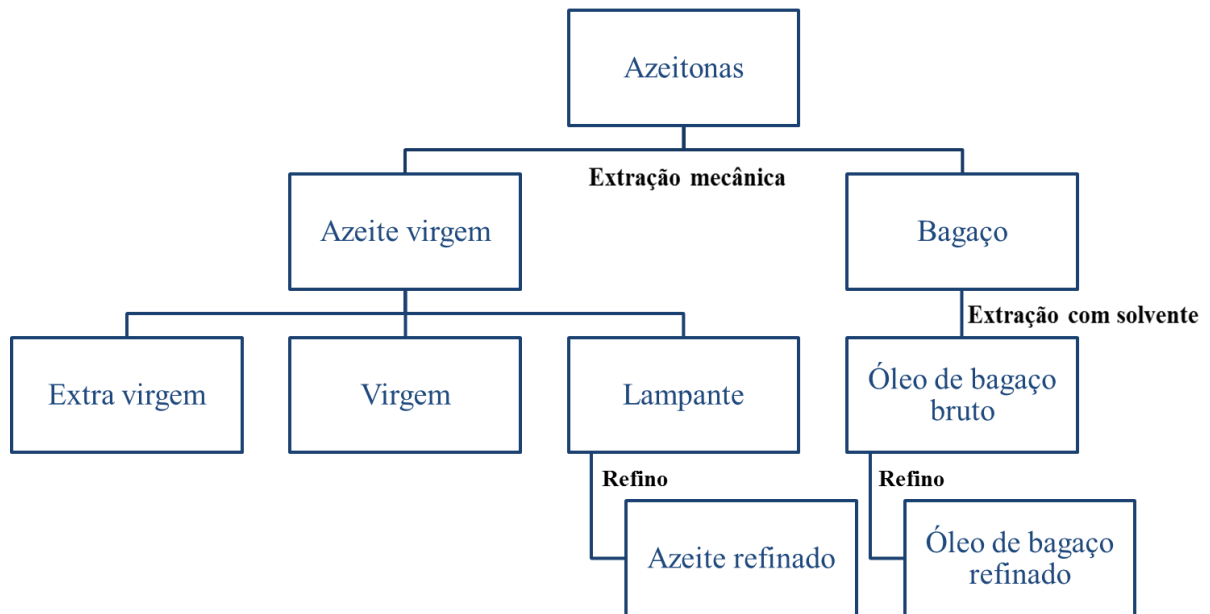
Óleo muito apreciado por seu gosto e seu aroma agradáveis, o azeite de oliva é um produto que pode ser consumido fresco ou ainda como ingrediente em diversos preparados alimentares, como conservas e molhos (Pérez-Castaño *et al.*, 2018; (Ruiz-Samblás *et al.*, 2012).

O azeite de oliva é composto majoritariamente (98 – 99%) por triacilgliceróis, ácidos graxos e fosfolipídeos, e minoritariamente (2 – 1%) por esteróis, álcoois, hidrocarbonetos e ceras. A distribuição dos ácidos graxos nas moléculas de triacilgliceróis e a composição em ácidos graxos no azeite de oliva variam de acordo com a região produtora, a variedade das azeitonas, a altitude de produção e os processos de extração do óleo, possibilitando a utilização destas informações para verificação da autenticidade do azeite de oliva (Ciafardini & Zullo, 2018; Harwood & Aparicio, 2000; Indelicato *et al.*, 2017). Os compostos minoritários são utilizados como parâmetros para identificação de adulterações em azeite de oliva por serem indicadores característicos da sua autenticidade e por sofrerem transformações durante o processamento do azeite de oliva (Boskou *et al.*, 2006; Ciafardini & Zullo, 2018; Harwood & Aparicio, 2000; Indelicato *et al.*, 2017).

O azeite de oliva é classificado em diferentes categorias em função do processo de produção, considerado desde a colheita e a lavagem das azeitonas até o armazenamento do óleo. O azeite virgem é obtido somente por processos mecânicos de extração, após moagem, sob condições que não alterem as características do produto. Estes processos incluem etapas de prensagem, percolação e centrifugação. Na categoria do azeite virgem estão incluídos os azeites extra virgem, virgem e lampante. Azeites de oliva de outras categorias são obtidos por meio de extrações com solventes (principalmente hexano) e de refino de

bagaços e extratos obtidos nas demais etapas (Figura 1) (Brasil, 2012; Harwood & Aparicio, 2000).

Figura 1 - Processo geral de obtenção do azeite de oliva.



FONTE: Adaptado de Brasil, 2012.

O azeite de oliva de melhor qualidade é denominado extra virgem e é obtido por extração mecânica em única etapa (com temperatura máxima de 28°C) de azeitonas recém colhidas e de boa qualidade. O azeite de oliva virgem apresenta qualidade semelhante ao extra virgem, diferindo apenas nos parâmetros acidez livre e absortividade molar (extinção específica), cujos valores são mais altos. Azeites de qualidade inferior incluem o azeite lampante e o azeite tipo único, sendo este uma mistura de azeite virgem e do azeite refinado e é o tipo mais encontrado no mercado brasileiro. O azeite lampante não é utilizado para consumo por apresentar sabor e aroma indesejáveis, originário de azeitonas de má qualidade (Brasil, 2012; Aued-pimentel *et al.*, 2008).

O consumo de azeite de oliva tem crescido em todo o mundo. Sua produção tem um grande impacto socioeconômico, alcançando a marca de mais de 3 milhões de toneladas produzidas na safra 2019/2020 (IOC, 2020).

O alto custo da produção faz com que o preço do azeite de oliva de maior qualidade seja elevado. Por isso, o azeite tem sido adulterado como uma forma de suprir uma grande procura a um produto de alto custo, adicionando ingredientes com propriedades compatíveis com o azeite de oliva de maior qualidade, como óleos vegetais e azeites de oliva de outras categorias, e que torna o custo de produção mais baixo e o lucro mais elevado (Aparicio *et al.*, 2013; Ciafardini & Zullo, 2018; Marcos Lorenzo *et al.*, 2002; Jiménez-Carvelo *et al.*, 2017).

Órgãos internacionais, como o IOC e o Codex Alimentarius, têm estabelecido critérios para garantir a autenticidade e a correta classificação e descrição dos tipos de azeite de oliva, e assim proteger o produto de possíveis adulterações, visando sua maior qualidade, para manter a confiança por parte dos consumidores (da Silveira *et al.*, 2017; Pereira *et al.*, 2018).

1.2. Adulterações em azeite de oliva

O tipo de adulteração mais comum em azeite de oliva é a adição de óleos vegetais, como o óleo de soja e o óleo de girassol, de menor valor comercial, principalmente quando o azeite é produzido em outros países e envasado no Brasil. Os métodos clássicos para identificação de adulterações em azeite de oliva, recomendados pelas legislações nacionais e internacionais, detectam apenas as adulterações mais grosseiras. Entretanto, outras adulterações, mais elaboradas e sofisticadas, são praticadas, sendo, em muitos casos, de difícil identificação devido às diferentes opções de adulterantes (Aued-pimentel *et al.*, 2008).

A presença de outros tipos de óleos vegetais no azeite de oliva extra virgem é normalmente avaliada por meio da determinação da composição em ácidos graxos e da determinação do coeficiente de absorvidade molar (extinção específica) (Brasil, 2012).

A análise da composição em ácidos graxos consiste na determinação da proporção dos diferentes ácidos graxos existentes na constituição das moléculas dos glicerídeos e de ácidos graxos livres, expresso em porcentagem relativa. As amostras analisadas que apresentam a composição em ácidos graxos diferente daquela determinada pela legislação como característica para o azeite de oliva, são classificadas como não genuína (adulterada). Já a determinação do coeficiente de

absortividade molar (extinção específica) avalia a adição de azeite de oliva de outras categorias ao azeite de oliva extra virgem por meio da absorção da luz ultravioleta a 232 nm e 270 nm, O aumento da absorção nestes comprimentos de onda indica a presença de compostos com duplas ligações conjugadas, de hidroperóxidos do ácido linoléico e de trienos conjugados, resultantes de processos oxidativos. Azeites de oliva de qualidade contém poucos produtos de oxidação e, portanto, valores baixos de absorção a 232 nm e 270 nm. Desta forma, valores maiores que 2,50 (232 nm) e 0,22 (270 nm) indicam que o azeite de oliva foi adulterado com azeites de outras categorias e óleos de sementes (Brasil, 2012; Antoniassi & Lago, 1998; Aued-pimentel *et al.*, 2008).

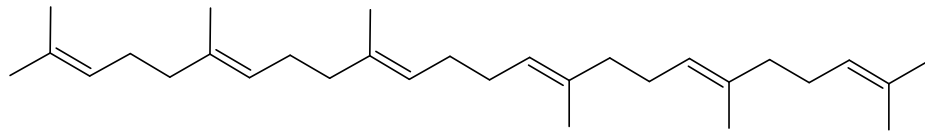
Para classificar o azeite de oliva, as legislações brasileiras consideram faixas amplas de composição e isso facilita a execução de adulterações ao azeite de oliva extra virgem com a adição de óleos de sementes e azeites de qualidades inferiores (Brasil, 2012). As adulterações do azeite de oliva têm implicações econômicas e nutricionais, pois o consumidor adquire o produto buscando seus benefícios nutricionais e não os encontra em azeites adulterados (Boskou *et al.*, 2006).

Novas metodologias têm sido empregadas para verificação da autenticidade de azeite de oliva, visando complementar informações e elucidar as adulterações. Entre essas metodologias estão a determinação do perfil dos triacilgliceróis, a evolução térmica do azeite de oliva adulterado, a determinação de compostos voláteis e a determinação do teor de esqualeno (Green *et al.*, 2020; Izquierdo *et al.*, 2020; Lima *et al.*, 2020; Ozcan-sinir, 2020; Pacetti *et al.*, 2019).

1.2.1. Determinação de esqualeno em azeite de oliva

O esqualeno é um triterpeno natural, com fórmula molecular $C_{30}H_{50}$ (Figura 2), principal componente (cerca de 90%) da fração de hidrocarbonetos do azeite de oliva (Samaniego-Sánchez *et al.*, 2010).

Figura 2 - Estrutura química do esqualeno.



Fonte: Própria autora.

O esqualeno tem recebido destaque na literatura por ser considerado o responsável pelos efeitos do azeite de oliva contra certos tipos de câncer (Salvo *et al.*, 2017) e um componente importante para a manutenção da estabilidade do azeite de oliva (Sagratini *et al.*, 2013).

O azeite de oliva é um dos alimentos mais ricos em esqualeno, representando cerca de 0,1 a 1,2% de sua composição, com teores variando de 1,0 a 12,0 g kg⁻¹, dependendo da variedade das azeitonas, estágio de amadurecimento da azeitona, técnica de extração do azeite e condições de armazenamento dos frutos (Aresta *et al.*, 2020). Em outros óleos, como o óleo de soja e o óleo de canola, o teor de esqualeno pode chegar a ser 70% menor em relação ao azeite de oliva extra virgem. Ainda, etapas de refino do azeite de oliva, que produzem, por exemplo, o azeite tipo único, causam a redução do teor de esqualeno do azeite em até 25%. Desta forma, a determinação do teor de esqualeno de amostras de azeite de oliva pode ser usada como uma ferramenta para identificação de adulterações pela adição de óleos de sementes ou de azeites de outras categorias no azeite de oliva extra virgem (Cayuela & García, 2018; Pacetti *et al.*, 2019).

O método oficial para determinação do teor de esqualeno em azeite de oliva envolve etapas de fracionamento e extrações em colunas de sílica, demandando grande quantidade de solvente e fase estacionária por amostra (AOAC, 1999). Assim, fazem-se necessários estudos que permitam a determinação de esqualeno em azeite de oliva por meio de métodos mais rápidos e que garantam menor consumo de solventes e insumos.

A determinação de esqualeno em azeite de oliva é realizada usualmente pelas técnicas de cromatografia líquida com detector por arranjo de diodos ou acoplada à espectrometria de massas, e a cromatografia gasosa com detector por ionização em

chama, aliadas a técnicas de preparo de amostras como a SPE (Aresta *et al.*, 2020; Grigoriadou *et al.*, 2007; Pacetti *et al.*, 2019; Sagratini *et al.*, 2013; Salvo *et al.*, 2017; Samaniego-Sánchez *et al.*, 2010). Neste trabalho, a proposta foi validar um método utilizando a SPE como técnica de extração e a GC-MS/MS para determinação de esqualeno em azeite de oliva.

1.3. Espectroscopia no Infravermelho Próximo (NIR) para análise de azeite de oliva

A Espectroscopia no Infravermelho Próximo (NIR) é um tipo de espectroscopia vibracional, cuja região compreende a faixa de radiação de 13300 a 4000 cm^{-1} (750 a 2500 nm). A NIR se caracteriza por ser uma técnica não invasiva e não destrutiva, capaz de analisar várias propriedades físicas e químicas das amostras (Pasquini, 2003; Bokobza, 1998).

Devido às suas vantagens em relação a outras técnicas, como menor tempo de análise e demanda mínima de preparo de amostra, a NIR tem se tornado uma importante ferramenta para estudos em áreas como a química, a farmácia e a agricultura. Sua grande desvantagem é a baixa sensibilidade a compostos minoritários das amostras (Blanco & Villarroya, 2002; Bokobza, 1998).

De forma geral, as bandas observadas na região do infravermelho próximo referem-se às vibrações das ligações C–H, N–H, O–H e S–H (Blanco & Villarroya, 2002). Como a grande maioria das moléculas possuem essas ligações, o número de informações espectrais geradas por uma amostra como azeite de oliva, quando analisada por NIR, é muito grande, fazendo-se necessária a utilização da quimiometria e seus métodos para interpretação dos resultados obtidos. Os métodos quimiométricos mais utilizados para tratamento dos dados obtidos por NIR são Análise de Componentes Principais (PCA), PLS-DA e PLS.

A utilização da técnica NIR para análise de azeite de oliva tem sido relatada. O trabalho desenvolvido por Sinelli e colaboradores (2010) mostrou que foi possível classificar amostras de azeites de oliva extra virgem monovarietais por NIR com base na composição de cada variedade. Já Jiménez-Carvelo e colaboradores (2019) mostraram que a utilização da técnica NIR permitiu identificar 100% das amostras de

azeite de oliva que foram adulteradas com óleo de milho, óleo de soja e óleo de girassol.

1.4. Cromatografia gasosa para análise de azeite de oliva

A cromatografia gasosa é um técnica analítica que permite separar e identificar compostos em diferentes matrizes. Técnica de alta detectabilidade e robusta, é capaz de quantificar os analitos em concentrações muito baixas (Collins *et al.*, 2006). Com diferentes formas de detecção, a técnica de cromatografia gasosa é muito utilizada para análises de alimentos, como o azeite de oliva (Borràs *et al.*, 2015; Aresta *et al.*, 2020; García-Nicolás *et al.*, 2020; Pérez-Castaño *et al.*, 2018).

A utilização da cromatografia gasosa para análise de azeite de oliva têm sido relatada na literatura, para identificação da origem geológica e botânica (Cecchi *et al.*, 2020; Revelou *et al.*, 2021), para determinação dos compostos voláteis (da Costa *et al.*, 2020) e para a classificação de azeites de oliva (García-Nicolás *et al.*, 2020).

A cromatografia gasosa com detector por ionização em chama é a técnica utilizada nos métodos clássicos de referência para verificação da autenticidade de azeite de oliva (IOC, 2003; IOC, 2018). Porém, novos métodos utilizando a cromatografia gasosa para identificação de adulterações em azeite de oliva, como por meio da determinação de ceras (Milani *et al.*, 2020) e da composição em ácidos graxos (Siano & Vasca, 2020) têm sido divulgados na literatura.

Neste trabalho a cromatografia gasosa acoplada à espectrometria de massas sequencial foi utilizada na validação de um método para determinação de esqualeno em azeite de oliva. A espectrometria de massas sequencial foi utilizada aliada à cromatografia gasosa devido à sua maior seletividade na determinação do analito em relação a outros detectores.

1.5. Quimiometria

A Quimiometria é uma ciência que utiliza métodos matemáticos e estatísticos para planejar, selecionar e/ou determinar condições ótimas de medidas e experimentos e extrair o máximo de informações de dados oriundos de um sistema ou processo químico (Ferreira, 2015).

A Quimiometria surge como uma ferramenta fundamental para a obtenção de respostas rápidas e precisas a partir de uma grande quantidade de dados gerados pelas técnicas instrumentais, como a espectroscopia e a cromatografia (Ferreira, 2015).

Os métodos mais utilizados na Quimiometria podem ser classificados em métodos de classificação não supervisionada, classificação supervisionada e calibração multivariada, de acordo com os objetivos das análises realizadas. Dentro destas classificações os métodos mais utilizados são a PCA para a classificação não supervisionada, a PLS-DA para a classificação supervisionada e a PLS para a calibração multivariada (Ferreira, 2015; Brereton, 2007).

1.5.1. Análise de componentes principais (PCA)

A análise de componentes principais é um método não supervisionado de classificação de amostras, inicialmente proposto em 1901 por Karl Pearson. O PCA também pode ser classificado com um método para análise exploratória de dados, que permite descobrir e interpretar de forma mais clara as diferenças existentes entre as variáveis e as relações entre as amostras (Ferreira, 2015; Brereton, 2007).

O método PCA se baseia na transformação de grandes matrizes de dados, como por exemplo dados espectroscópicos, em matrizes menores, extraindo os fatores mais importantes, os Componentes Principais (PC) dos dados, preservando a maior parte da variância, reduzindo a dimensionalidade do conjunto de dados, sem que as relações existentes entre as amostras sejam afetadas (Ferreira, 2015; Brereton, 2007).

Em termos matriciais, o PCA é um método que decompõe a matriz de dados em duas matrizes menores, como mostrado na equação 1:

$$X = TP^T = t_1p_1^T + t_2p_2^T + \dots + t_Rp_R^T \quad (1)$$

Onde:

X é a matriz de dados espectroscópicos;

T é a matriz que representa os escores;

P é a matriz que representa os pesos (“loadings”);
 t é o vetor dos escores;
 p é o vetor dos pesos (“loadings”).

Nesta decomposição, os escores expressam as relações entre as amostras e os pesos indicam as relações entre as variáveis.

Uma componente principal é, portanto, definida por um par de vetores escores e pesos t_R, p_R . As componentes principais descrevem o máximo de variância e são calculadas em ordem decrescente de importância. Desta forma, similaridades e diferenças entre as amostras podem ser vistas através de gráficos em que os escores ou os pesos das componentes principais são plotados uns contra os outros (Ferreira, 2015).

1.5.2. Regressão por mínimos quadrados parciais (PLS)

A regressão por mínimos quadrados parciais foi desenvolvida por Herman Wold no anos 1970 e é o método mais aplicado em análises por espectroscopia no infravermelho próximo (Ferreira, 2015; Geladi & Kowalski, 1986).

O método PLS consiste em encontrar a máxima correlação entre dois conjuntos de variáveis, encontrados na matriz X (contendo os dados espectroscópicos e/ou cromatográficos) e a matriz Y (contendo as concentrações das amostras). O PLS então decompõe as matrizes de dados em uma soma de variáveis latentes, de acordo com as equações 2 e 3.

As variáveis latentes são combinações lineares das variáveis originais, que descrevem a maior parte das informações das duas matrizes de dados (Ferreira, 2015; Geladi & Kowalski, 1986; Brereton, 2007).

$$X = TP^T = \sum t_R p_R^T \quad (2)$$

Onde:

X é a matriz de dados espectroscópicos e/ou cromatográficos;

T é a matriz que representa os escores;

P é a matriz que representa os pesos (“loadings”) da matriz X;
 t é o vetor dos escores;
 p é o vetor dos pesos (“loadings”).

$$Y = TQ^T = \sum t_R q_R^T \quad (3)$$

Onde:

Y é a matriz de concentrações das amostras;

T é a matriz que representa os escores;

Q é a matriz que representa os pesos (“loadings”) da matriz Y;

t é o vetor dos escores;

q é o vetor dos pesos (“loadings”).

O Número de Variáveis Latentes (nLv) do modelo é escolhido por meio da validação cruzada, de acordo com os menores valores de RMSECV encontrados. O número de variáveis latentes que corresponde ao modelo de menor RMSECV deve ser escolhido.

O desempenho dos modelos de calibração é avaliado pelo cálculo dos parâmetros RMSEC, RMSECV e RMSEP (equações 4 a 6), assim como os respectivos Coeficientes de Correlação (R) entre os valores de referência e previstos (Rcv, Rc e Rp).

$$\text{RMSEC} = \sqrt{\frac{\sum_{i=1}^I (y_i - \hat{y}_i)^2}{I - v}} \quad (4)$$

Onde:

y_i = valor de referência;

\hat{y}_i = valor estimado;

I = número de amostras no conjunto de calibração;

v = número de variáveis latentes do modelo + 1, se os dados forem centrados na média.

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^I (y_i - \hat{y}_i)^2}{I}} \quad (5)$$

Onde:

y_i = valor de referência;

\hat{y}_i = valor estimado;

I = número de amostras no conjunto de validação cruzada.

$$\text{RMSEP} = \sqrt{\frac{\sum_{p=1}^P (y_p - \hat{y}_p)^2}{P}} \quad (6)$$

Onde:

y_p = valor de referência;

\hat{y}_p = valor estimado;

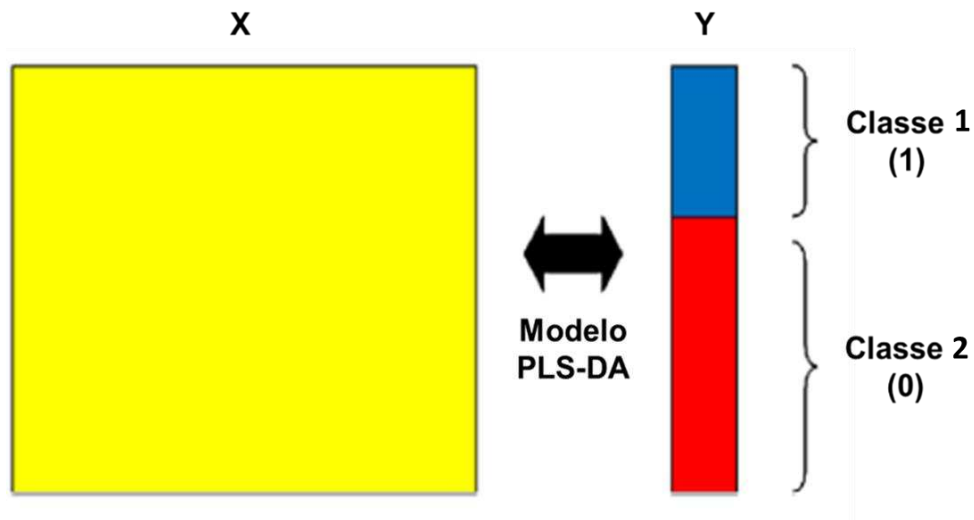
P = número total de amostras.

1.5.3. Análise discriminante por mínimos quadrados parciais (PLS-DA)

O PLS-DA é um método supervisionado de classificação de amostras. Ele foi uma adaptação baseada no método de calibração multivariada dos mínimos quadrados parciais, inicialmente proposto por Herman Wold, na década de 1960 (Ferreira, 2015).

Na Figura 3 é mostrada a representação gráfica do método PLS-DA.

Figura 3 - Representação gráfica do modelo PLS-DA.



FONTE: Adaptado de Brereton e Lloyd, 2014.

Na representação do método PLS-DA, a matrix X tem dimensões $I \times L$, onde I são as amostras de cada classe e L são as classes envolvidas. A matriz X contém as variáveis independentes (dados espectrais). A matriz Y consiste em um conjunto de vetores y que contem números inteiros que indicam a classe à qual pertence a amostra, sendo um vetor para cada classe (número de colunas da matriz igual ao número de classes).

Para a primeira classe ($l = 1$), o número 1 é atribuído às amostras desta classe, e o número 0 é atribuído a todas as outras amostras, não pertencente àquela classe. Em seguida, o mesmo é realizado na próxima coluna da matriz Y , e assim por diante, até que cada uma das classes seja representada pelo número 1, como mostrado na tabela representativa da atribuição dos valores binários para a matriz Y no modelo PLS-DA (Figura 4).

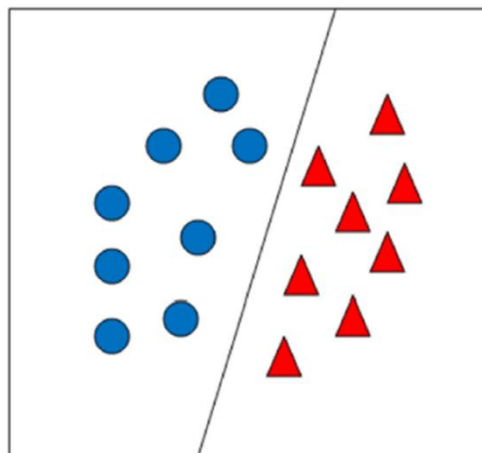
Figura 4 - Tabela representativa da atribuição dos valores binários para a matriz Y no modelo PLS-DA.

		Classes						
		1	2	3	4	5	6	7
Amostras	1	0	0	0	0	0	0	1
	2	0	0	1	0	0	0	0
	.							
	.							
	.							
	.							
	n							

Fonte: Própria autora.

Um vez atribuídos os números às classes, a etapa seguinte é a construção de um modelo de regressão utilizando o método PLS, em que as matrizes de dados são decompostas em uma soma de variáveis latentes. Essas variáveis latentes são responsáveis por determinar um novo espaço, que contém as informações das amostras e permite criar uma delimitação capaz de discriminá-las (Figura 5) (Brereton & Lloyd, 2014).

Figura 5 - Discriminação das amostras de duas classes de acordo com as classes definidas pela função discriminante determinada pelo método PLS-DA.



FONTE: Adaptado de Brereton e Lloyd, 2014.

Para a construção de modelos PLS-DA são utilizados dois conjuntos diferentes de amostras: o conjunto de treinamento, que contém os dados das amostras que serão utilizadas para a construção do modelo, e o conjunto teste, que contém amostras que não foram utilizadas para o cálculo dos parâmetros do modelo e serão utilizadas para verificar o seu desempenho (Ferreira, 2015).

1.5.4. Seleção de variáveis

As respostas analíticas geradas por técnicas como a espectroscopia e a cromatografia contém grande quantidade de variáveis. Quando um modelo é construído com um número grande de variáveis, sua interpretação é dificultada e se torna imprecisa. Dentro deste contexto, os métodos de seleção de variáveis têm sido uma ferramenta importante na análise dos dados (Ferreira, 2015; Yun *et al.*, 2015).

Os métodos de seleção de variáveis se baseiam no princípio de que a escolha de um número pequeno de variáveis de uma medida, em que são consideradas apenas as regiões espectrais e cromatográficas de interesse, pode produzir modelos com melhor capacidade preditiva, aumentando a robustez, a simplicidade e a precisão dos modelos, eliminando informações redundantes e ruidosas (Ferreira, 2015; Andersen & Bro, 2010; Balabin & Smirnov, 2011).

Existem vários métodos utilizados na literatura para a seleção de variáveis de dados de origens espectroscópicas e cromatográficas. Dentre eles, o mais utilizados são o OPS (Teófilo *et al.*, 2009) e o GA (Broadhurst *et al.*, 1997).

1.5.4.1. Seleção de preditores ordenados (OPS)

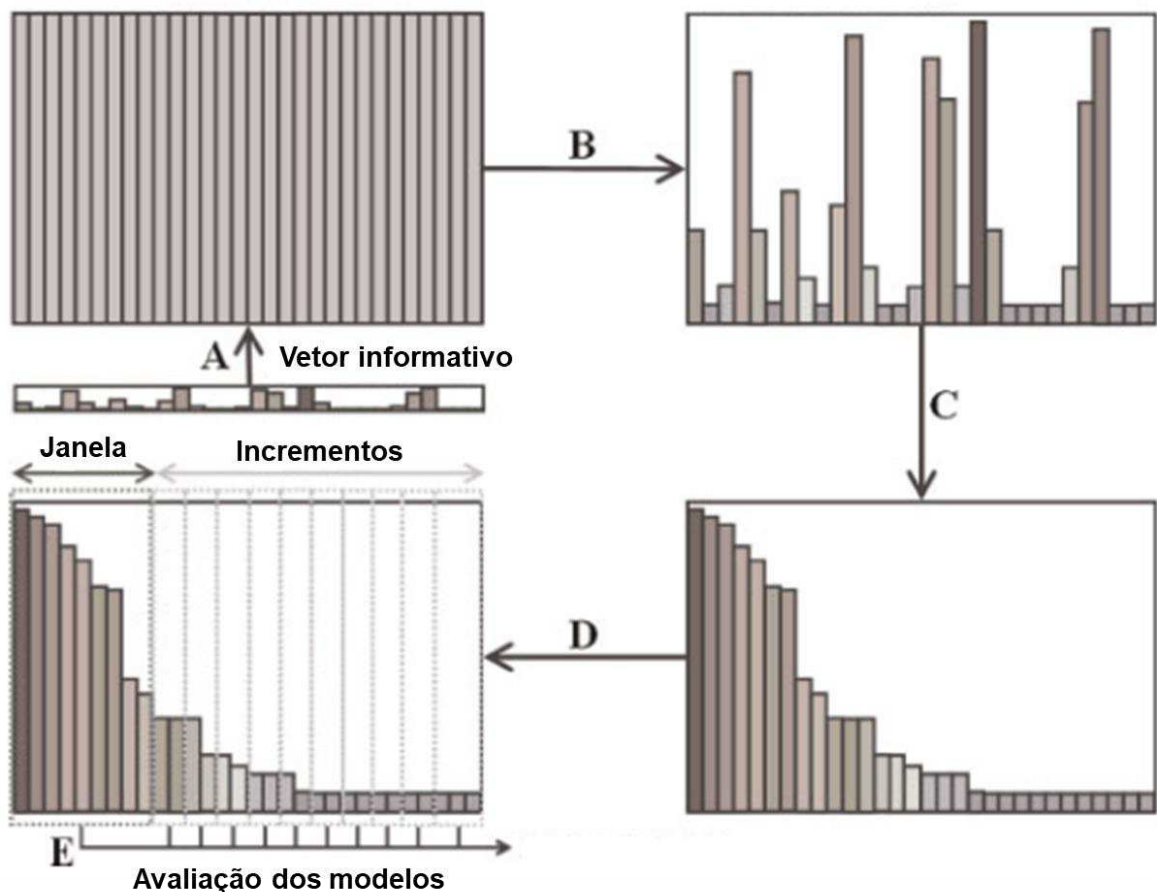
O método de seleção de preditores ordenados foi desenvolvido por Teófilo e colaboradores (2009) e tem como objetivo utilizar vetores informativos sobre as variáveis mais preditivas nos dados originais completos, realizando uma seleção destas variáveis (Teófilo *et al.*, 2009; Ferreira, 2015).

O método OPS utiliza um algoritmo que, por meio de análises sucessivas de subconjuntos formados por variáveis diferenciadas pelos vetores informativos, constrói modelos que apresentam menores erros de predição (menores valores de RMSECV durante a validação cruzada). Os vetores informativos, como os de

coeficientes de regressão, são estimados a partir de parâmetros do modelo e expressam a legítima variabilidade do modelo (originária exclusivamente dos componentes das amostras) e excluem toda a informação irrelevante e ruidosa (Teófilo *et al.*, 2009; Ferreira, 2015).

Na Figura 6 é mostrada a representação gráfica do método OPS.

Figura 6 - Representação gráfica do método OPS. (A): Construção do vetor informativo indicando as variáveis mais preditivas; (B): Diferenciação das variáveis segundo os valores absolutos dos elementos do vetor informativo; (C): Organização das variáveis diferenciadas em ordem decrescente de valor absoluto; (D): Definição dos subconjuntos formados pela janela de variáveis e seus incrementos; (E): Avaliações dos modelos para selecionar aquele que produz menor valor de RMSECV.



FONTE: Adaptado de Teófilo *et al.*, 2009 e Ferreira, 2015.

A primeira etapa (Figura 6A) consiste na construção do vetor informativo indicando as variáveis mais preditivas. O vetor informativo tem comprimento igual ao número de variáveis dos espectros originais completos. Na segunda etapa (Figura 6B), as variáveis originais são diferenciadas de acordo com os valores absolutos dos elementos do vetor informativo. Quanto maior o valor absoluto, mais importante é a variável. A terceira etapa (Figura 6C) consiste em organizar as variáveis diferenciadas em ordem decrescente, definindo um subconjunto ótimo de variáveis. Na quarta etapa (Figura 6D e 6E), a janela inicial (um conjunto das variáveis selecionadas) é utilizada para a construção de um modelo, e durante a validação cruzada, o valor de RMSECV é determinado. A partir daí, um número de variáveis (incrementos) é adicionado a este modelo e um novo valor de RMSECV é calculado. Este procedimento de adição de novos incrementos é realizado repetidas vezes até que seja obtido um valor mínimo de RMSECV. O conjunto das variáveis (janela inicial + incrementos) cujo modelo gera o menor valor de RMSECV, será o conjunto ótimo de variáveis selecionadas que serão utilizadas para a construção dos modelos finais (Teófilo *et al.*, 2009; Ferreira, 2015).

A grande vantagem da aplicação do método OPS para seleção de variáveis está na utilização de vetores informativos, pois estes possibilitam a diferenciação das variáveis que são mais facilmente interpretáveis de forma eficiente e totalmente automatizada (Teófilo *et al.*, 2009; Ferreira, 2015).

1.5.4.2. Algoritmo genético (GA)

O algoritmo genético, proposto pelo pesquisador John H. Holland nos anos 1960, é uma técnica de busca e otimização, inspirada no princípio Darwiniano de seleção natural e reprodução genética, muito utilizada para otimizar análises de dados espectroscópicos (Da Costa Filho & Poppi, 1999; Goicoechea *et al.*, 2010; Xiaobo *et al.*, 2010).

O GA consiste em cinco etapas: 1) codificação das variáveis, 2) criação da população inicial, 3) avaliação da resposta, 4) cruzamento e 5) mutação. Na etapa de codificação das variáveis, número binários são atribuídos às variáveis, sendo representadas por uma sequência de zeros e um, onde 0 indica que o intervalo codificado é descartado, enquanto 1 indica a inclusão das informações no modelo

PLS. Assim, na segunda etapa, por meio de um gerador randômico, encontra-se uma população inicial, formada pela combinação de todas as variáveis testadas. Pelo fato da população inicial ser criada de maneira aleatória, garante-se que não haverá nenhum tipo de influência tendenciosa (Da Costa Filho & Poppi, 1999; Goicoechea *et al.*, 2010; Xiaobo *et al.*, 2010).

Na terceira etapa, as respostas são avaliadas e indicam a probabilidade de uma variável ser selecionada, ou seja, apresentar maior capacidade de previsão. A resposta é obtida por meio do cálculo de um modelo de regressão para cada variável e estimando-se o valor do erro por validação cruzada (RMSECV). As variáveis cujos modelos apresentarem valores altos de RMSECV são descartados (Da Costa Filho & Poppi, 1999; Goicoechea *et al.*, 2010; Xiaobo *et al.*, 2010).

Na próxima etapa, as melhores variáveis são recombinadas, formando uma nova população pelo cruzamento aleatório. Com isso, prevalecerão as variáveis mais preditivas, o que leva à convergência do algoritmo (condição ótima).

A etapa final de mutação consiste em uma perturbação das informações contidas na nova população. De forma prática, a mutação ocorre a uma taxa baixa, sendo realizada aleatoriamente pela troca de um por zero, e vice-versa, em algumas variáveis, por vezes consecutivas (gerações). Após uma série de n gerações que tenham sido avaliadas, o subconjunto de variáveis que apresentar os melhores parâmetros estatísticos é definido e utilizado para construção do modelo (Da Costa Filho & Poppi, 1999; Goicoechea *et al.*, 2010; Xiaobo *et al.*, 2010).

1.5.5. Fusão de dados

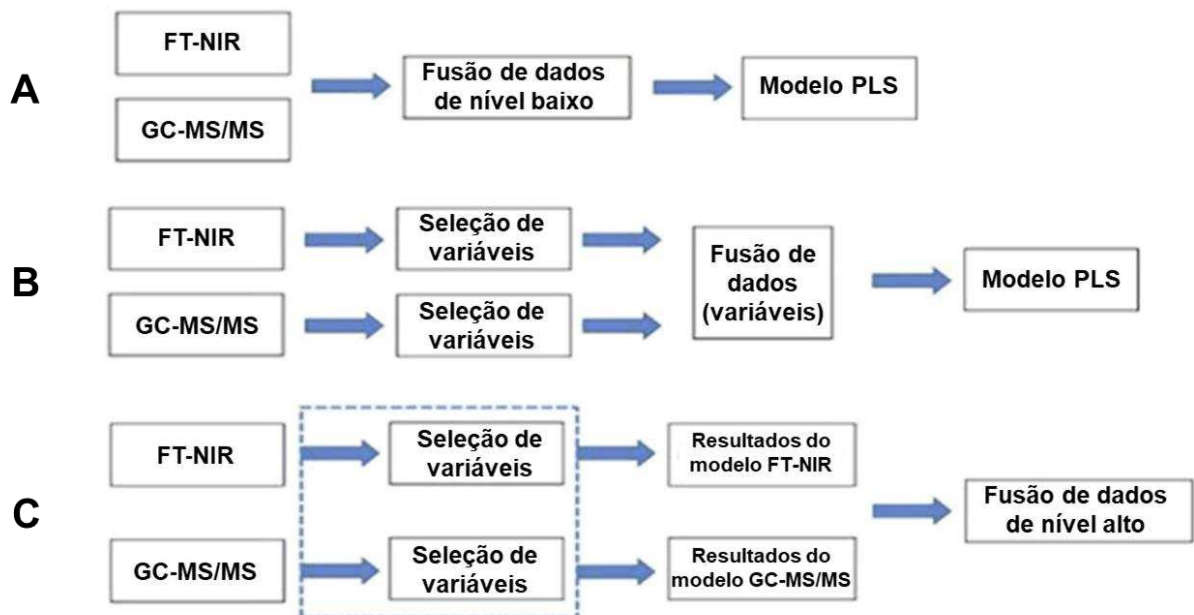
A fusão de dados é uma estratégia que combina no mesmo modelo medidas obtidas de diferentes fontes, tais como as técnicas analíticas instrumentais espectroscopia no infravermelho próximo, espectroscopia Raman, espectrometria de massas, cromatografia líquida e cromatografia gasosa. Os dados gerados por essas técnicas podem ser combinados de diferentes formas com o objetivo de construir modelos com melhores parâmetros estatísticos, se comparados com modelos que utilizam apenas uma técnica instrumental (Borràs *et al.*, 2015). Isso ocorre porque as técnicas fornecem informações distintas sobre as amostras e a fusão de dados

utiliza essas informações e abrange as vantagens de todas as técnicas instrumentais (Bevilacqua *et al.*, 2017).

A fusão de dados tem sido uma estratégia muito utilizada em investigações de adulterações em alimentos, como azeite de oliva. Esta ferramenta torna as análises cada vez mais específicas, pois a fusão reduz o impacto de fontes espúrias de variabilidade e potenciais interferentes nos dados, causando uma diminuição nos erros de predição dos modelos (Li *et al.*, 2019; Bevilacqua *et al.*, 2017).

A fusão de dados pode ocorrer em três níveis: baixo, médio e alto. A escolha do melhor nível de fusão é realizada após a geração dos dados e depende de vários fatores, como as técnicas instrumentais utilizadas, a quantidade de dados gerados, entre outros. Na Figura 7 é apresentado um esquema representativo do processo de fusão de dados, considerando a espectroscopia no infravermelho próximo e a cromatografia gasosa acoplada à espectrometria de massas como técnicas instrumentais cujos dados serão fundidos.

Figura 7 - Esquema representativo do processo de fusão de dados: A: Fusão de dados de nível baixo; B: Fusão de dados de nível médio; C: Fusão de dados de nível alto.



FONTE: Adaptado de Li *et al.*, 2019.

Na fusão de dados de nível baixo, as variáveis obtidas pelas diferentes fontes são diretamente unidas em uma única matriz. Dessa forma, essa informação é então utilizada para construir um modelo único PLS que fornece a calibração final. Neste nível a fusão ocorre diretamente nas matrizes de dados originais, fazendo com que o número de variáveis do modelo seja muito alto (Li *et al.*, 2019; Bevilacqua *et al.*, 2017).

Na fusão de dados de nível médio, a primeira etapa é a seleção de variáveis de cada fonte. Em seguida, estas informações são unidas em uma única matriz, que pode ser utilizada para classificação ou calibração multivariada. O principal desafio, portanto, é encontrar a combinação ideal de informações extraídas que forneça o melhor modelo final. Dessa forma, como esta técnica extrai apenas a variação relevante nas diferentes matrizes de dados, a estratégia de nível médio geralmente é mais eficaz do que a de baixo nível (Li *et al.*, 2019; Bevilacqua *et al.*, 2017).

Já na fusão de dados de nível alto, modelos de classificação ou calibração multivariada são construídos separadamente para cada fonte de dados e os resultados finais de cada modelo individual são unidos, gerando uma única resposta final. O principal desafio é construir modelos que funcionem adequadamente para cada bloco de modo que sua combinação tenha um desempenho consideravelmente melhor que os modelos individuais (Li *et al.*, 2019; Bevilacqua *et al.*, 2017).

Neste trabalho, foi utilizada a estratégia da fusão de dados de nível baixo, com etapa posterior de seleção de variáveis. A fusão de dados de nível baixo apresentou os melhores resultados porque as técnicas de espectroscopia e cromatografia geram dados de diferentes magnitudes, e para os demais níveis pouca informação foi explicada para uma variação de dados como as encontradas neste trabalho.

1.6. Validação de métodos qualitativos

A validação de métodos qualitativos é uma ferramenta importante, que confirma e fornece evidências objetivas de que o método atende aos requisitos para o qual se destina, sendo confiável quanto à sua capacidade preditiva (Trullols *et al.*, 2004; Ferreira, 2015).

O desempenho de métodos qualitativos é avaliado pelo cálculo de figuras de mérito utilizando os valores apresentados pela tabela de contingência, como exemplificada na Tabela 1.

Tabela 1 - Exemplo de tabela de contingência para avaliação do desempenho de métodos qualitativos.

		Resultado verdadeiro	
		Verdadeiro	Falso
Resultado previsto	Verdadeiro	VP	FP
	Falso	FN	VN

Nas tabelas de contingência, os resultados das classificações das amostras são apresentadas como: número de amostras classificadas corretamente como verdadeiras, Verdadeiras Positivas (VP); número de amostras classificadas corretamente como falsas, Verdadeiras Negativas (VN); número de amostras classificadas erroneamente como verdadeiras, Falsas Positivas (FP); e número de amostras classificadas erroneamente como falsas, Falsas Negativas (FN) (Trullols *et al.*, 2004).

A partir dos resultados obtidos pela tabela de contingência, as figuras de mérito do método são calculadas, por meio das equações 7 a 12.

$$SEN = \frac{VP}{VP + FN} \quad (7)$$

$$Esp = \frac{VN}{VN + FP} \quad (8)$$

$$TFN = \frac{FN}{VP + FN} \quad (9)$$

$$TFP = \frac{FP}{VN + FP} \quad (10)$$

$$Exa = \frac{VP + VN}{VP + VN + FP + FN} \quad (11)$$

$$Con = SEN + Esp - 1 \quad (12)$$

A Sensibilidade (SEN) (equação 7) e a Especificidade (Esp) (equação 8) representam as taxas de classificações corretas, e seus valores devem ser tão altos quanto possíveis (máximo de 100%). Já as Taxas de Falsos Negativos (TFN) (equação 9) e as Taxas de Falsos Positivos (TFP) (equação 10) são uma representação matemática da razão de amostras classificadas erroneamente e devem ser valores baixos (idealmente iguais a 0%). A Exatidão (Exa) (equação 11) é uma representação das classificações corretas, considerando-se o número total de amostras. Já a Confiabilidade (Con) (equação 12) indica o desempenho do método para a classificação das amostras, considerando as previsões corretas de cada classe independentemente. Ambos os parâmetros devem ter o maior valor possível (100%) (Trullols *et al.*, 2004).

***CAPÍTULO 2 – CONSTRUÇÃO DE MODELOS POR
FT-NIR PARA VERIFICAÇÃO DA AUTENTICIDADE
DE AZEITE DE OLIVA***

2.1. INTRODUÇÃO

O exigente processo de produção do azeite de oliva extra virgem faz com que esse produto seja caro. Por esse motivo, ele é adulterado com óleos vegetais (como óleo de girassol e óleo de soja) e azeites das demais categorias, que são mais baratos, fazendo com que os fraudadores tenham ganhos econômicos (Jiménez-Carvelo *et al.*, 2017; Marcos Lorenzo *et al.*, 2002; Aparicio *et al.*, 2013). Órgãos internacionais, como o IOC, a União Europeia (EU) e o Codex Alimentarius, têm estabelecido critérios para garantir a autenticidade e a correta classificação e descrição dos tipos de azeite de oliva, e assim proteger o consumidor de adquirir um produto adulterado (Jiménez-Carvelo *et al.*, 2017; Pereira *et al.*, 2018).

Os métodos utilizados pelos órgãos oficiais para verificação da autenticidade de azeite de oliva utilizam técnicas clássicas como a cromatografia gasosa e a cromatografia líquida. Ainda, os métodos incluem etapas de fracionamento e purificação dos constituintes do azeite de oliva em colunas de sílica, tornando-os demorados e demandando grande quantidade de solvente e fase estacionária por amostra (da Silveira *et al.*, 2017; IOC, 2003; IOC, 2010). Devido às suas vantagens em relação a outras técnicas, como menor tempo de análise e demanda mínima de preparo de amostra, técnicas vibracionais como a NIR (Cayuela & García, 2018; Mustorgi *et al.*, 2020; Gertz *et al.*, 2020; Sohng *et al.*, 2020; Yan *et al.*, 2019; Jiang & Chen, 2019; Mossoba *et al.*, 2017; Azizian *et al.*, 2016; Wu *et al.*, 2016) têm sido utilizadas para a verificação da autenticidade de azeite de oliva. Como uma grande quantidade de dados é obtida nas análises por técnicas vibracionais, ferramentas quimiométricas, como PCA e PLS-DA, são utilizadas para estabelecer modelos matemáticos que permitam diferenciar as amostras (Vaclavik *et al.*, 2009).

O objetivo deste capítulo foi construir modelos sensíveis e confiáveis baseados no uso de FT-NIR associada à análise exploratória dos dados, à classificação supervisionada e à calibração multivariada para classificar e calibrar amostras de azeite de oliva extra virgem adulteradas com óleos vegetais e azeites de outras categorias.

2.2. OBJETIVOS

2.2.1. Objetivo geral

Utilizar a espectroscopia no infravermelho próximo associada à análise exploratória dos dados, à classificação supervisionada e à calibração multivariada para verificação da autenticidade de azeite de oliva.

2.2.2. Objetivos específicos

- Preparar amostras de azeites de oliva extra virgem adulteradas com óleos de sementes e azeite de outras categorias;
- Analisar as amostras preparadas por FT-NIR;
- Realizar uma análise exploratória dos dados de FT-NIR pelo método PCA para diferenciar amostras adulteradas de azeite de oliva extra virgem (varietais e *blend*);
- Construir e validar modelos utilizando FT-NIR, PLS-DA e os métodos OPS e GA para diferenciar amostras adulteradas de azeite de oliva extra virgem (varietais e *blend*);
- Construir modelos robustos utilizando FT-NIR, PLS-DA e o método OPS para a classificação dos adulterantes óleo de soja, óleo de milho, óleo de canola, óleo de girassol, azeite virgem, azeite lampante e azeite tipo único em azeite de oliva extra virgem (varietais e *blend*);
- Construir modelos individuais e robustos utilizando FT-NIR, PLS e o método OPS para quantificar amostras de azeite de oliva extra virgem (varietais e *blend*) adulteradas.

2.3. MATERIAL E MÉTODOS

2.3.1. Amostras

Foram utilizadas para a realização das análises amostras de azeite de oliva extra virgem produzidas por empresas localizadas em Minas Gerais e no Rio Grande do Sul, a partir das variedades de azeitonas Arbequina e Koroneiki,

respectivamente. Foi utilizado também um *blend* dos azeites das duas empresas, produzido no laboratório, que consistia em uma mistura de 50% m/m de cada variedade. As amostras adulteradas foram preparadas utilizando sete adulterantes: óleo de soja, óleo de milho, óleo de canola, óleo de girassol, adquiridos em comércios locais de Belo Horizonte, MG; e de azeites de oliva virgem, lampante e tipo único, adquiridos no Laboratório de Identidade e Qualidade de Alimentos (IQA)/ Laboratório Federal de Defesa Agropecuária – Minas Gerais (LFDA-MG). As amostras adulteradas (azeite de oliva extra virgem com cada adulterante) foram preparadas adicionando-se quantidades crescentes (1%, 2%, 5%, 10%, 15% e 20% m/m) de cada adulterante em cada variedade e no *blend* preparado (Anexo I, Figura 1A). Cada mistura preparada continha uma massa total próxima de 3,0000 g. As massas das amostras adulteradas foram medidas em balança analítica (Kern, modelo ALJ 220-4, com precisão de 0,0001 g). Cada amostra adulterada foi preparada em quintuplicata. Desta forma, foram preparadas 630 amostras adulteradas. Os azeites e os óleos foram analisados no IQA/LFDA-MG (Belo Horizonte – MG) e eles foram aprovados nos parâmetros determinados pela legislação de fiscalização federal do Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Os azeites de oliva puros (varietais e *blend*), as amostras adulteradas e os adulterantes puros foram submetidos à análise por FT-NIR. Essas mesmas amostras foram utilizadas também nos demais capítulos.

2.3.2. Análise exploratória dos dados de FT-NIR

Os espectros gerados por todas as análises por FT-NIR foram exportados na extensão .xls e importados pelo software OriginPro 2018 (OriginLab Corporation, Northampton, USA). Os modelos foram desenvolvidos no pacote Principal Component Analysis for Spectroscopy, versão 1.1 (OriginLab Technical Support, Northampton, USA).

O PCA foi utilizado para desenvolver, inicialmente, um modelo de análise exploratória dos dados para discriminar as amostras (varietais e *blend*), utilizando todas as amostras preparadas (com e sem adulteração).

2.3.3. Construção dos modelos de classificação e calibração

2.3.3.1. Desenvolvimento e validação de modelos para classificação supervisionada e calibração multivariada das amostras

Os espectros gerados por todas as análises por FT-NIR foram exportados na extensão .xls e importados pelo software Matlab 7.13 (Math Works, Natick, EUA). Os modelos foram desenvolvidos no pacote PLS Toolbox, versão 6.7.1 (Eigenvector Technologies, Manson, WA, EUA).

O PLS-DA foi utilizado para desenvolver, inicialmente, um modelo de classificação supervisionado para discriminar as amostras (varietais e *blend*), utilizando todas as amostras preparadas (com e sem adulteração). A partir deste, foram construídos modelos individuais para classificação dos adulterantes.

O PLS foi utilizado para desenvolver modelos individuais e robustos (que englobam as amostras varietais e *blend* em um só modelo) para a calibração multivariada das amostras adulteradas.

Para divisão do conjunto de dados em calibração e validação, as amostras foram divididas em 2/3 para o conjunto de calibração e 1/3 para o conjunto de validação, através de planejamento experimental. Dessa forma, as amostras foram separadas pela ordem Calibração, Validação, Calibração (CVC), garantindo a presença de amostras em toda a faixa de concentração. Para a escolha do melhor n_{LV} , a validação cruzada foi aplicada, utilizando subconjuntos aleatórios, com 10 subconjuntos e 20 interações.

Diferentes pré-processamentos foram testados para a matriz, tais como: (1) primeira derivada; (2) Correção do Espalhamento Multiplicativo (MSC) e (3) Variação Normal Padrão (SNV). O vetor y foi centrado na média em todos os modelos.

O desempenho dos modelos de classificação PLS-DA foi avaliado pelo cálculo de figuras de mérito, utilizando os valores apresentados pela tabela de contingência: VP, VN, FP e FN.

A SEM e a Esp também foram calculadas. Elas representam as taxas de classificações corretas, e seus valores devem ser tão altos quanto possíveis (máximo de 100%). Já as TFN e as TFP são uma representação matemática da

razão de amostras classificadas erroneamente e devem ser valores baixos (idealmente igual a 0%).

Neste trabalho, os valores de Exa e Con foram calculados. A exatidão é uma representação das classificações corretas, considerando-se o número total de amostras. Já a confiabilidade indica o desempenho do método para a classificação das amostras, considerando as previsões corretas de cada classe independentemente.

Para os modelos de calibração PLS, o desempenho foi avaliado pelo cálculo dos parâmetros RMSEC, RMSECV e RMSEP, assim como os respectivos coeficientes de correlação entre os valores de referência e previstos (R_{cv} , R_c e R_p).

Para a seleção de variáveis, diferentes condições foram testadas e aquelas que apresentaram menores valores de RMSECV foram selecionadas. Os métodos OPS (Teófilo *et al.*, 2009) e GA (Broadhurst *et al.*, 1997) foram aplicados para identificar subconjuntos de variáveis com maior poder de predição. A rotina OPS está disponível em <http://www.deq.ufv.br/chemometrics>. O OPS foi executado com as seguintes configurações otimizadas: vetor informativo: coeficientes de regressão; janela inicial: 5; e incrementos: 5. Para o GA, as condições otimizadas foram: população (64); gerações (120); taxa de mutação (0.005); largura de janela: 1; convergência: 50; incrementos: 10.

2.3.4. Análises por FT-NIR

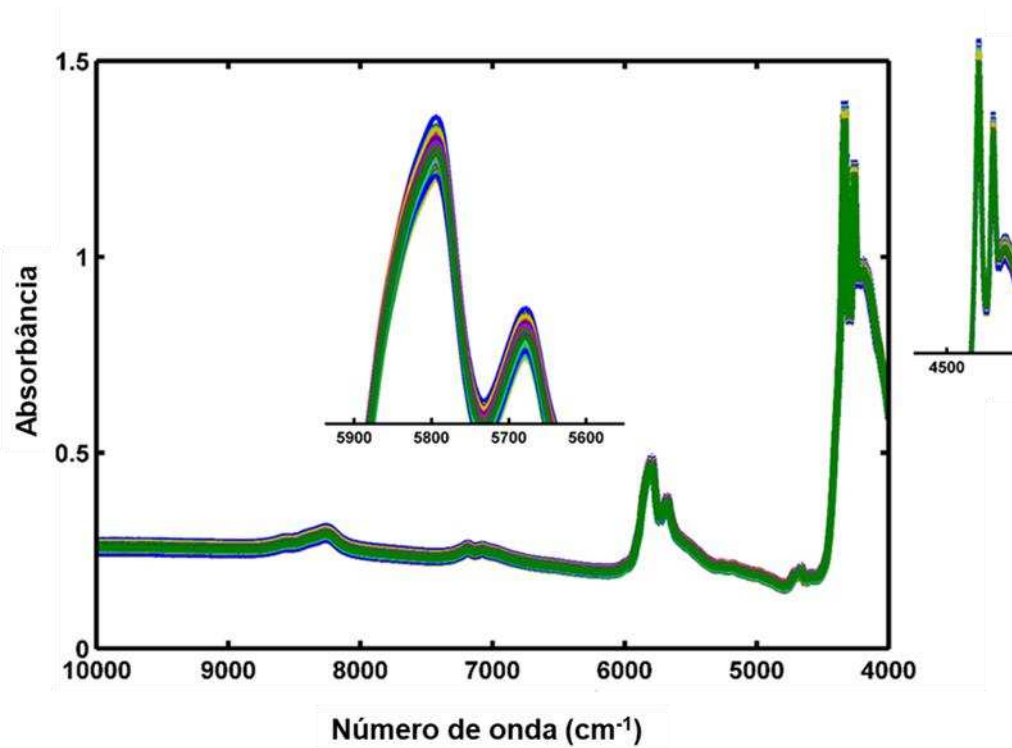
Os azeites de oliva (varietais e *blend*) puros, as amostras adulteradas e os adulterantes puros foram analisadas utilizando um espectrofotômetro (Büchi NIRFlex N-500) com transformada de Fourier, acessório em cromo e suporte para amostras do tipo XL (Anexo I, Figuras 2A a 4A). A faixa analisada foi de $10000 - 4000 \text{ cm}^{-1}$, com resolução de 8 cm^{-1} . Os espectros foram coletados utilizando o software Büchi NIR Operator. Para cada amostra, 32 varreduras foram realizadas. Os espectros foram obtidos no modo de transfectância e os dados foram convertidos para absorbância. Os espectros das amostras foram gerados de forma aleatória, visando não incluir variações sistemáticas no modelo.

2.4. RESULTADOS E DISCUSSÃO

2.4.1. Análise exploratória dos dados de FT-NIR para as amostras

Os espectros obtidos por FT-NIR para amostras de azeite de oliva extra virgem varietais e *blend* puros, para os adulterantes lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja puros e para todas as amostras adulteradas, na faixa de 10.000 - 4.000 cm^{-1} , são apresentados na Figura 8. Em todos os modelos, o melhor pré-processamento foi MSC seguido de centragem dos dados na média.

Figura 8 - Espectros obtidos por FT-NIR para amostras de azeite de oliva extra virgem varietais e *blend* puros, para os adulterantes lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja puros e para todas as amostras adulteradas.



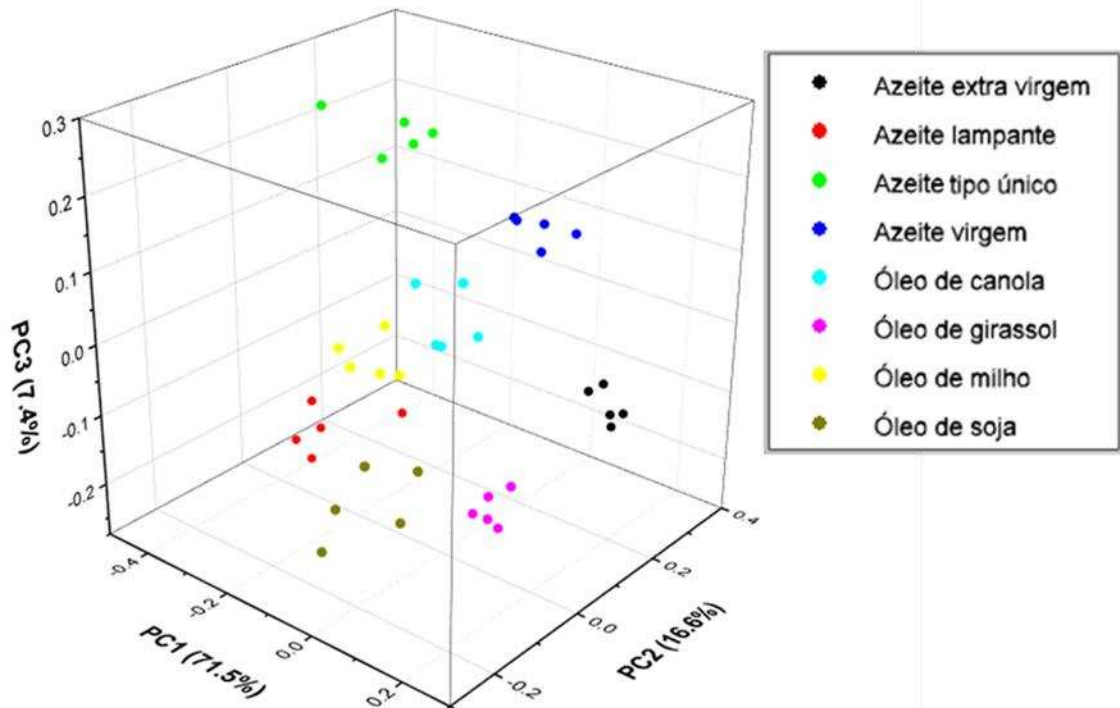
Observando a Figura 8, pode-se notar que os perfis espectrais brutos de todas as amostras de azeite de oliva são muito semelhantes e é muito difícil distinguir as amostras. As diferenças estão, essencialmente, nas intensidades de absorção associadas a cada número de onda. As bandas intensas em 4260 cm^{-1} e 4336 cm^{-1} estão associadas à combinação de vibrações do modo de estiramento de ligações C–H com outros modos vibracionais (Hourant *et al.*, 2000). As bandas discretas presentes em 4664 cm^{-1} e 4788 cm^{-1} são resultado de combinações de vibrações do modo de estiramento de ligações duplas C=C e ligações C–H de ácidos graxos insaturados (Cozzolino *et al.*, 2005). Uma banda muito discreta se apresenta em 4704 cm^{-1} , resultado de combinações de vibrações do modo de estiramento do grupo COOR (Christy *et al.*, 2004). As bandas em 5796 cm^{-1} e 5680 cm^{-1} ocorrem devido ao primeiro harmônico de ligações C–H de grupos CH_2 . Essas bandas estão associadas aos triacilgliceróis, principais componentes do azeite de

oliva (Christy *et al.*, 2004; Sato *et al.*, 1991). Em 8264 cm^{-1} e 7188 cm^{-1} observa-se bandas como resultado do segundo harmônico do modo de estiramento de ligações C–H de grupos CH_2 e como combinações de vibrações de ligações C–H de grupos CH_3 terminais, respectivamente. Essas regiões espectrais são características dos ácidos graxos oleico e linoleico, cujos teores são diferentes no azeite de oliva e nos demais óleos (Christy *et al.*, 2004; Holman & Edmondson, 1956). As variações nas intensidades das bandas de absorção acima descritas se relacionam às diferenças de composição entre o azeite de oliva extra virgem e os demais adulterantes.

A fim de entender as diferenças espectrais entre as amostras e verificar a possível discriminação das amostras por FT-NIR, uma análise exploratória de dados pelo método PCA foi realizada para as variedades Koroneiki e Arbequina e para o *blend* das variedades.

A Figura 9 apresenta o *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Koroneiki puro e para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após aplicação do método PCA.

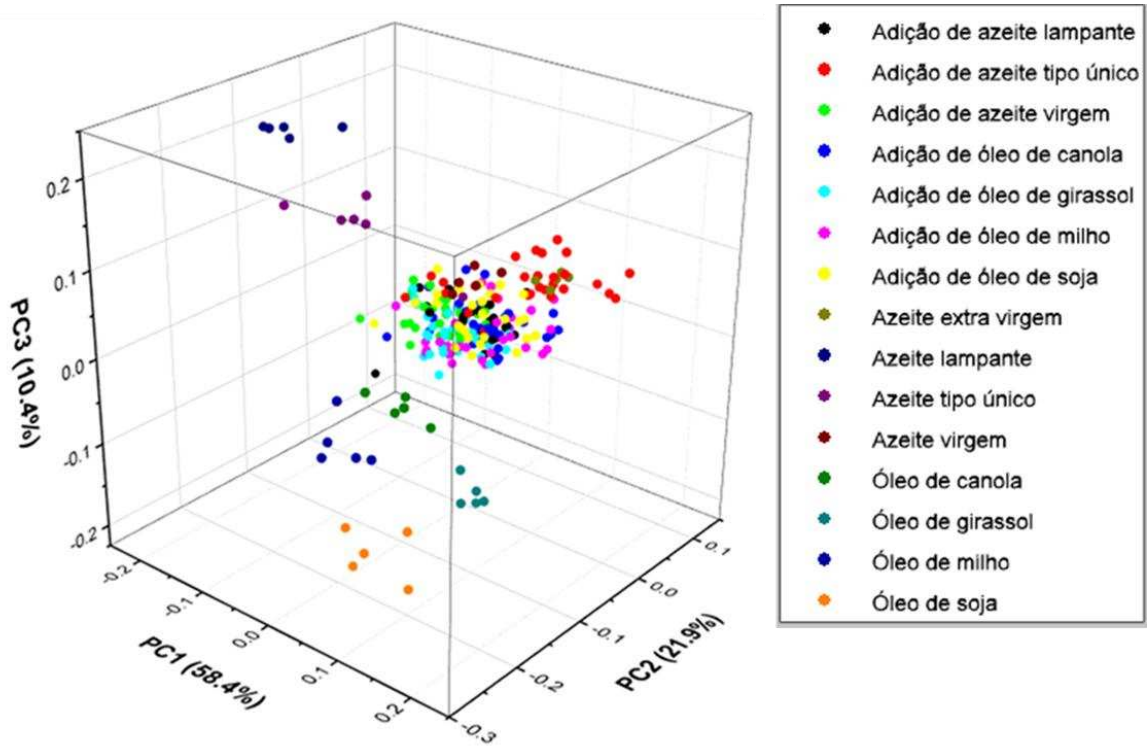
Figura 9 - *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Koroneiki puro e para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após a aplicação do método PCA.



Os resultados apresentados na Figura 9 mostram que as três PCs explicam 95,5% (71,5% + 16,6% + 7,4%) dos dados e estes estão com grande dispersão entre as replicatas. O *Score plot* indica ainda que houve a separação dos grupos referentes às amostras de azeite de oliva extra virgem da variedade Koroneiki e aos adulterantes puros. Isso indica que a técnica FT-NIR foi capaz de diferenciar o azeite extra virgem dos demais óleos.

A Figura 10 apresenta o *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Koroneiki puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.

Figura 10 - *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Koroneiki puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.



O *Score plot* (Figura 10) mostra que por FT-NIR foi possível diferenciar os adulterantes puros do azeite de oliva extra virgem, com a formação de seus grupos, com três PCs explicando 90,7% (58,4% + 21,9% + 10,4%) dos dados. Porém, não foi possível discriminar as amostras de azeite de oliva extra virgem das amostras adulteradas.

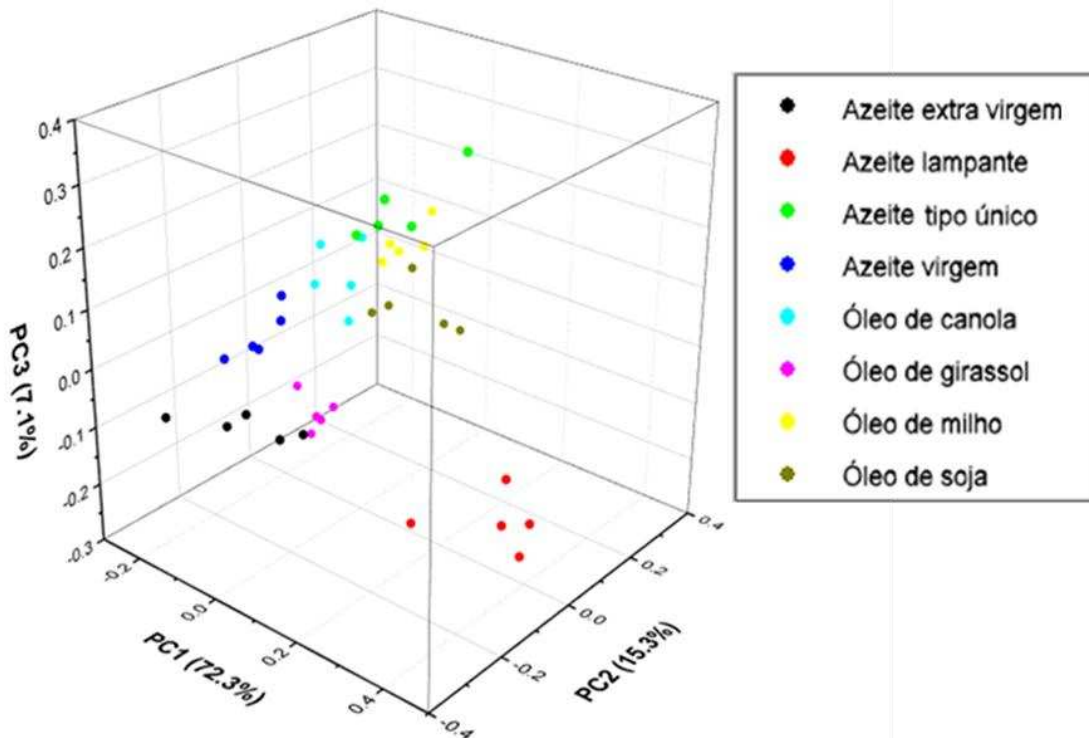
A adição dos adulterantes no azeite de oliva extra virgem nas porcentagens de adulteração analisadas não causaram diferenças espectrais suficientes para que fossem discriminados aplicando o método PCA para o azeite extra virgem da variedade Koroneiki. Como o número de informações espectrais é grande, as alterações nos valores de absorvância gerados pela adição de algum adulterante no azeite de oliva extra virgem foram pequenas, gerando valores de *scores* muito similares. Estes valores foram interpretados pelo modelo, que indicou as amostras

como semelhantes, dificultando, assim, os agrupamentos e a discriminação entre o azeite extra virgem da variedade Koroneiki e o azeite adulterado. Desta forma, a utilização de outros métodos quimiométricos pode fazer com o que o modelo seja capaz de discriminar melhor as amostras e aprimorar os agrupamentos, permitindo verificar a autenticidade do azeite de oliva.

Uma análise exploratória de dados pelo método PCA também foi realizada para a variedade Arbequina.

A Figura 11 apresenta o *Score plot* dos dados de FT-NIR para as amostras de azeite de oliva extra virgem da variedade Arbequina puro e para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após aplicação do método PCA.

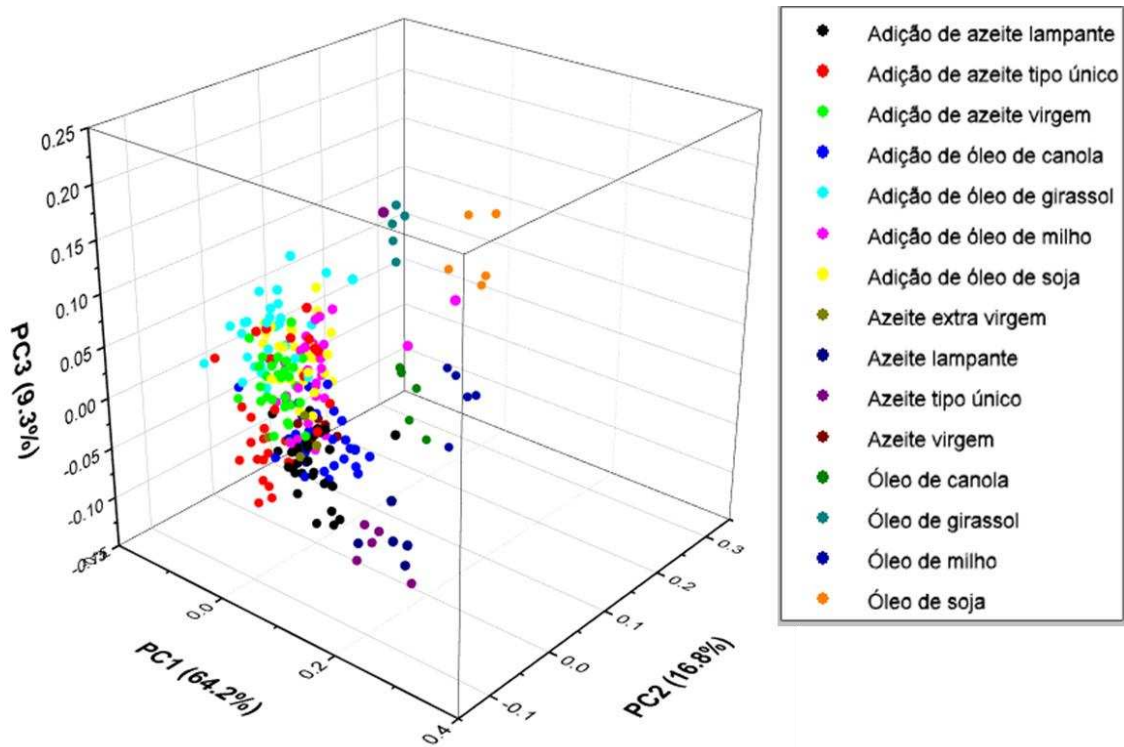
Figura 11 - *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Arbequina e para os adulterantes lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após aplicação do método PCA.



A Figura 11 mostra que as três PCs explicam 94,7% (72,3% + 15,3% + 7,1%) dos dados e estes estão com aceitável dispersão entre as replicatas. O gráfico *Score plot* indica que houve a separação dos grupos referentes à amostra de azeite de oliva extra virgem da variedade Arbequina puro e dos adulterantes puros, mostrando que a técnica foi capaz de discriminar o azeite extra virgem dos demais óleos.

A Figura 12 apresenta o *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Arbequina puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.

Figura 12 - *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem da variedade Arbequina puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.



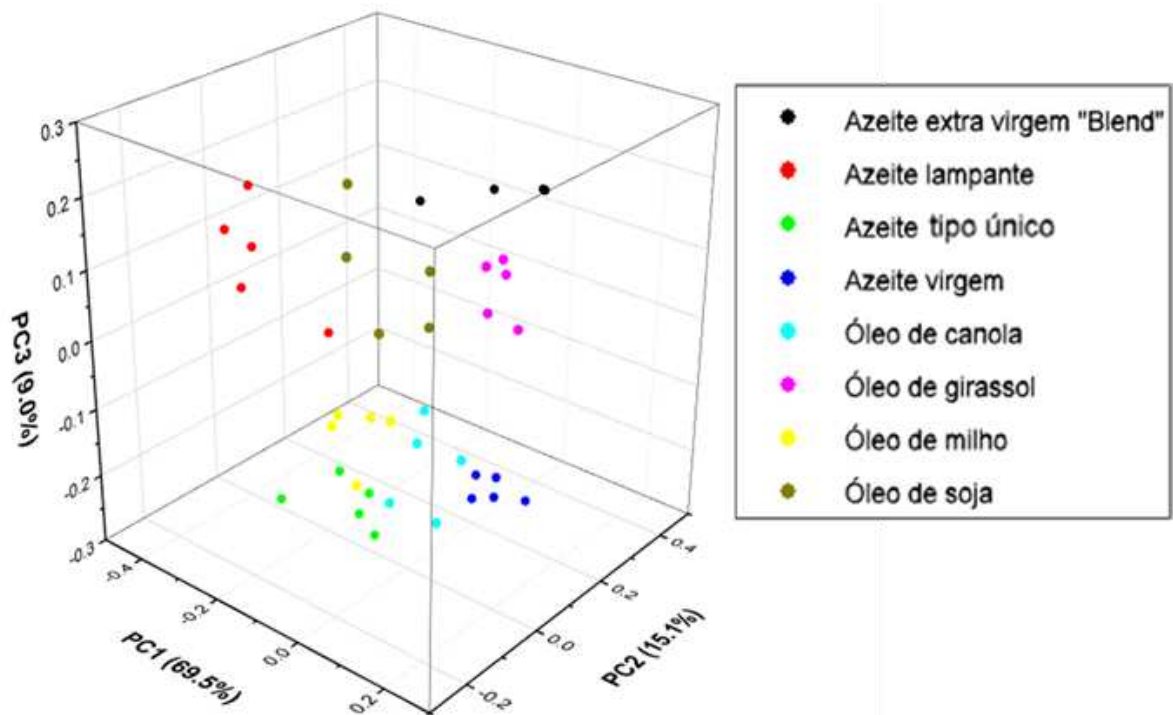
Os resultados apresentados na Figura 12 mostram que por FT-NIR foi possível diferenciar as adulterantes puros do azeite de oliva extra virgem, com a formação de seus grupos, com três PCs explicando 90,3% (64,2% + 16,8% + 9,3%) dos dados. Porém, a discriminação entre o azeite de oliva extra virgem e as misturas deste azeite com os adulterantes não foi evidente.

As diferenças nas composições do azeite extra virgem da variedade Arbequina e das amostras adulteradas não foram suficientes para que as amostras fossem discriminadas por FT-NIR. Os diferentes valores de absorvância obtidos para as misturas em relação ao azeite de oliva extra virgem geraram valores de *scores* diferentes e por isso as alterações foram interpretadas pelo modelo como gerados por amostras similares, o que dificultou a diferenciação entre as amostras. Desta forma, assim como para a variedade Koroneiki, a utilização de outros métodos quimiométricos pode fazer com o que o modelo seja capaz de discriminar melhor as amostras.

Para o *blend* das variedades, uma análise exploratória de dados pelo método PCA também foi realizada.

A Figura 13 apresenta o *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem do *blend* dos azeites puro e para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após aplicação do método PCA.

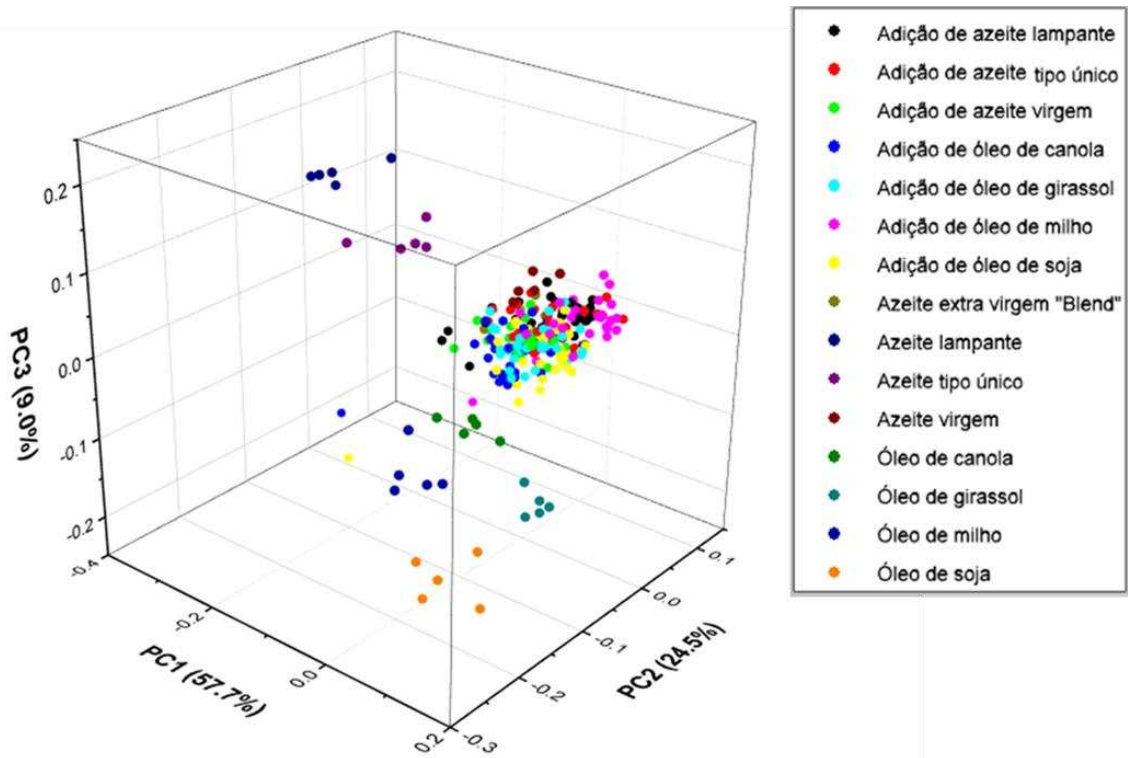
Figura 13 - *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem do *blend* dos azeites puro e para os adulterantes lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, após aplicação do método PCA.



Os resultados apresentados na Figura 13 mostram que as três PCs explicam 93,6% (69,5% + 15,1% + 9,0%) dos dados e estes estão com grande dispersão entre as replicatas. O *Score plot* indica ainda que houve a separação dos grupos referentes às amostras de azeite de oliva extra virgem do *blend* e aos adulterantes puros. Isso indica que a técnica FT-NIR foi capaz de diferenciar o azeite extra virgem do *blend* dos demais óleos, de forma semelhante às variedades Arbequina e Koroneiki.

A Figura 14 apresenta o *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem do *blend* puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.

Figura 14 - *Score plot* dos dados de FT-NIR para a amostra de azeite de oliva extra virgem do *blend* puro, para os adulterantes puros lampante, tipo único, virgem, óleo de canola, óleo de girassol, óleo de milho e óleo de soja, e para as misturas do azeite de oliva extra virgem com os adulterantes com diferentes porcentagens de adulteração, após aplicação do método PCA.



De forma semelhante aos estudos realizados para as variedades Arbequina e Koroneiki, as diferenças nas composições do azeite extra virgem do *blend* e das amostras adulteradas não foram suficientes para que as amostras fossem discriminadas pelo método PCA. Novamente, os resultados indicam que as alterações nos valores dos *scores* foram interpretadas pelo modelo como gerados por amostras similares, o que dificultou a diferenciação entre as amostras.

Os resultados obtidos na etapa de análise exploratória dos dados mostram que foram gerados modelos diferentes para as variedades Arbequina e Koroneiki e para o *blend*. Isso porque, além das diferenças existentes entre os azeites de oliva extra virgens e os adulterantes, existem diferenças na composição dos azeites de oliva extra virgens, devido a fatores como características do solo onde as oliveiras foram cultivadas, processo produtivo e pré-tratamento das azeitonas antes da

extração, os quais não foram avaliados nesse estudo. Isso mostra como é grande a variabilidade de informações que são obtidas por meio de análises de azeites de oliva por FT-NIR, o que dificulta a geração de um modelo que discrimine qualquer tipo de adulteração em qualquer tipo de azeite de oliva.

O método PCA permitiu uma análise exploratória dos dados, com uma diferenciação das amostras de azeite de oliva extra virgem dos adulterantes puros, mas ainda com pouca diferenciação das amostras adulteradas. Visando uma análise mais criteriosa dos dados e uma maior discriminação das amostras, os métodos PLS-DA e PLS foram aplicados para a construção de modelos de classificação e calibração multivariada das amostras de azeite de oliva.

2.4.2. Construção dos modelos de classificação e calibração

2.4.2.1. Desenvolvimento e validação de modelos para classificação supervisionada dos azeites de oliva varietais e *blend*

Após a análise preliminar exploratória dos dados utilizando PCA, diferentes modelos PLS-DA foram construídos para fins de classificação supervisionada.

Os valores calculados para as figuras de mérito para o modelo PLS-DA para classificar os azeites de oliva varietais e *blend*, utilizando os dados de FT-NIR para o modelo completo e para modelo com seleção de variáveis por meio do método OPS são apresentados na Tabela 2.

Tabela 2 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA, completo e pelo método OPS, para classificação das variedades Arbequina, Koroneiki e para o *blend*, utilizando os dados de FT-NIR.

	MODELO COMPLETO						OPS					
	TREINAMENTO						TREINAMENTO					
	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)
Arbequina	98,6	1,4	100,0	0,0	99,5	98,1	100,0	0,0	100,0	0,0	100,0	100,0
Koroneiki	99,7	0,3	98,4	0,0	98,8	98,5	99,7	0,3	100,0	0,0	99,9	99,5
<i>Blend</i>	96,9	3,1	99,1	0,0	98,4	95,2	100,0	0,0	99,8	0,2	99,9	99,9
	TESTE						TESTE					
Arbequina	100,0	0,0	100,0	0,0	100,0	100,0	100,0	0,0	100,0	0,0	100,0	100,0
Koroneiki	100,0	0,0	99,3	0,0	99,5	99,5	100,0	0,0	99,6	0,4	99,8	99,8
<i>Blend</i>	98,6	1,4	100,0	0,0	99,5	98,1	99,3	0,7	100,0	0,0	99,8	99,1

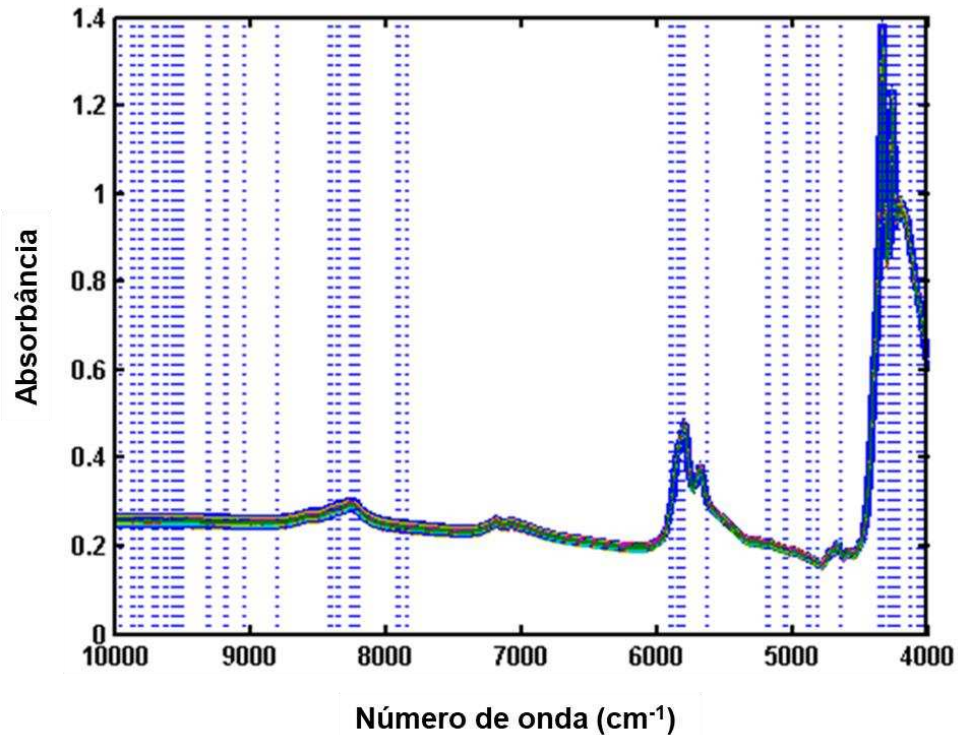
SEN: Sensibilidade; TFN: Taxa de Falsos Negativos; Esp: Especificidade; TFP: Taxa de Falsos Positivos; Exa: Exatidão; Con: Confiabilidade.

A Tabela 2 apresenta os valores obtidos para as figuras de mérito para o modelo completo e para o OPS, ambos construídos com 6 variáveis latentes. O OPS, que utilizou 120 variáveis (valor 10% menor que a quantidade original), teve resultados melhores que o modelo completo para ambas as variedades e para o *blend*. O modelo OPS apresentou TFP e TFN abaixo de 0,3% para o conjunto de treinamento, além de valores de Sen e Esp superiores a 99,7%, valores de Exa acima de 99,9% e Con superiores a 99,5%. Para o conjunto teste, os valores de Exa e de Con foram superiores a 99,1% e as TFP e TFN foram inferiores a 0,7%. Desta forma, com um número consideravelmente menor de variáveis, característica inerente ao OPS, este método proporcionou um aumento da capacidade preditiva do modelo.

As variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para classificação das variedades Arbequina, Koroneiki e para o *blend*,

utilizando os dados de FT-NIR, são mostradas na Figura 15. As variáveis selecionadas estão indicadas na figura pelos pontos azuis.

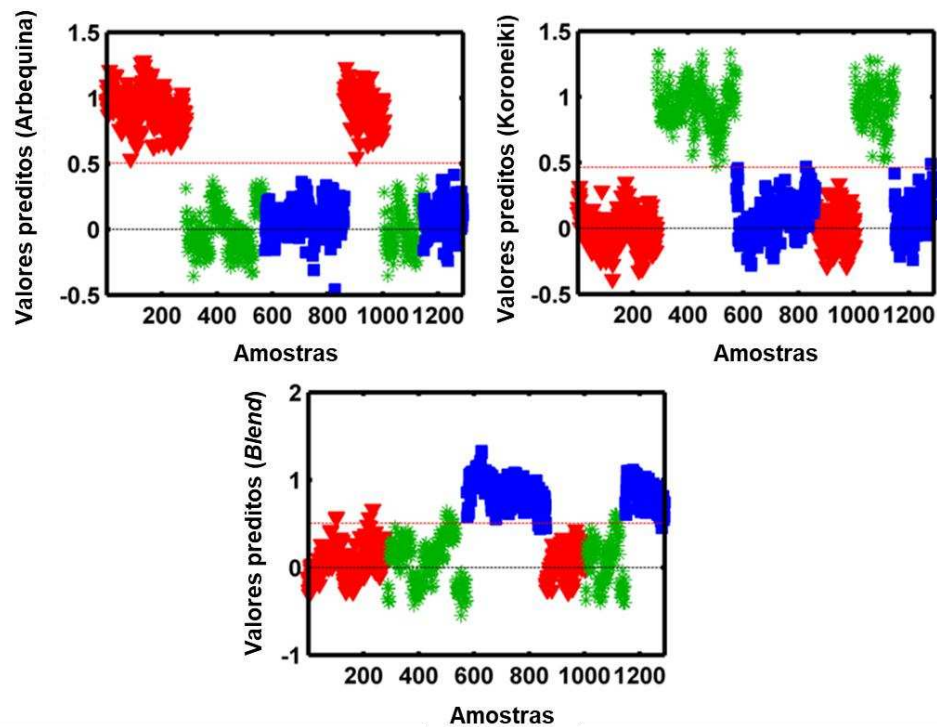
Figura 15 - Variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para classificação das variedades Arbequina, Koroniewski e para o *blend*, utilizando os dados de FT-NIR.



Analisando a Figura 15 pode-se notar que regiões espectrais relacionadas aos ácidos graxos (4664 cm^{-1} e 4788 cm^{-1}), aos triacilgliceróis (5796 cm^{-1} e 5680 cm^{-1}) e aos ácidos oleico e linoleico (8264 cm^{-1}) foram selecionadas. As variações nas intensidades de absorção destas regiões se relacionam às principais diferenças de composição entre o azeite de oliva extra virgem das diferentes variedades e do *blend*, justificando a seleção destas variáveis.

A Figura 16 ilustra a discriminação das amostras de cada classe (Arbequina, Koroniewski e *blend*) pelos valores preditos obtidos por meio da atribuição dos valores binários do método PLS-DA-OPS para cada amostra, utilizando os dados de FT-NIR.

Figura 16 – Discriminação das amostras de cada classe (Arbequina, Koroneiki e *blend*) pelos valores preditos obtidos por meio da atribuição dos valores binários do método PLS-DA-OPS para cada amostra, utilizando os dados de FT-NIR. Em vermelho, variedade Arbequina (Classe 1); em verde, variedade Koroneiki (Classe 2); em azul, *Blend* (Classe 3).



A Figura 16 mostra a discriminação das classes Arbequina, Koroneiki e *blend* pelo método PLS-DA. A atribuição dos valores binários (1 e 0) para as amostras de cada classe permitiu a diferenciação completa entre àquelas pertencentes a cada variedade e ao *blend*, o que pode ser observado pelos conjuntos formados por pontos de mesma cor.

2.4.2. Desenvolvimento e validação de modelos para classificação supervisionada dos adulterantes

Modelos PLS-DA para classificação supervisionada dos adulterantes foram construídos para as amostras adulteradas (azeite de oliva extra virgem com cada adulterante) para cada variedade e para o *blend* nas concentrações preparadas. Os resultados para os modelos (completos e com aplicação dos métodos de seleção de

variáveis) com melhores desempenhos são apresentados a seguir. Em todos os modelos, o melhor pré-processamento foi MSC seguido de centragem dos dados na média.

2.4.2.1.1. Variedade Koroneiki

Os valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas da variedade Koroneiki, com sete adulterantes são apresentados na Tabela 3.

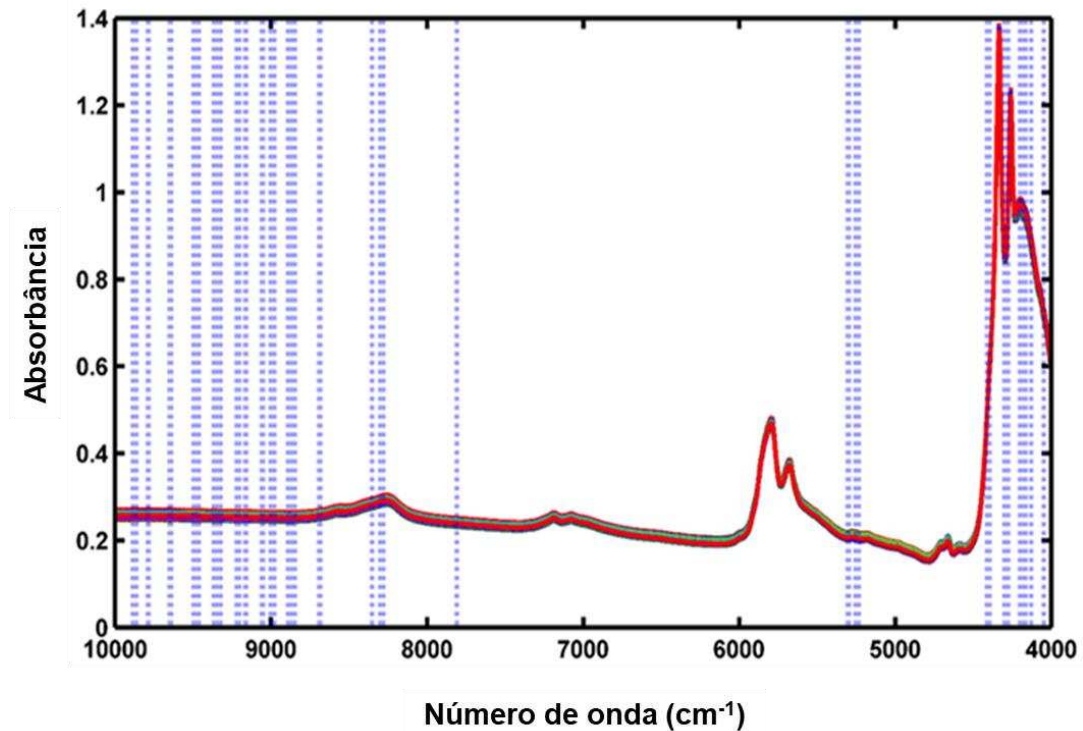
Tabela 3 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas da variedade Koroneiki, com sete adulterantes, utilizando os dados de FT-NIR.

	MODELO COMPLETO						OPS					
	TREINAMENTO						TREINAMENTO					
	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)
Azeite puro	100	0	99,6	0,4	99,6	99,6	100	0	100	0	100	100
Lampante	75	25	96,7	3,3	93,6	68,6	90	10	98,3	1,7	97,1	87,1
Virgem	92,5	7,5	97	3	96,3	88,8	95	5	98,7	1,3	98,2	93,2
Tipo único	98	3	99	1	99	96	97,5	2,5	99,1	0,9	98,9	96,4
Óleo de canola	87,5	12,5	98,7	1,3	97	84,5	95	5	99,2	0,8	98,6	93,6
Óleo de milho	90	10	99,1	0,9	97,8	87,8	92,5	7,5	99,2	0,8	98,2	90,7
Óleo de girassol	100	0	100	0	100	100	100	0	100	0	100	100
Óleo de soja	95	5	99,6	0,4	98,9	93,9	92,5	7,5	99,2	0,8	98,2	90,7
	TESTE						TESTE					
Azeite puro	100	0	99,2	0,8	99,2	99,2	100	0	100	0	100	100
Lampante	70	30	99,1	0,9	94,8	64,8	100	0	98,3	1,7	98,6	98,6
Virgem	100	0	96,4	3,6	97	97	100	0	98,3	1,7	98,6	98,6
Tipo único	95	5	99,1	0,9	98,5	93,5	100	0	99,2	0,8	99,3	99,3
Óleo de canola	85	15	99,1	0,9	97	82	94,7	5,3	100	0	99,3	94
Óleo de milho	75	25	99,1	0,9	95,5	70,5	85	15	99,2	0,8	97,2	82,2
Óleo de girassol	100	0	99,1	0,9	99,2	99,2	100	0	100	0	100	100
Óleo de soja	100	0	98,2	1,8	98,5	98,5	95	5	100	0	99,3	94,3

SEN: Sensibilidade; TFN: Taxa de Falsos Negativos; Esp: Especificidade; TFP: Taxa de Falsos Positivos; Exa: Exatidão; Con: Confiabilidade.

As variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para a variedade Koroneiki com os sete adulterantes são apresentadas na Figura 17.

Figura 17 - Variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para a variedade Koroneiki com os sete adulterantes.



A variedade Koroneiki é composta principalmente por álcoois alifáticos e o ácido linolênico. A seleção de variáveis para este modelo incluiu regiões espectrais em 4260 cm^{-1} and 4336 cm^{-1} , associadas à combinação de vibrações do modo de estiramento de ligações C–H, referentes às longas cadeias dos componentes. As bandas em torno de 5186 cm^{-1} são associadas à combinação da deformação e estiramento do primeira harmônico de ligações O-H (Harwood & Aparicio, 2000; Hourant *et al.*, 2000; Cozzolino *et al.*, 2005).

O modelo OPS, com 85 variáveis, apresentou valores de Esp superiores a 98,3% e valores de Exa acima de 97,2% para os conjuntos de treinamento. Para o conjunto teste, os valores de Exa foram superiores a 98,3%, com baixas TFP e TFN. Ainda, valores de Con maiores que 87,1% foram obtidos para o conjunto de treinamento, mostrando um ótimo desempenho do método.

O modelo utilizando o método OPS selecionando as variáveis para a variedade Koroneiki apresentou excelentes resultados para todas as classes, sendo considerado robusto e confiável.

2.4.2.1.2. Variedade Arbequina

A Tabela 4 apresenta os valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas da variedade Arbequina, com sete adulterantes.

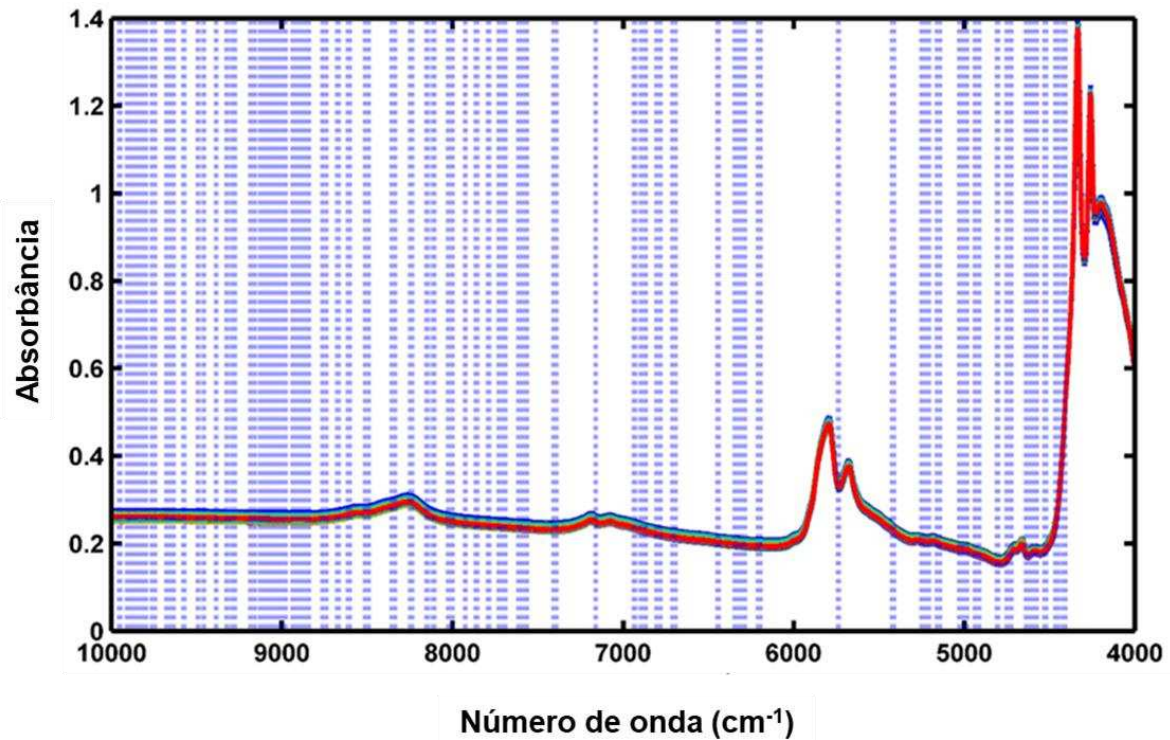
Tabela 4 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas da variedade Arbequina, com sete adulterantes, utilizando os dados de FT-NIR.

	MODELO COMPLETO						OPS					
	TREINAMENTO						TREINAMENTO					
	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)
Azeite puro	100	0	99,6	0,4	99,6	99,6	100	0	99,6	0,4	99,6	99,6
Lampante	90	10	99,1	0,9	97,6	87,6	95	5	97,7	2,3	97,3	92,3
Virgem	80	20	93,8	6,2	91,8	71,8	92,5	7,5	96	4	95,5	88
Tipo único	87,5	12,5	96,3	3,7	95	82,5	90	10	98,6	1,4	97,3	87,3
Óleo de canola	87,5	12,5	99,1	0,9	97,2	84,7	79,5	20,5	98,2	1,8	95,5	75
Óleo de milho	90	10	98,1	1,9	96,8	86,8	82,5	17,5	97,8	2,2	95,5	78
Óleo de girassol	87,5	12,5	97,7	2,3	96,1	83,6	95	5	99,1	0,9	98,4	93,4
Óleo de soja	72,5	27,5	99,1	0,9	95	67,5	85	15	98,2	1,8	96,2	81,2
	TESTE						TESTE					
Azeite puro	100	0	99,1	0,9	99,1	99,1	100	0	99,2	0,8	99,3	99,3
Lampante	95	5	98	2	97,5	92,5	85	15	99,2	0,8	97,1	82,1
Virgem	82,4	17,6	91,8	8,2	90,6	72,9	90	10	93,5	6,5	93,1	83,1
Tipo único	80	20	97,1	2,9	94,3	74,3	85	15	98	2	95,9	80,9
Óleo de canola	75	25	99	1	95	70	75	25	96,2	3,8	92,9	67,9
Óleo de milho	85	15	96,1	3,9	94,3	79,3	70	30	96,3	3,7	92,1	62,1
Óleo de girassol	85	15	97	3	95	80	100	0	98	2	98,3	98,3
Óleo de soja	73,7	26,3	96,2	3,8	92,7	66,4	65	35	98,1	1,9	92,9	57,9

SEN: Sensibilidade; TFN: Taxa de falsos negativos; Esp: Especificidade; TFP: Taxa de falsos positivos; Exa: Exatidão; Con: Confiabilidade.

As variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para a variedade Arbequina com os sete adulterantes são apresentadas na Figura 18.

Figura 18 - Variáveis selecionadas para a construção do modelo pelo método PLS-DA-OPS para classificação das variedades Arbequina, utilizando os dados de FT-NIR.



A variedade Arbequina possui principalmente esteróis e ácidos graxos em sua composição. A seleção de variáveis para este modelo incluiu regiões espectrais em 8264 cm^{-1} e 7188 cm^{-1} , características do ácido linoleico, e regiões em torno de 7502 cm^{-1} e 6102 cm^{-1} , características dos esteróis campesterol e estigmasterol, respectivamente (Harwood & Aparicio, 2000; Christy *et al.*, 2004; Holman & Edmondson, 1956; Özdemir *et al.*, 2018).

Para a variedade Arbequina, o modelo OPS (235 variáveis) apresentou valores de Exa superiores a 95,5% e valores de Esp acima de 96% para os conjuntos de treinamento. Para o conjunto teste, os valores de Exa foram superiores a 92,1%. Valores de Con maiores que 75% (conjunto de treinamento) obtidos no método PLS-DA-OPS para a variedade Arbequina mostram que este foi um melhor modelo em relação ao completo.

2.4.2.1.3. Blend

Os valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método GA, para classificação das amostras adulteradas do *blend*, com sete adulterantes são apresentados na Tabela 5.

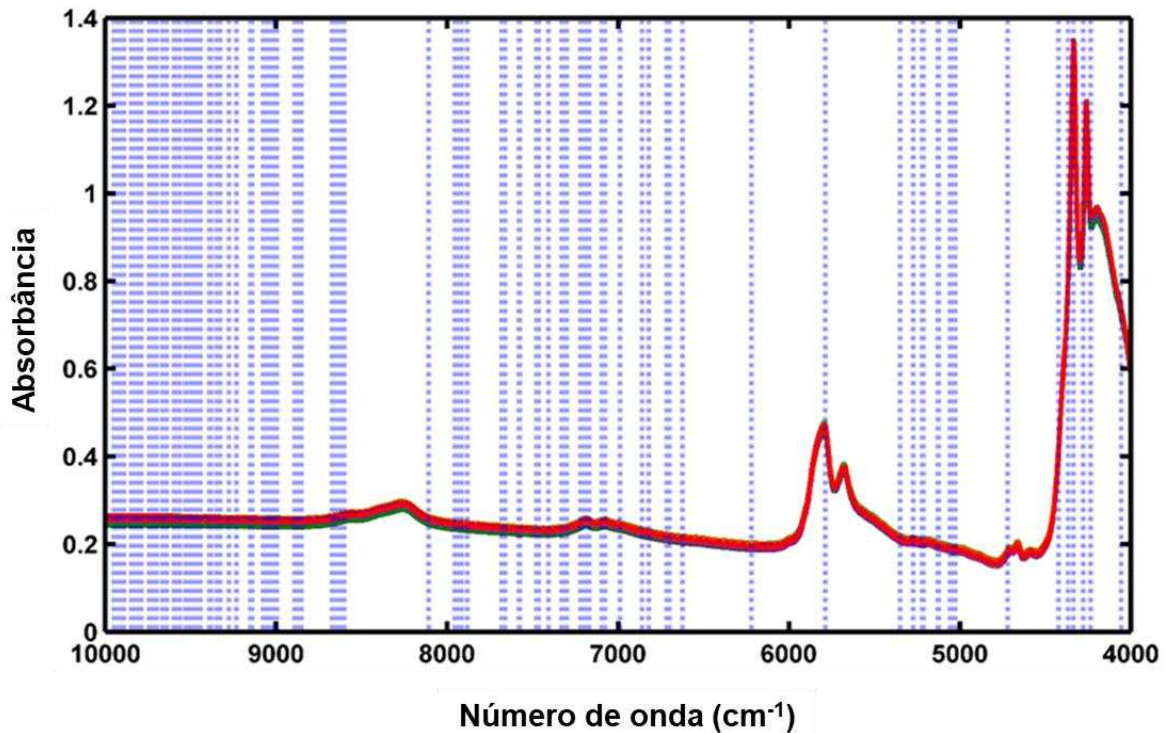
Tabela 5 - Valores calculados para as figuras de mérito referentes ao modelo PLS-DA completo e pelo método OPS, para classificação das amostras adulteradas do *blend*, com sete adulterantes, utilizando os dados de FT-NIR.

	MODELO COMPLETO						GA					
	TREINAMENTO						TREINAMENTO					
	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)	SEN (%)	TFN (%)	Esp (%)	TFP (%)	Exa (%)	Con (%)
Azeite puro	100	0	100	0	100	100	100,0	0,0	99,6	0,4	99,6	99,6
Lampante	82,5	17,5	99,1	0,9	96,7	79,2	95,0	5,0	99,2	0,8	98,6	93,6
Virgem	97,5	2,5	98,2	1,8	98,1	95,6	97,5	2,5	100,0	0,0	99,6	97,1
Tipo único	97,5	2,5	96,1	3,9	96,3	93,8	100,0	0,0	98,8	1,2	98,9	98,9
Óleo de canola	80	20	98,3	1,7	95,6	75,6	92,5	7,5	99,2	0,8	98,2	90,7
Óleo de milho	97,5	2,5	100	0	99,6	97,1	97,5	2,5	100,0	0,0	99,6	97,1
Óleo de girassol	97,5	2,5	97,4	2,6	97,4	94,9	100,0	0,0	99,6	0,4	99,6	99,6
Óleo de soja	85	15	100	0	97,8	82,8	95,0	5,0	100,0	0,0	99,3	94,3
	TESTE						TESTE					
Azeite puro	66,7	33,3	100	0	99,2	65,9	100,0	0,0	99,3	0,7	99,3	99,3
Lampante	90	10	99,1	0,9	97,8	87,8	100,0	0,0	100,0	0,0	100,0	100,0
Virgem	95	5	98,2	1,8	97,8	92,8	100,0	0,0	100,0	0,0	100,0	100,0
Tipo único	95	5	97,4	2,6	97	92	100,0	0,0	97,5	2,5	97,9	97,9
Óleo de canola	85	15	98,3	1,7	96,3	81,3	90,0	10,0	99,2	0,8	97,9	87,9
Óleo de milho	95	5	100	0	99,2	94,2	95,0	5,0	100,0	0,0	99,3	94,3
Óleo de girassol	100	0	98,2	1,8	98,5	98,5	95,0	5,0	100,0	0,0	99,3	94,3
Óleo de soja	85	15	99,1	0,9	97	82	95,0	5,0	100,0	0,0	99,3	94,3

SEN: Sensibilidade; TFN: Taxa de Falsos Negativos; Esp: Especificidade; TFP: Taxa de Falsos Positivos; Exa: Exatidão; Con: Confiabilidade.

A Figura 19 apresenta as variáveis selecionadas para a construção do modelo pelo método PLS-DA-GA para o *blend* com os sete adulterantes.

Figura 19 - Variáveis selecionadas para a construção do modelo pelo método PLS-DA-GA para classificação do *blend*, utilizando os dados de FT-NIR.



O *blend* consiste em uma mistura de 50% m/m para cada variedade. Assim, a seleção de variáveis para este modelo inclui regiões espectrais importantes para ambas as variedades.

Para o conjunto teste, o modelo GA (107 variáveis) apresentou valores de Exa e Esp superiores a 97,9% e a 97,5%, respectivamente. Ainda, apresentou valores de Esp superiores a 98,8% e valores de Exa acima de 98,2% para os conjuntos de treinamento. O modelo apresenta ainda baixos valores de TFN e TFP.

Selecionando as variáveis, o modelo PLS-DA-GA aplicado ao *blend* apresentou ótimos resultados para todas as classes, sendo considerado um melhor modelo em relação ao completo e bastante sensível e confiável.

2.4.2.2. Desenvolvimento e validação de modelos para calibração multivariada das amostras adulteradas

Modelos PLS foram construídos para calibração das amostras adulteradas. Além dos modelos específicos (varietais e *blend*), também foram construídos

modelos robustos, que incorporaram todas as amostras das variedades e do *blend*. Os resultados para os modelos (completos e com aplicação do método de seleção de variáveis OPS) com melhores desempenhos são apresentados a seguir. Em todos os modelos, o melhor pré-processamento foi MSC seguido de centragem dos dados na média.

2.4.2.2.1. Azeite lampante

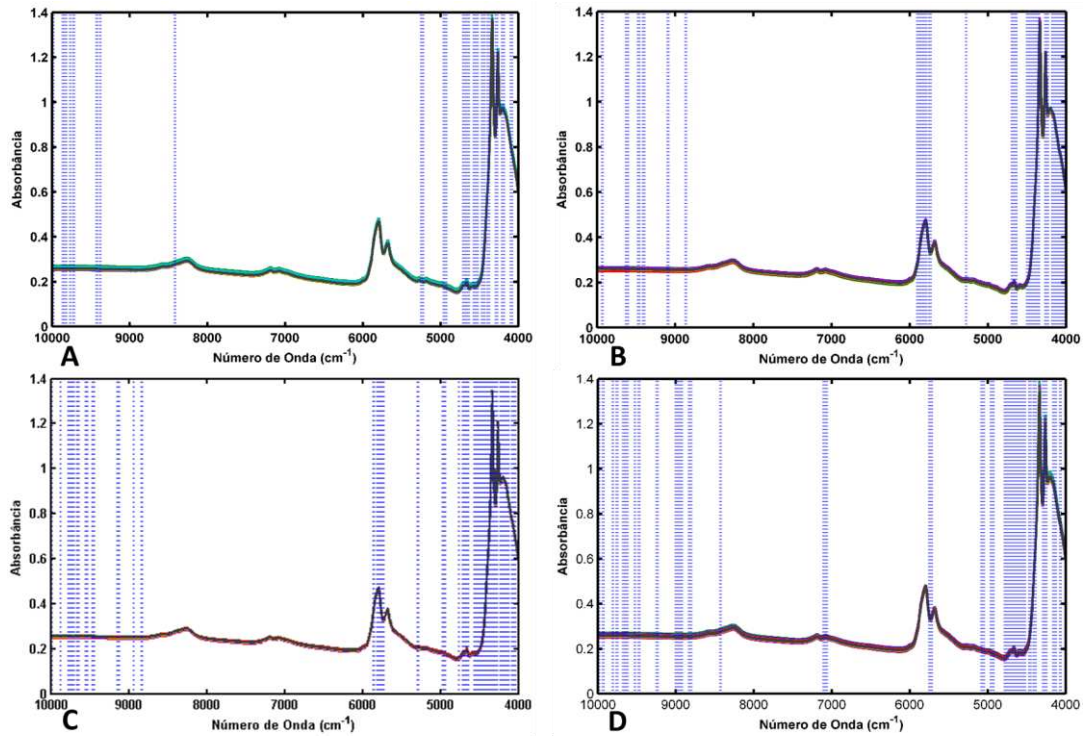
A Tabela 6 apresenta os parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante azeite lampante.

Tabela 6 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante azeite lampante.

	Arbequina		Kononeiki		Blend		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	6	6	6	6	5	5	7	7
nvars	1501	103	1501	125	1501	189	1501	167
RMSEC (%)	1.45	1.03	0.56	0.59	1.01	1.04	1.43	0.90
RMSECV (%)	4.18	2.13	2.30	1.15	4.40	2.18	2.29	1.40
RMSEP (%)	2.02	1.47	1.60	0.89	2.50	1.75	2.10	1.14
Rc	0.98	0.99	0.99	0.99	0.99	0.99	0.98	0.99
Rcv	0.82	0.96	0.95	0.99	0.80	0.96	0.95	0.98
Rp	0.95	0.98	0.98	0.99	0.97	0.98	0.96	0.99

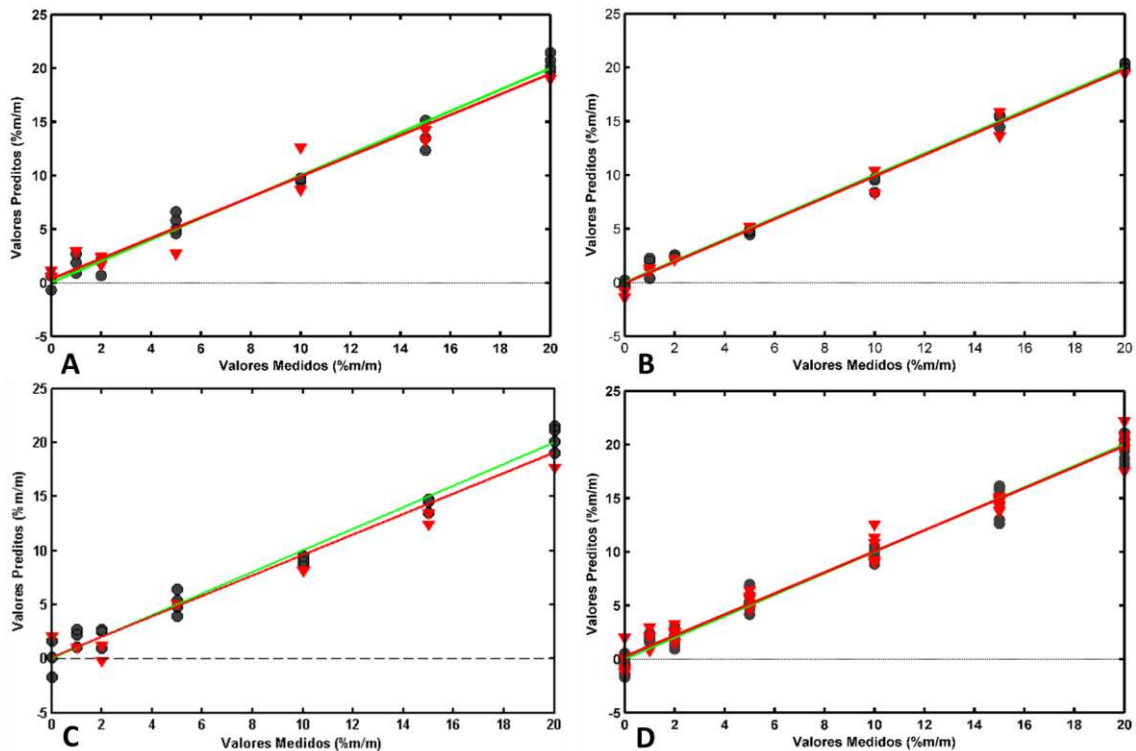
A Figura 20 apresenta as variáveis selecionadas para a construção dos modelos de calibração PLS pelo método OPS para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante azeite lampante.

Figura 20 - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante azeite lampante. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



A Figura 21 apresenta o gráfico de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante azeite lampante.

Figura 21 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite lampante. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



Para o adulterante azeite lampante, o modelo PLS-OPS para a variedade Koroneiki apresentou melhores valores dos coeficientes de correlação. Já o modelo robusto apresentou maior redução nos valores dos erros RMSEC, RMSECV e RMSEP em relação aos demais modelos.

2.4.2.2.2. Azeite virgem

A Tabela 7 apresenta os parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante azeite virgem.

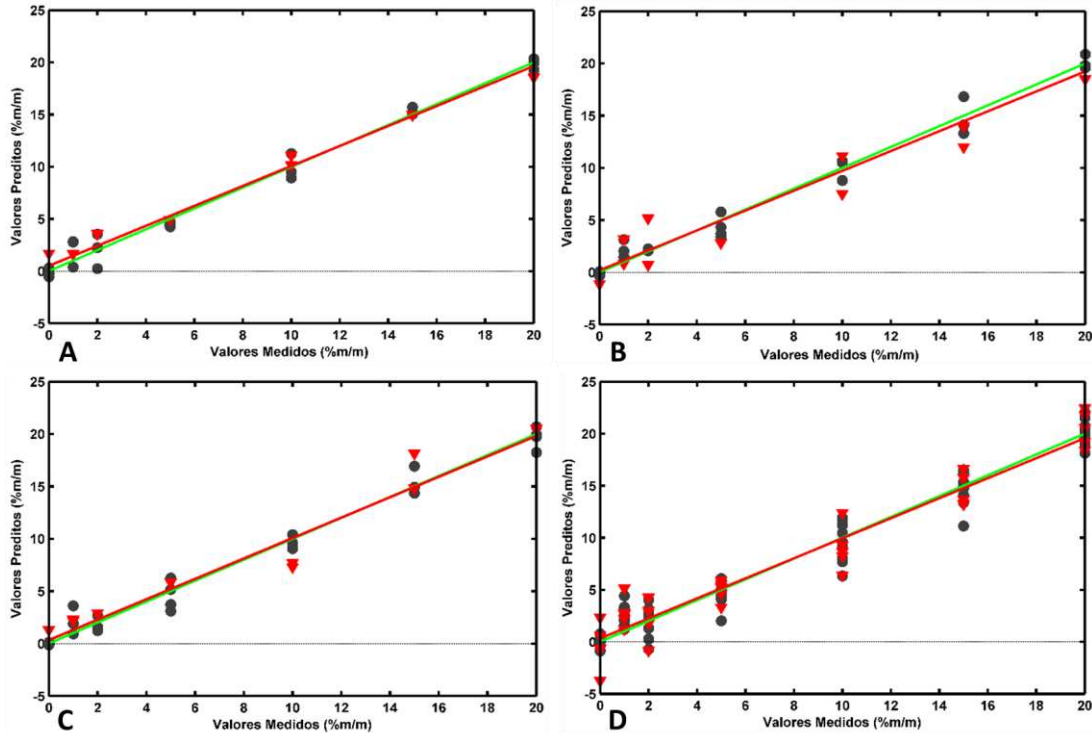
Tabela 7 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante azeite virgem.

	Arbequina		Koroneiki		Blend		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	6	5	6	5	5	5	8	8
nvars	1501	289	1501	219	1501	169	1501	145
RMSEC (%)	1.02	0.90	1.54	0.97	1.09	1.03	1.90	1.40
RMSECV (%)	3.41	2.14	4.73	2.29	5.00	2.35	3.40	2.17
RMSEP (%)	2.53	1.08	2.92	1.96	3.20	1.70	4.10	1.83
Rc	0.99	0.99	0.97	0.99	0.99	0.99	0.96	0.98
Rcv	0.89	0.96	0.77	0.96	0.75	0.95	0.88	0.95
Rp	0.95	0.99	0.91	0.96	0.90	0.97	0.84	0.97

A Figura 6A (Anexo I) apresenta as variáveis selecionadas para a construção dos modelos de calibração pelo método PLS-OPS para as variedades Arbequina, Koroneiki e *blend* e robusto para o adulterante azeite virgem.

A Figura 22 apresenta o gráfico de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo e pelo método OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante azeite virgem.

Figura 22 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite virgem. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



Menores valores dos erros de calibração, validação e predição foram obtidos pelos modelos da variedade Arbequina para o adulterante azeite virgem. Ainda, para o modelo robusto utilizou um menor número de variáveis após aplicação do método OPS.

2.4.2.2.3. Azeite tipo único

A Tabela 8 apresenta os parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante azeite tipo único.

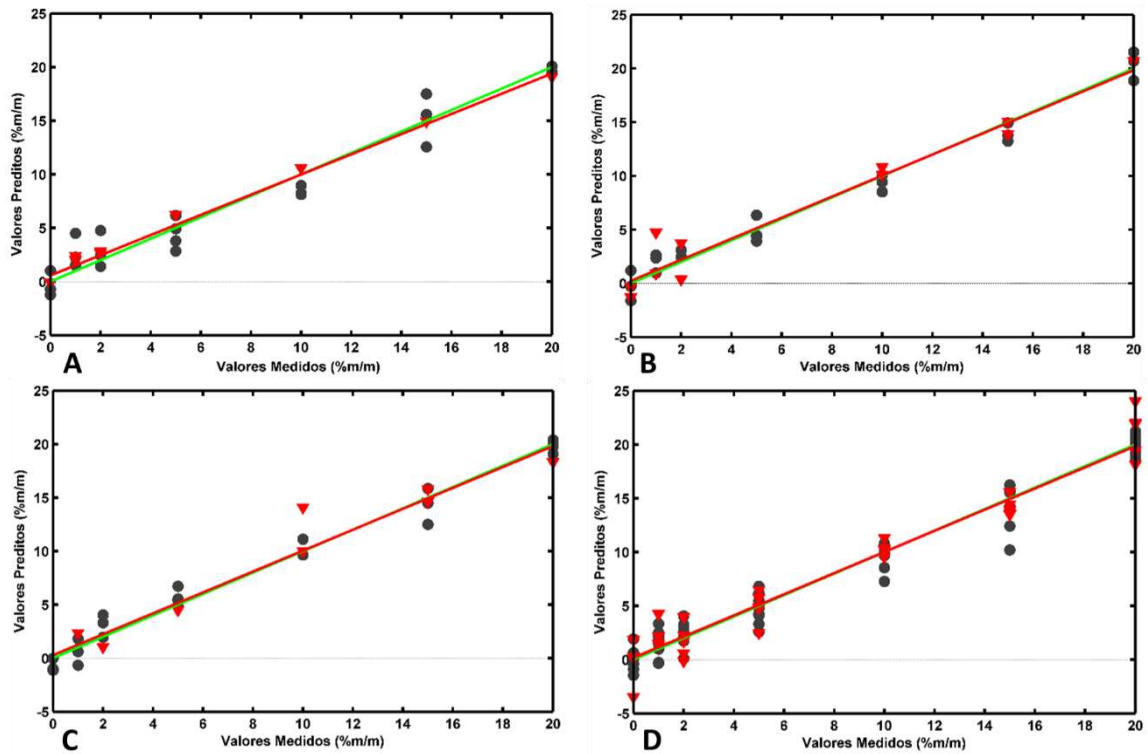
Tabela 8 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante azeite tipo único.

	Arbequina		Koroneiki		Blend		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	6	6	4	4	6	5	10	8
nvars	1501	55	1501	120	1501	90	1501	195
RMSEC (%)	1.15	1.56	1.40	1.07	0.80	1.04	1.26	1.24
RMSECV (%)	6.84	2.90	2.95	1.92	4.29	2.38	3.62	1.95
RMSEP (%)	4.10	0.84	2.42	1.47	3.65	1.70	2.83	1.61
Rc	0.98	0.97	0.98	0.99	0.99	0.99	0.98	0.98
Rcv	0.45	0.92	0.92	0.97	0.82	0.95	0.87	0.96
Rp	0.80	0.99	0.93	0.98	0.91	0.96	0.92	0.98

A Figura 7A (Anexo I) apresenta as variáveis selecionadas para a construção dos modelos de calibração pelo método PLS-OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante azeite tipo único.

A Figura 23 apresenta o gráfico de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante azeite tipo único.

Figura 23 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite tipo único. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



Para o adulterante azeite tipo único, a aplicação do método OPS no modelo da variedade Arbequina reduziu em 79,5% o valor de RMSEP. Já o modelo robusto apresentou maiores valores dos coeficientes de correlação após a seleção de variáveis.

2.4.2.2.4. Óleo de canola

A Tabela 9 apresenta os parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante óleo de canola.

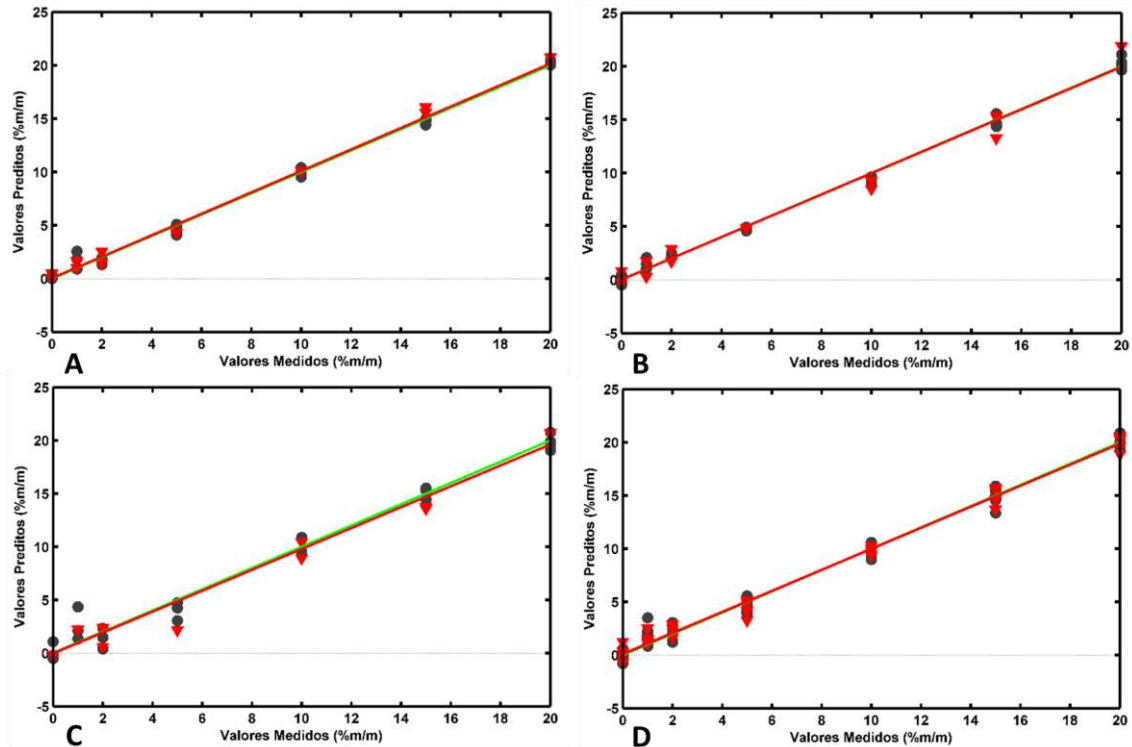
Tabela 9 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de canola.

	Arbequina		Koroneiki		Blend		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	4	4	4	4	3	3	5	5
nvars	1501	295	1501	85	1501	220	1501	150
RMSEC (%)	0.61	0.51	0.78	0.51	1.46	1.07	0.89	0.67
RMSECV (%)	1.00	0.83	1.32	1.04	1.98	1.69	1.06	0.93
RMSEP (%)	0.60	0.56	1.03	1.01	1.21	1.31	1.60	0.70
Rc	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99
Rcv	0.99	0.99	0.98	0.99	0.96	0.97	0.99	0.99
Rp	0.99	0.99	0.99	0.99	0.99	0.98	0.97	0.99

A Figura 8A (Anexo I) apresenta as variáveis selecionadas para a construção dos modelos de calibração pelo método PLS-OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de canola.

A Figura 24 apresenta o gráfico de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de canola.

Figura 24 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de canola. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



Para o adulterante óleo de canola destaca-se o resultado da aplicação do método OPS para o modelo da variedade Koroneiki, que reduziu o número de variáveis para menos de 5,66% das variáveis originais.

2.4.2.2.5. Óleo de girassol

A Tabela 10 apresenta os parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante óleo de girassol.

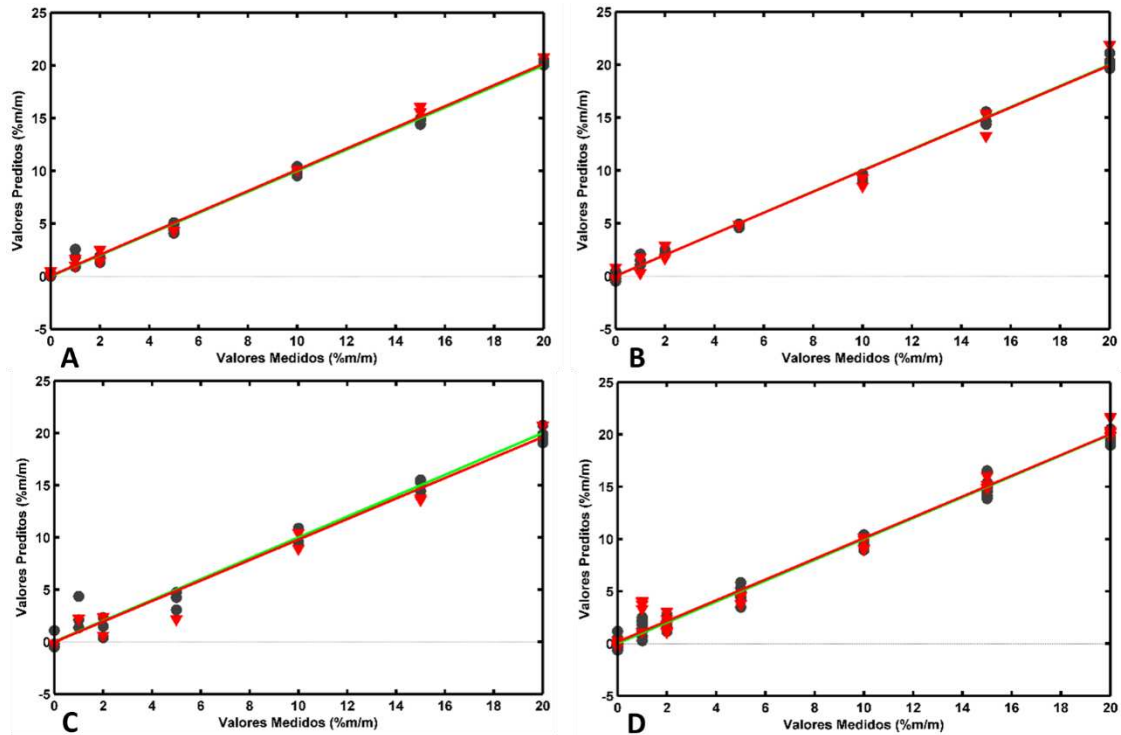
Tabela 10 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de girassol.

	Arbequina		Koroneiki		Blend		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	3	3	3	3	2	2	5	5
nvars	1501	25	1501	30	1501	70	1501	80
RMSEC (%)	0.57	0.50	0.78	0.47	1.23	1.11	0.71	0.63
RMSECV (%)	0.80	0.67	0.99	0.65	1.60	1.35	0.83	0.85
RMSEP (%)	0.88	0.71	0.98	0.50	1.20	1.02	1.18	1.16
Rc	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99
Rcv	0.99	0.99	0.99	0.99	0.97	0.98	0.99	0.99
Rp	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99

A Figura 9A (Anexo I) apresenta as variáveis selecionadas para a construção dos modelos de calibração pelo método PLS-OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de girassol.

A Figura 25 apresenta o gráfico de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de girassol.

Figura 25 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de girassol. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



Os modelos construídos para o adulterante óleo de girassol apresentaram, em geral, menores valores de RMSEC e RMSECV e, após aplicação do método OPS, o número de variáveis foram reduzidos a menos de 80.

2.4.2.2.6. Óleo de milho

Na Tabela 11 são apresentados os parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Kononeiki, *blend* e robusto para o adulterante óleo de milho.

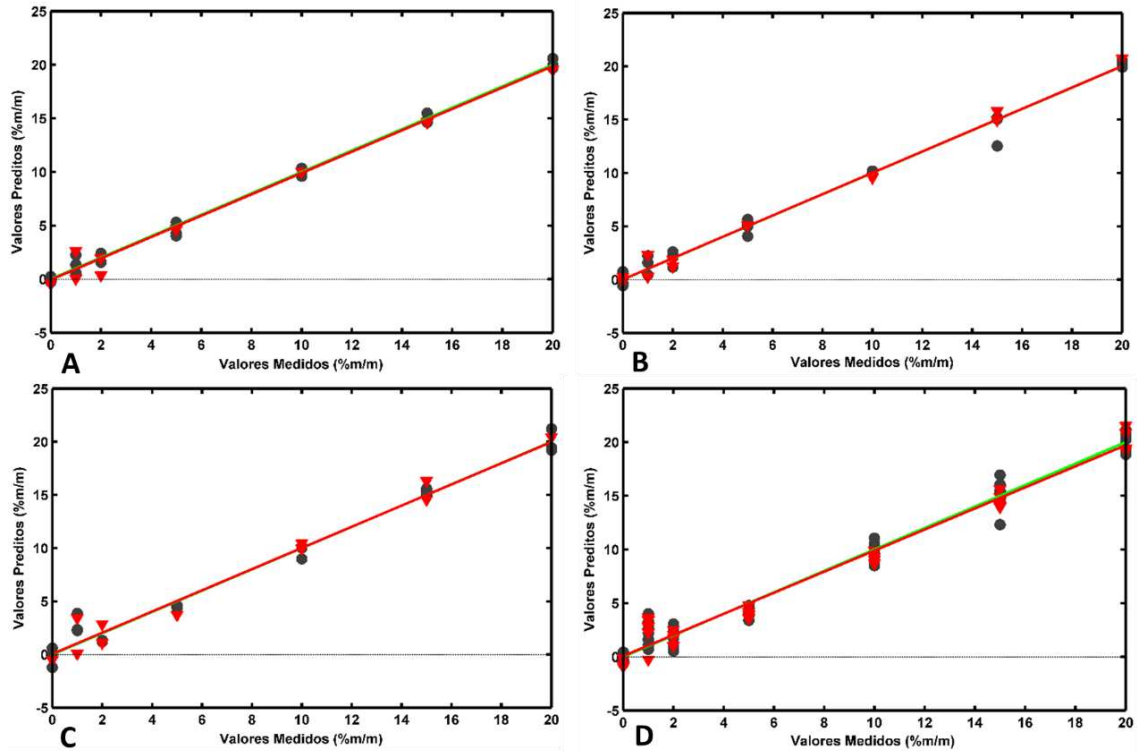
Tabela 11 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de milho.

	Arbequina		Koroneiki		Blend		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	4	4	4	4	3	3	4	4
nvars	1501	105	1501	280	1501	20	1501	210
RMSEC (%)	0.70	0.47	0.80	0.72	1.01	0.95	1.07	1.04
RMSECV (%)	1.33	1.00	1.48	1.32	1.41	1.32	1.18	1.15
RMSEP (%)	0.96	0.84	0.83	0.61	1.39	1.03	1.23	1.09
Rc	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Rcv	0.99	0.99	0.98	0.98	0.98	0.98	0.99	0.99
Rp	0.99	0.99	0.99	0.99	0.98	0.99	0.98	0.99

A Figura 10A (Anexo I) apresenta as variáveis selecionadas para a construção dos modelos de classificação pelo método PLS-OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de milho.

A Figura 26 apresenta o gráfico de valores preditos *versus* valores medidos (% m/m) para os modelos de classificação PLS completo e pelo método OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de milho.

Figura 26 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de milho. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



Para o adulterante óleo de milho destaca-se o resultado da aplicação do método OPS para o modelo do *blend*, que reduziu o número de variáveis para menos de 1,33% das variáveis originais.

2.4.2.2.7. Óleo de soja

Na Tabela 12 são apresentados os parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de soja.

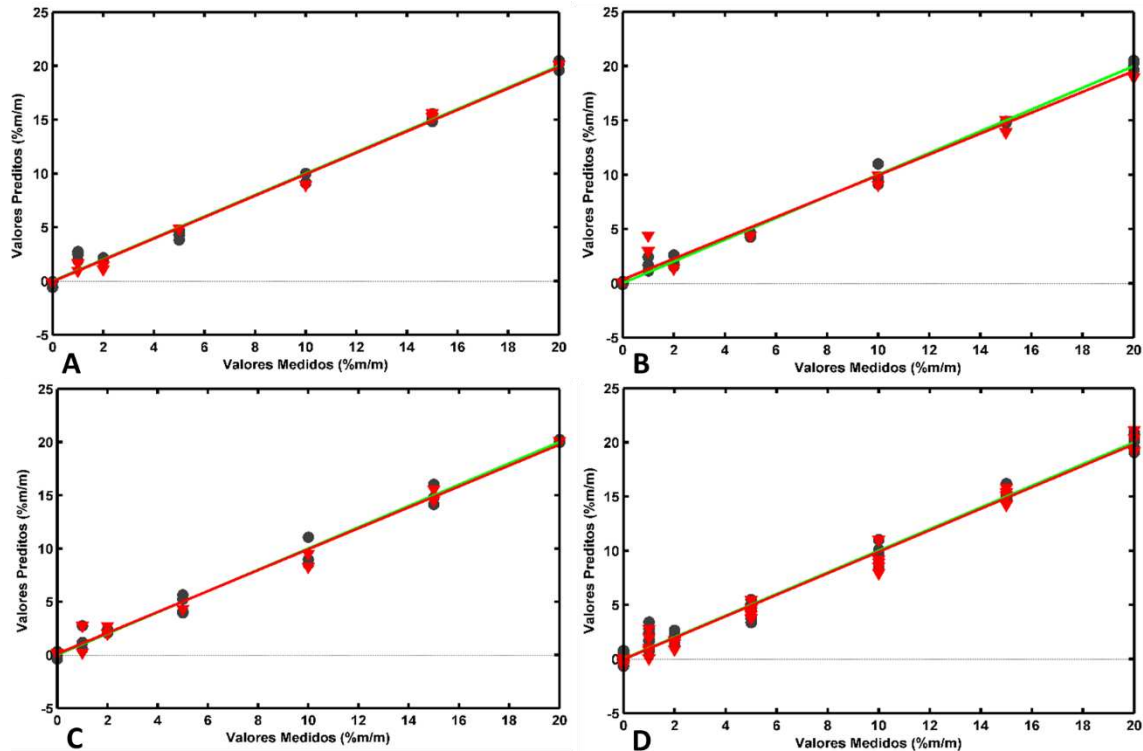
Tabela 12 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de soja.

	Arbequina		Koroneiki		Blend		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	3	3	3	3	4	3	5	5
nvars	1501	30	1501	750	1501	50	1501	125
RMSEC (%)	0.76	0.71	0.60	0.56	0.46	0.76	0.87	0.73
RMSECV (%)	1.09	0.87	0.78	0.73	0.83	1.06	1.02	0.91
RMSEP (%)	0.77	0.62	1.41	1.39	1.02	0.83	1.02	0.86
Rc	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Rcv	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Rp	0.99	0.99	0.98	0.98	0.99	0.99	0.99	0.99

A Figura 11A (Anexo I) apresenta as variáveis selecionadas para a construção dos modelos de calibração pelo método PLS-OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de milho.

Na Figura 27 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para as variedades Arbequina, Koroneiki, *blend* e robusto para o adulterante óleo de soja.

Figura 27 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de soja. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



Todos os modelos para o adulterante óleo de soja apresentaram coeficientes de correlação maiores que 0,98. Ainda, o modelo PLS-OPS para a variedade Arbequina foi construído com 2,0% do número de variáveis originais.

2.4.2.3. Discussão dos resultados

Os modelos PLS construídos para calibração das amostras adulteradas apresentaram bom desempenho. Em geral, os valores de RMSEC, RMSECV e RMSEP foram mais altos para os modelos robustos em relação aos demais modelos. Este resultado era esperado visto que os modelos robustos apresentam maior complexidade devido à quantidade de dados utilizados, sendo construídos com números maiores de nL_v (variando de 4 a 10 entre os adulterantes).

A seleção de variáveis pelo método OPS melhorou consideravelmente os modelos para todos os adulterantes. Quando comparados aos modelos robustos

completos, os modelos PLS-OPS reduziram para menos de 14% o número de variáveis em relação aos espectros originais. Os resultados satisfatórios obtidos após a seleção de variáveis indicam que as variáveis selecionadas para cada modelo representam os conjuntos de informações espectrais mais importantes para prever a composição das amostras adulteradas, eliminando muitas variáveis ruidosas e não informativas, e aumentando a capacidade preditiva dos modelos.

Os modelos robustos PLS-OPS apresentaram uma quantidade maior de variáveis selecionadas em relação aos modelos das variedades e do *blend* devido à grande complexidade dos dados utilizados, pois os modelos robustos abrangem os espectros de todas as amostras adulteradas e dos adulterantes puros. Em geral, o número de variáveis selecionadas para os modelos robustos dos adulterantes azeites virgem, lampante e tipo único foram maiores que para os demais adulterantes devido à semelhança de composição entre estes e o azeite extra virgem, sendo necessário um maior número de informações espectrais para distinção das amostras.

Para os modelos robustos dos adulterantes azeites virgem, lampante e tipo único foram selecionadas principalmente variáveis na região abaixo de 6000 cm^{-1} , cujas bandas são associadas principalmente aos triacilgliceróis, principais componentes dos azeites de oliva e a alguns compostos minoritários, como as ceras (ésteres de cadeia longa), cujos teores são diferentes entre o azeite de oliva extra virgem e os demais azeites, sendo estas variáveis responsáveis por permitir a diferenciação entre o azeite de oliva extra virgem e os adulterantes (Christy *et al.*, 2004; Sato *et al.*, 1991; Cozzolino *et al.*, 2005). As bandas na região acima de 7000 cm^{-1} foram as principais variáveis selecionadas nos modelos robustos para os demais adulterantes. Uma das principais diferenças entre o azeite de oliva e os óleos de sementes está na composição em ácidos graxos (Brasil, 2012), com regiões espectrais características em números de onda acima de 7000 cm^{-1} (Christy *et al.*, 2004; Holman & Edmondson, 1956).

Analisando os valores de RMSEP para os modelos robustos vê-se que eles foram reduzidos após a seleção de variáveis. Para o modelo do adulterante óleo de canola a redução foi de 56% (de 1,60% para 0,70% m/m); para o adulterante azeite virgem, a redução foi de 55% (de 4,10% para 1,83% m/m); o modelo PLS-OPS para

o adulterante azeite lampante teve uma redução de 46% (de 2,10% para 1,14% m/m) no valor de RMSEP. Os valores de RMSEC e RMSECV também foram significativamente menores e os coeficientes de correlação foram maiores, em geral, para os modelos PLS-OPS para todos os adulterantes.

A linearidade dos modelos de calibração pode ser analisada por meio dos gráficos dos valores preditos *versus* valores medidos (Figuras 14 a 20). Para os modelos PLS-OPS robustos, todos os valores de coeficientes de correlação foram superiores a 0,95, mostrando uma boa linearidade em toda a faixa de 1 a 20% m/m de adulteração.

Trabalhos encontrados na literatura utilizam a calibração multivariada para verificação da autenticidade de azeite de oliva. Oussama e colaboradores (2012) construíram um modelo PLS com variáveis pré-selecionadas para calibração de amostras de azeite de oliva extra virgem adulteradas com óleo de soja e óleo de girassol na faixa de 1 a 24% m/m de adulteração, utilizando dados de Espectroscopia no Infravermelho com Transformada de Fourier (FT-IR). O modelo apresentou baixo valor de RMSECV e foi capaz de identificar e classificar corretamente 100% de novas amostras adicionadas. Ainda, Jiang & Chen (2019) utilizaram dados de FT-NIR para verificação da autenticidade de azeite de oliva adulterado com óleo de amendoim, óleo de girassol, óleo de soja, óleo de milho e óleo de gergelim na faixa de 2,5 a 50% m/m. O modelo robusto PLS foi construído aplicando o método de seleção de variáveis BOSS, reduzindo o número de variáveis a menos de 10%, resultando em um modelo com grande capacidade preditiva.

2.5. CONCLUSÕES

Neste capítulo modelos robustos utilizando FT-NIR foram construídos para identificação de adulterações em azeite de oliva.

Modelos PLS-DA foram otimizados utilizando uma etapa anterior de seleção de variáveis, sendo essencial para eliminar informações redundantes e ruidosas, aumentando a robustez, a simplicidade e a precisão das classificações. Utilizando o método de seleção de variáveis OPS, um modelo permitiu discriminar amostras adulteradas de azeite de oliva extra virgem genuíno de duas diferentes variedades e o *blend* destas variedades. Ainda, modelos PLS-DA com aplicação do método OPS

foram construídos com confiabilidade, permitindo a diferenciação dos sete adulterantes em baixas concentrações para cada variedade (Arbequina e Koroneiki) de azeite de oliva extra virgem genuínos, com altos valores de especificidade e sensibilidade. Para o *blend* das variedades, o melhor modelo utilizou o método de seleção de variáveis GA, obtendo excelentes valores para todos os parâmetros de validação.

Modelos robustos para calibração das amostras adulteradas foram construídos com bom desempenho. Para os sete adulterantes, os modelos robustos PLS-OPS construídos apresentaram baixos valores de RMSEC, RMSECV e RMSEP e boa linearidade, caracterizada por coeficientes de correlação maiores que 0,95, mostrando a boa capacidade preditiva dos modelos mesmo quando os adulterantes se encontram em baixas concentrações.

***CAPÍTULO 3 – VALIDAÇÃO DE MÉTODO PARA
DETERMINAÇÃO DO TEOR DE ESQUALENO EM
AZEITE DE OLIVA POR GC/MS-MS***

3.1. INTRODUÇÃO

O consumo de azeite de oliva tem crescido em todo o mundo devido ao efeito nutricional da sua composição, relacionado principalmente à presença de ácidos graxos como o ácido oleico e a compostos com importantes efeitos antioxidantes, como o esqualeno (Sagrati *et al.*, 2013).

O esqualeno é um triterpeno com fórmula molecular $C_{30}H_{50}$, principal componente da fração de hidrocarbonetos do azeite de oliva. Devido aos seus benefícios para a saúde, o esqualeno tem recebido destaque na literatura, sendo considerado o responsável pelos efeitos do azeite de oliva contra certos tipos de câncer (Salvo *et al.*, 2017). Ainda, o esqualeno tem sido indicado como um componente importante para a manutenção da estabilidade do azeite de oliva (Sagrati *et al.*, 2013).

O azeite de oliva é um dos alimentos mais ricos em esqualeno, com teores variando de 1,0 a 12,0 g kg⁻¹. Em outros óleos, como o óleo de soja e o óleo de canola, o teor de esqualeno pode chegar a ser 25 vezes menor do que no azeite de oliva. Ainda, etapas de refino do azeite de oliva, que produzem, por exemplo, o azeite tipo único, causam a redução do teor de esqualeno do azeite. Desta forma, a determinação do teor de esqualeno pode ser usada como ferramenta para identificação de adulterações pela adição de óleos de sementes ou de azeites de outras categorias no azeite de oliva extra virgem (Cayuela & García, 2018; Pacetti *et al.*, 2019).

O método oficial para determinação do teor de esqualeno em azeite de oliva envolve etapas de fracionamento e extrações em colunas de sílica, demandando grande quantidade de solvente e fase estacionária por amostra (AOAC, 1999). Assim, desenvolver métodos mais rápidos que permitam a determinação de esqualeno em azeite de oliva com menor consumo de solventes e insumos traria uma grande contribuição para a determinação da autenticidade de azeites.

O objetivo deste capítulo foi validar um método para determinação do teor de esqualeno em azeite de oliva utilizando a extração em fase sólida e a cromatografia gasosa acoplada à espectrometria de massas sequencial, e utilizar o método validado para avaliar a autenticidade de azeites de oliva extra virgens pelo teor de esqualeno.

3.2. OBJETIVOS

3.2.1. Objetivo geral

Validar um método para determinação de esqualeno em azeite de oliva, utilizando a cromatografia gasosa acoplada à espectrometria de massas sequencial.

3.2.2. Objetivos específicos

- Utilizar a extração em fase sólida e a cromatografia gasosa acoplada à espectrometria de massas sequencial para determinar o teor de esqualeno em azeite de oliva;
- Validar o método adaptado de IOC, 2003 para determinação de esqualeno, utilizando os parâmetros de desempenho linearidade, exatidão/recuperação, precisão (repetitividade e precisão intermediária), Limite de Detecção (LD) e Limite de Quantificação (LQ) (INMETRO, 2016);
- Utilizar o método validado para determinar o teor de esqualeno de amostras do *blend* de azeites de oliva extra virgem adulteradas e de amostras de azeite de oliva coletadas para fiscalização, realizada pelo IQA/LFDA-MG.

3.3. MATERIAL E MÉTODOS

3.3.1. Amostras

Amostras do *blend* dos azeites de oliva extra virgem descritas no item 2.3.1 (Capítulo 2) foram utilizadas para a validação do método. As amostras do *blend* dos azeites de oliva extra virgem adulteradas (210 amostras) (Capítulo 2) foram analisadas utilizando o método validado para determinação do teor de esqualeno destas amostras. Ainda, os teores de esqualeno de amostras de azeites de oliva coletadas para fiscalização, realizada pelo IQA/LFDA-MG, foram determinados pelo método validado. As amostras de fiscalização foram coletadas em diferentes localidades no Brasil.

3.3.2. Padrões

Foi utilizada uma solução padrão de esqualeno (Sigma-Aldrich, $\geq 98\%$) com concentração de $1,0 \text{ mg mL}^{-1}$ em hexano, para as análises por GC-MS/MS.

3.3.2. Extração em fase sólida (SPE)

Para o método de extração, 40 mg de amostra de azeite de oliva foram adicionados em cartuchos de sílica Bond II (200 mg, 3 mL), dissolvidos em cerca de 2 mL de hexano. Um volume de 10 mL de hexano percolou por cada cartucho para eluição do esqualeno, com fluxo de $2,0 \text{ mL min}^{-1}$. Todo volume de 10 mL foi recolhido em tubos de ensaio, em sistema para realização da extração em fase sólida com recipiente de vidro com *manifold* (Figura 5A – Anexo I). O eluato foi evaporado com fluxo de nitrogênio, recuperado em 2 mL de hexano e analisado por GC-MS/MS.

A eluição foi realizada sob vácuo utilizando um recipiente de vidro com *manifold*, onde foram coletadas as frações eluídas dos cartuchos.

3.3.3. Validação do método

Para a validação do método os seguintes parâmetros de desempenho foram avaliados: linearidade, exatidão/recuperação, precisão (repetitividade e precisão intermediária), LD e LQ.

3.3.3.1. Linearidade

A linearidade foi avaliada pela curva analítica obtida pela análise de soluções padrões com concentrações de esqualeno variando de $1,0 - 8,0 \text{ ng mL}^{-1}$ (sete pontos em triplicata). A relação matemática entre o sinal analítico (área do pico) e a concentração de esqualeno foi expressa pela equação da curva analítica e o seu Coeficiente de Determinação (R^2).

3.3.3.2. Exatidão/Recuperação

Para a determinação da exatidão do método, 100 μL de esqualeno foram adicionadas à amostra do *blend* de azeites de oliva contendo $3,2 \text{ g kg}^{-1}$ de esqualeno, em três níveis de concentrações, para avaliação da recuperação. Os

resultados foram expressos em porcentagem de recuperação e foram considerados adequados com valores entre 80% e 110% (MAPA, 2015).

3.3.3.3. Precisão (repetitividade e precisão intermediária)

A repetitividade foi avaliada por meio da análise de seis replicatas de amostras do *blend* dos azeites de oliva extra virgem, no mesmo dia, em três níveis: 0,5 g kg⁻¹, 2,0 g kg⁻¹ e 4,0 g kg⁻¹. Para avaliar a precisão intermediária, seis replicatas para três níveis (0,5 g kg⁻¹, 2,0 g kg⁻¹ e 4,0 g kg⁻¹) foram preparadas e analisadas em três dias consecutivos. Os resultados foram expressos em porcentagem do desvio padrão relativo dos teores de esqualeno encontrados nas replicatas e foram considerados adequados com valores menores que 10% (MAPA, 2015).

3.3.3.4. Limite de Detecção (LD) e Limite de Quantificação (LQ)

Os limites de detecção e de quantificação foram calculados pela relação entre a estimativa do desvio padrão da resposta do branco e a inclinação (coeficiente angular) da curva analítica. O valor resultante foi multiplicado por 3,3 para obtenção do LD e por 10 para obtenção do LQ (INMETRO, 2016).

3.3.4. Análise de amostras de azeite de oliva adulteradas

As amostras do *blend* dos azeites de oliva extra virgem adulteradas descritas foram analisadas para determinação do teor de esqualeno (em g kg⁻¹) destas amostras.

3.3.5. Análise de amostras de azeite de oliva coletadas para fiscalização

O método validado foi utilizado para determinar os teores de esqualeno de amostras de azeites de oliva coletadas para fiscalização, pelo IQA/LFDA-MG. As amostras foram identificadas com os códigos AM01 a AM10. Essas amostras foram analisadas no IQA/LFDA-MG por outros métodos que classificam as amostras de azeites de oliva como extra virgem genuíno ou não genuíno, e os resultados encontrados foram comparados ao resultados dos teores de esqualeno (em g kg⁻¹) destas amostras.

3.3.6. Equipamentos

Para as análises cromatográficas foi utilizado um cromatógrafo a gás Agilent 7890B, equipado com coluna cromatográfica Rtx® – 5MS - 5% difenil/95% dimetil polissiloxano (30 m x 0,25 mm d.i. x 0,25 µm), com injetor do tipo *Multimode Inlet* (MMI), operando com temperatura constante de 350 °C, com modo de injeção *Splitless*. Hélio foi utilizado como fase móvel, com fluxo de 1,0 mL min⁻¹. Acoplado ao cromatógrafo foi utilizado um espectrômetro de massas do tipo triplo quadrupolo Agilent 7000C, com fonte de ionização por elétrons, operando com temperatura de 350 °C e com 70 eV de energia.

O esqualeno foi analisado no modo *Multiple Reaction Monitoring* (MRM), com as transições de Razão massa/carga (m/z) 367,2 > 299,2 e m/z 367,2 > 149,1 e energia de colisão de 5 eV. As integrações foram realizadas utilizando o software MassHunter Qualitative Workstation (Agilent Technologies, Palo Alto, CA). O esqualeno foi identificado por comparação do tempo de retenção do sinal analítico presente nos extratos das amostras com o tempo de retenção do padrão analítico e pela razão entre os íons de cada transição (intensidade relativa ± 20% em relação ao pico base).

3.4. RESULTADOS E DISCUSSÃO

3.4.1. Análises iniciais

As condições utilizadas foram adaptadas do método oficial utilizado pelo IOC para determinação de ceras em azeite de oliva (IOC, 2003). O método descreve uma primeira etapa de extração que consiste na retirada dos hidrocarbonetos da amostra de azeite de oliva, pois são interferentes na análise das ceras. Como o esqualeno pertence a essa classe, essa etapa do método oficial seria um referência para uma técnica de extração de esqualeno em amostras de azeite de oliva.

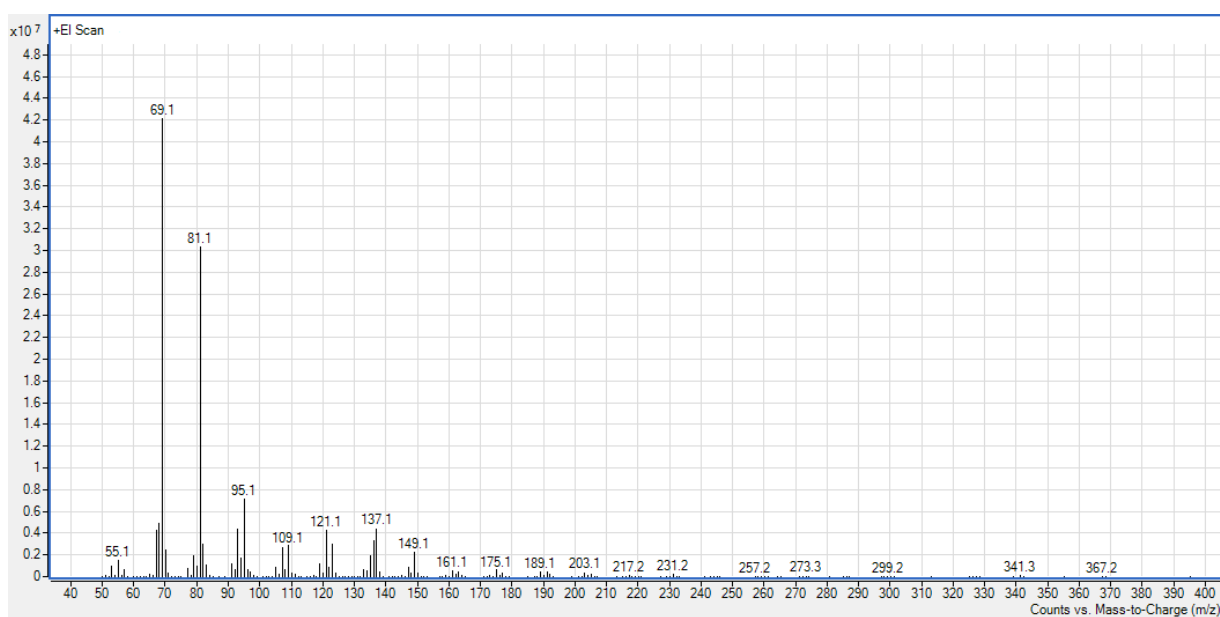
O método oficial utiliza sílica gel (15 g) e 500 mL de hexano para fracionar 500 mg de amostra. Visando utilizar uma menor quantidade de amostra e reduzir o volume de solvente utilizado, foi escolhida a extração em fase sólida em cartuchos de sílica para miniaturizar o processo de extração. Então, o método utiliza 40 mg de

amostra de azeite de oliva e um volume de 10 mL de hexano para extração de esqualeno.

As condições cromatográficas utilizadas pelo método também foram adaptadas do método oficial (IOC, 2003). Visando aumentar a seletividade na identificação e quantificação do esqualeno, foi utilizado um espectrômetro de massas sequencial acoplado ao cromatógrafo à gás, operando com modo de aquisição MRM, com duas transições.

Para definir as duas transições ideais para identificação e quantificação do esqualeno, uma solução padrão de esqualeno na concentração de 0,1 mg mL⁻¹ foi injetada no GC-MS/MS no modo de aquisição *Scan*, para determinar o perfil de fragmentação do analito por meio do espectro de massas (Figura 28).

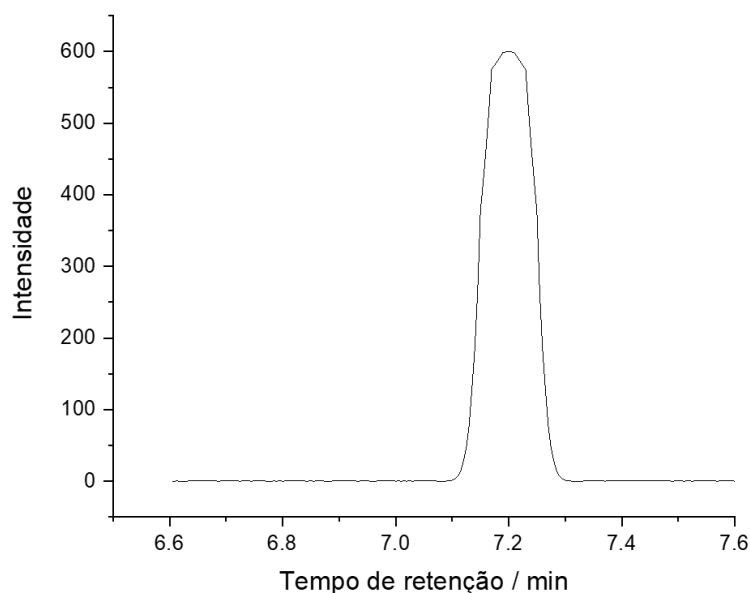
Figura 28 - Espectro de massas do esqualeno obtido pelo modo de aquisição *Scan*.



O íon com m/z 367,2 foi escolhido como íon precursor. O íons mais intensos do espectro de massas estavam na região abaixo da razão m/z 100, porém muitos compostos apresentam uma intensidade grande de íons nessa região. Desta forma, para uma melhor identificação do esqualeno, foram escolhidos os íons com m/z 299,2 e m/z 149,1 como íons produto.

Utilizar o MRM faz com que o sinal analítico do composto de interesse seja mais intenso e diminui a possibilidade de haver uma resposta analítica de algum interferente no tempo de retenção do esqualeno. Isso pode ser observado na Figura 29, onde é apresentado o cromatograma para a amostra do *blend* dos azeites de oliva contendo esqualeno, obtido pelo modo de aquisição MRM, com as duas transições m/z 367,2 > 299,2 e m/z 367,2 > 149,1. O tempo de retenção do esqualeno foi de 7,2 minutos.

Figura 29 - Cromatograma para a amostra do *blend* dos azeites de oliva contendo esqualeno, obtido pelo modo de aquisição MRM, com as transições 367,2 > 299,2 e 367,2 > 149,1, com tempo de retenção de 7,2 minutos.



O cromatograma mostra que o sinal analítico para o esqualeno é intenso e que não há nenhum interferente com resposta no tempo de retenção deste analito. Desta forma, as condições cromatográficas e espectrométricas foram consideradas adequadas e utilizadas para a validação do método.

3.4.2. Validação do método

A curva analítica foi obtida por regressão linear, expressa pela equação $c = 269,75A + 4371,1$. Foram utilizados sete níveis de concentração, obtendo um R^2 de

0,9926, mostrando uma boa linearidade do método (INMETRO, 2016). Os LD e LQ do método, estimados pelos parâmetros da curva analítica, foram iguais a 0,07 e 0,22 g kg⁻¹, respectivamente.

As recuperações foram de 95 ± 15% para nível de 0,5 g kg⁻¹, 102 ± 12% para o nível de 2,0 g kg⁻¹ e 104 ± 10% para o nível de 4,0 g kg⁻¹. Os valores de referência indicam uma recuperação adequada entre 80 e 110%, com coeficientes de variação menores que 10% (MAPA, 2015). Os valores obtidos atendem a estas condições, portanto a recuperação do método foi considerada adequada.

A repetitividade do método foi avaliada nos níveis 0,5 g kg⁻¹, 2,0 g kg⁻¹ e 4,0 g kg⁻¹. Os valores obtidos para o desvio padrão relativo foram de 5,4%, 6,5% e 5,6%, respectivamente. A precisão intermediária também foi avaliada pelo desvio padrão relativo, obtido após análises realizadas em três dias consecutivos. Seis amostras de cada nível de concentração (0,5 g kg⁻¹, 2,0 g kg⁻¹ e 4,0 g kg⁻¹) foram analisadas pelo mesmo analista, apresentando coeficientes de variação de 8,7%, 6,5% e 5,9%, respectivamente, indicando que o método validado apresentou uma boa precisão, com coeficientes de variação menores que 10% (MAPA, 2015).

3.4.3. Análise de amostras de azeite de oliva adulteradas

O método validado foi utilizado para analisar amostras de azeites de oliva adulteradas. Na Tabela 13 são apresentados os resultados dos teores médios de esqualeno das amostras do *blend* dos azeites de oliva extra virgem adulteradas em cada porcentagem de adulteração, para os sete adulterantes.

Tabela 13 - Teores médios de esqualeno das amostras do *blend* dos azeites de oliva extra virgem adulteradas em cada porcentagem de adulteração, para os sete adulterantes.

Amostras adulteradas	Teor de esqualeno / g kg⁻¹	
<i>Blend puro</i>		2,93 ± 0,10
	1%	2,87 ± 0,20
	2%	2,65 ± 0,03
<i>Blend + Lampante</i>	5%	2,45 ± 0,05
	10%	2,22 ± 0,50
	15%	1,97 ± 0,40
	20%	1,58 ± 0,30
	1%	2,90 ± 0,05
<i>Blend + Virgem</i>	2%	2,73 ± 0,08
	5%	2,60 ± 0,10
	10%	2,45 ± 0,90
	15%	2,32 ± 0,08
	20%	2,00 ± 0,06
	1%	2,85 ± 0,05
<i>Blend + Tipo único</i>	2%	2,58 ± 0,05
	5%	2,31 ± 0,02
	10%	2,04 ± 0,04
	15%	1,73 ± 0,03
	20%	1,22 ± 0,50
	1%	2,68 ± 0,50
<i>Blend + Óleo de canola</i>	2%	2,43 ± 0,30
	5%	2,07 ± 0,05
	10%	1,72 ± 0,06
	15%	1,37 ± 0,05
	20%	0,82 ± 0,03
	1%	2,73 ± 0,07
<i>Blend + Óleo de girassol</i>	2%	2,45 ± 0,05
	5%	2,23 ± 0,80
	10%	1,95 ± 0,09
	15%	1,73 ± 0,05
	20%	1,20 ± 0,07
	1%	2,72 ± 0,90
<i>Blend + Óleo de milho</i>	2%	2,42 ± 0,07
	5%	2,14 ± 1,00
	10%	1,82 ± 0,08
	15%	1,55 ± 0,75
	20%	0,87 ± 0,15
	1%	2,73 ± 0,09
<i>Blend + Óleo de soja</i>	2%	2,25 ± 0,07
	5%	2,01 ± 0,98
	10%	1,75 ± 0,76
	15%	1,47 ± 0,45
	20%	1,18 ± 0,70

Os resultados apresentados na Tabela 13 mostram que, à medida que a porcentagem de adulteração aumentou, o teor de esqualeno das amostras do *blend* dos azeites de oliva adulteradas diminuiu. Este resultado era esperado visto que o azeite de oliva extra virgem possui um teor alto de esqualeno, enquanto que para os adulterantes, devido a processos de refino, no caso dos azeites de outras categorias, e às diferenças na composição, no caso dos óleos de sementes, o teor de esqualeno é mais baixo (Cayuela & García, 2018; Pacetti *et al.*, 2019).

É necessário ressaltar que o teor de esqualeno pode variar de acordo com a variedade das azeitonas e com os processos de extração e de refino do óleo (Aresta *et al.*, 2020). Desta forma, é conveniente que a determinação do teor de esqualeno seja associada a outros métodos para que a verificação da autenticidade do azeite de oliva ocorra de forma inequívoca, garantindo que as variações devido a fatores associados à produção e ao processamento não sejam mais significativas do que as variações do teor de esqualeno provenientes da adulteração.

Grigoriadou e colaboradores (2007) desenvolveram um método utilizando a extração em fase sólida para análise do teor de esqualeno em azeites de oliva virgens, encontrando teores de esqualeno variando de 2,04 a 4,66 g kg⁻¹. Ainda, Pacetti e colaboradores (2019) encontraram teores de esqualeno variando de 3,5 a 10,2 g kg⁻¹ para azeites de oliva extra virgens de diferentes variedades. As diferenças encontradas entre estes valores e àqueles encontrados neste trabalho estão relacionados às diferenças entre as variedades de azeitonas utilizadas na obtenção dos azeites de oliva.

3.4.4. Análise de amostras de azeite de oliva coletadas para fiscalização

Amostras de azeites de oliva coletadas para fiscalização pelo IQA/LFDA-MG foram analisadas pelo método validado. Na Tabela 14 são apresentados os resultados dos teores de esqualeno das amostras de azeites de oliva coletadas para fiscalização.

Tabela 14 - Teores de esqualeno de amostras de azeites de oliva coletadas para fiscalização realizada pelo IQA/LFDA-MG.

Amostras de fiscalização	Teor de esqualeno / g kg⁻¹
AM01	2,62
AM02	1,53
AM03	1,05
AM04	2,47
AM05	0,87
AM06	1,17
AM07	0,75
AM08	2,48
AM09	3,00
AM10	2,83

As amostras coletadas para fiscalização apresentaram teores de esqualeno de 0,87 a 3,00 g kg⁻¹, mostrando grande variação nos valores encontrados.

As amostras AM1 a AM10 foram analisadas quanto à absorvidade molar (extinção específica) no ultravioleta, composição em ácidos graxos, acidez livre, umidade e matéria volátil, índice de peróxidos e índice de refração, utilizando os métodos oficiais do IOC para verificação da autenticidade de azeites de oliva. A partir dos resultados encontrados por estas análises, as amostras foram classificadas como azeite de olive extra virgem genuíno ou não genuíno.

Na Tabela 15 é apresentada a classificação das amostras de azeites de oliva coletadas para fiscalização realizada pelo IQA/LFDA-MG, para verificação da autenticidade.

Tabela 15 - Classificação de amostras de azeites de oliva coletadas para fiscalização realizada pelo IQA/LFDA-MG, após análises para verificação de autenticidade de azeite de oliva.

Amostras de fiscalização	Classificação pelo IQA/LFDA-MG
AM01	Azeite de oliva extra virgem genuíno
AM02	Azeite de oliva extra virgem não genuíno
AM03	Azeite de oliva extra virgem não genuíno
AM04	Azeite de oliva extra virgem genuíno
AM05	Azeite de oliva extra virgem não genuíno
AM06	Azeite de oliva extra virgem não genuíno
AM07	Azeite de oliva extra virgem não genuíno
AM08	Azeite de oliva extra virgem genuíno
AM09	Azeite de oliva extra virgem genuíno
AM10	Azeite de oliva extra virgem genuíno

As amostras AM01, AM04, AM08, AM09 e AM10 foram classificadas como azeite de oliva extra virgem genuínos. Estas amostras apresentaram os maiores teores de esqualeno determinados pelo método, variando de 2,47 a 3,00 g kg⁻¹. As demais foram classificadas como não genuínos, indicando a inautenticidade destes como azeites de oliva extra virgem. Para estas amostras os teores de esqualeno foram menores, variando de 0,87 a 1,53 g kg⁻¹. Como esperado, amostras de azeites de oliva adulteradas apresentaram menores valores de esqualeno.

Os resultados encontrados para os teores de esqualeno das amostras coletadas para fiscalização estão em concordância com as classificações realizadas pelo IQA/LFDA-MG, mostrando que a determinação do teor de esqualeno pelo método validado foi eficiente para verificação da autenticidade de azeites de oliva.

3.5. CONCLUSÕES

O método validado utilizando a SPE e análise por GC-MS/MS mostrou-se simples e eficiente, utilizando uma pequena quantidade de amostra e um volume de solvente muito menor do que o método de referência para determinação de esqualeno. Os parâmetros de desempenho mostraram que o método apresentou uma boa repetitividade, com desvios padrões relativos menores que 6,5%, e uma

faixa grande de linearidade. Além disso, o método apresentou uma boa exatidão nos resultados, com porcentagens de recuperação entre 95% e 104%.

Como esperado, as análises das amostras adulteradas mostraram que a adição dos adulterantes ao *blend* dos azeites de oliva extra virgens gerou uma diminuição dos teores de esqualeno. Em relação às amostras coletadas para fiscalização, os valores dos teores de esqualeno encontrados pelo método corroboraram com as classificações realizadas pelo IQA/LFDA-MG. Portanto, o método validado mostrou-se eficiente na determinação de esqualeno em azeites de oliva.

***CAPÍTULO 4 – CONSTRUÇÃO DE MODELOS POR
FUSÃO DE DADOS OBTIDOS POR FT-NIR E GC-
MS/MS PARA VERIFICAÇÃO DA AUTENTICIDADE
DE AZEITE DE OLIVA***

4.1. INTRODUÇÃO

O azeite de oliva é um alimento com alto valor agregado e que recentemente tem despertado a atenção dos consumidores devido a seus benefícios nutricionais e suas qualidades sensoriais. Por isso, diversos métodos e técnicas têm sido desenvolvidas visando analisar a qualidade deste produto (Rohman & Man, 2010; Hodaifa *et al.*, 2012).

O azeite de oliva é uma matriz complexa, com composição química muito variada. Desta forma, as técnicas utilizadas para análise de azeite de oliva geram uma grande quantidade de dados, demandando novas ferramentas para melhor interpretação desses resultados. Uma destas ferramentas é a criação de modelos por fusão de dados (Biancolillo *et al.*, 2019).

A fusão de dados pode ser definida como a combinação de blocos de dados gerados por diferentes técnicas analíticas em modelos únicos. Os dados gerados pelas diferentes técnicas podem ser combinados de diferentes formas para melhorar a classificação de amostras e a predição de suas propriedades (Borràs *et al.*, 2015).

As análises por técnicas como a espectroscopia no infravermelho próximo e a cromatografia gasosa acoplada à espectrometria de massas podem fornecer diferentes informações sobre as amostras e a fusão de dados utiliza essas informações e abrange as vantagens das tecnologias instrumentais (Bevilacqua *et al.*, 2017).

A fusão de dados tem sido uma estratégia muito utilizada em investigações de adulterações em alimentos, como o azeite de oliva (Bevilacqua *et al.*, 2017; Borràs *et al.*, 2015; Li *et al.*, 2019). Esta ferramenta torna as análises cada vez mais específicas, pois a fusão reduz o impacto de fontes espúrias de variabilidade e potenciais interferentes nos dados, causando uma diminuição nos erros de predição dos modelos (Li *et al.*, 2019; Bevilacqua *et al.*, 2017).

O objetivo deste capítulo foi construir modelos sensíveis e confiáveis por meio da fusão de dados, associando os dados de FT-NIR e GC-MS/MS, PLS e o método OPS para seleção de variáveis para calibrar amostras de azeites de oliva extra virgem adulteradas.

4.2. OBJETIVOS

4.2.1. Objetivo geral

Utilizar a fusão de dados associada a métodos quimiométricos e calibração multivariada para verificação da autenticidade de azeites de oliva.

4.2.2. Objetivos específicos

- Construir modelos por fusão de dados, associando os dados obtidos pelas técnicas de espectroscopia (Capítulo 2) e cromatografia gasosa (Capítulo 3);
- Construir modelos individuais e robusto utilizando a fusão de dados de FT-NIR e GC-MS/MS, PLS e o método OPS para seleção de variáveis para calibrar amostras do *blend* de azeites de oliva extra virgem adulteradas.

4.3. MATERIAL E MÉTODOS

4.3.1. Amostras

Foram empregadas as amostras do *blend* dos azeites de oliva extra virgem adulteradas descritas no item 2.3.1 (Capítulo 2). Os dados gerados pelas análises dessas amostras por FT-NIR (Capítulo 2) e GC-MS/MS (Capítulo 3) foram utilizados para a construção dos modelos por fusão de dados.

4.3.2. Construção dos modelos de calibração

O PLS foi utilizado para desenvolver modelos individuais e robusto para a calibração das amostras adulteradas.

Para divisão do conjunto de dados em calibração e validação, as amostras foram separadas em 2/3 para o conjunto de calibração e 1/3 para o conjunto de validação, através de planejamento experimental. Dessa forma, as amostras foram separadas pela ordem CVC, garantindo a presença de amostras em toda a faixa de concentração. Para a escolha do melhor número de variáveis latentes, a técnica de validação cruzada foi aplicada, utilizando subconjuntos aleatórios, com 10 subconjuntos e 20 interações.

Para a técnica de FT-NIR foram utilizadas as variáveis espectrais (número de onda e absorvância) e para a técnica de GC-MS/MS foram utilizados todos os valores de intensidades dos picos em todo o cromatograma (0 a 9,0 min), incluindo o pico referente ao esqualeno em 7,2 min.

Os dados foram pré-processados por meio de autoescalamento.

O método OPS foi aplicado para identificar sub-grupos de variáveis com maior poder de predição e executado com as seguintes configurações otimizadas: vetor informativo: coeficientes de regressão; janela: 5 e incrementos: 5.

4.4. RESULTADOS E DISCUSSÃO

4.4.1. Construção dos modelos de calibração

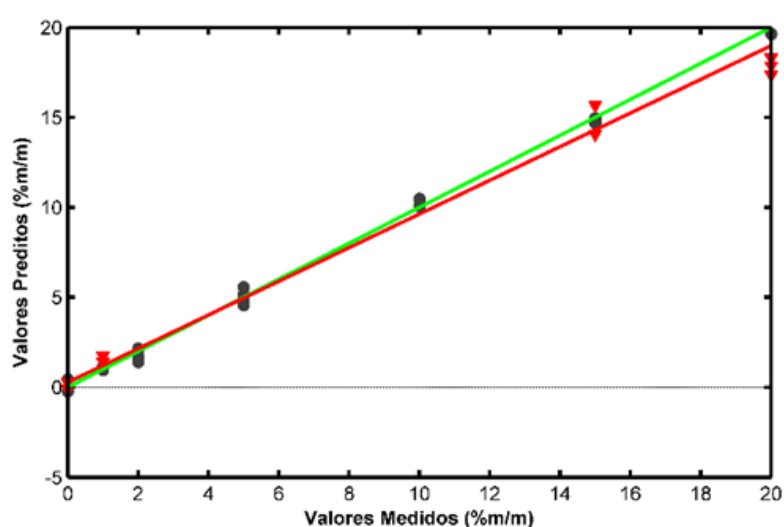
Modelos PLS obtidos por fusão de dados foram construídos para calibração das amostras do *blend* de azeite de oliva extra virgem adulteradas. Além dos modelos específicos, também foi construído um modelo robusto que incorporou todas as amostras. Os modelos foram construídos pelos valores preditos (valores calculados pelo modelo) *versus* os valores medidos (valores teóricos) para as amostras adulteradas. Os resultados para os modelos com melhores desempenhos são apresentados a seguir. Em todos os modelos, o melhor pré-processamento foi MSC seguido de centragem dos dados na média.

Nas Figuras 30 a 37 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para os sete adulterantes e para o modelo robusto, unindo todos os adulterantes em um mesmo modelo. Nas tabelas 16 e 17 são apresentados os parâmetros estatísticos para os modelos de calibração dos adulterantes. Os resultados serão discutidos no item 4.4.2.

4.4.1.1. Azeite lampante

Na Figura 30 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o adulterante azeite lampante.

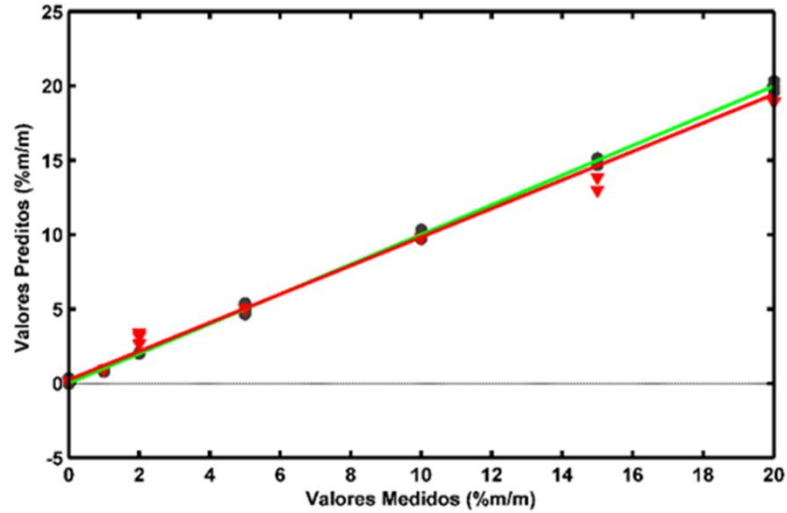
Figura 30 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite lampante.



4.4.1.2. Azeite virgem

Na Figura 31 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o adulterante azeite virgem.

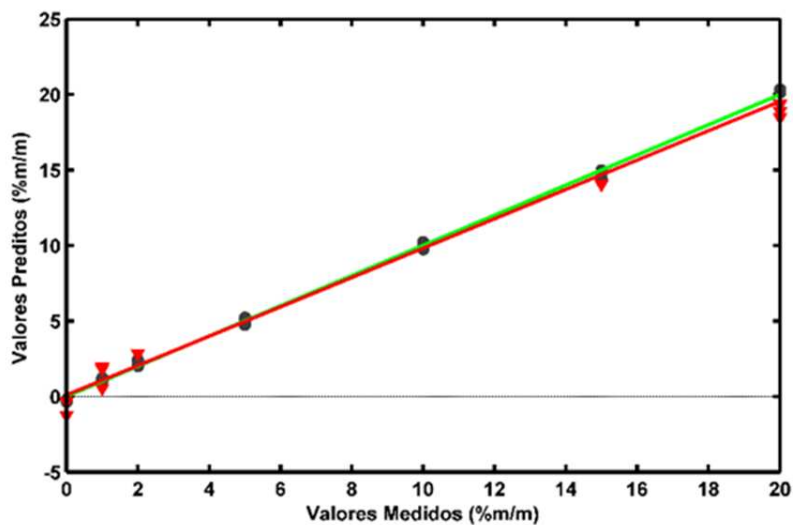
Figura 31 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite virgem.



4.4.1.3. Azeite tipo único

Na Figura 32 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o adulterante azeite tipo único.

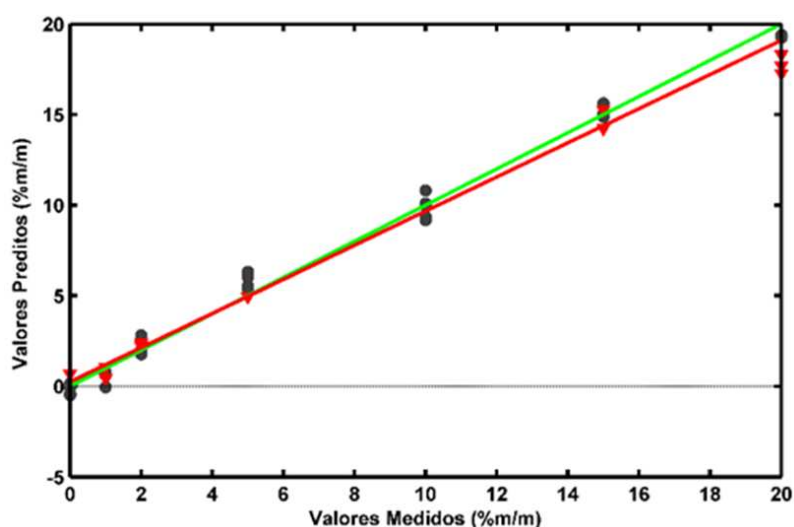
Figura 32 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante azeite tipo único.



4.4.1.4. Óleo de canola

Na Figura 33 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o adulterante óleo de canola.

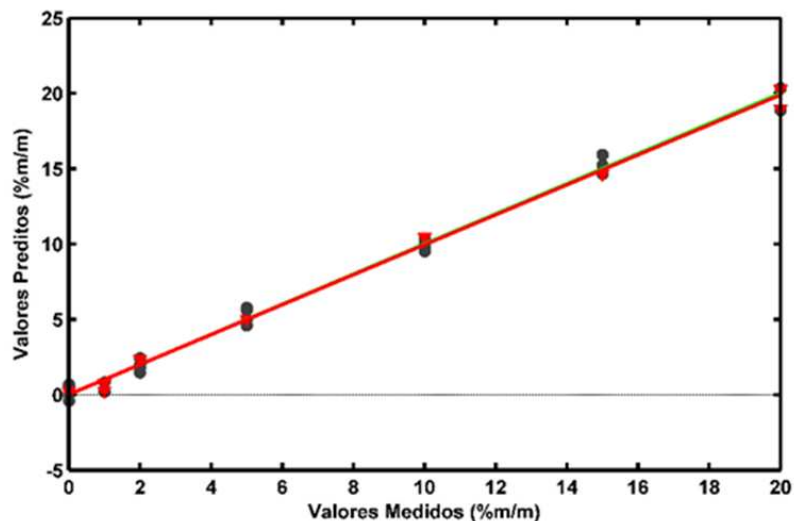
Figura 33 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de canola.



4.4.1.5. Óleo de girassol

Na Figura 34 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o adulterante óleo de girassol.

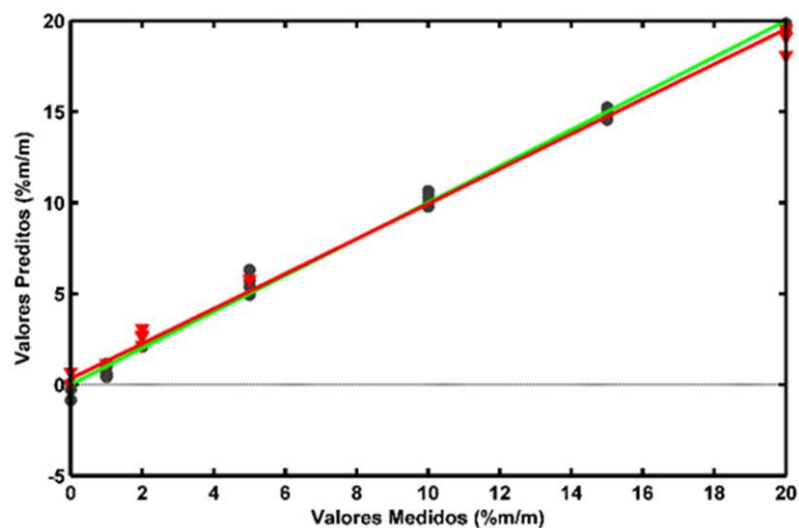
Figura 34 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de girassol.



4.4.1.6. Óleo de milho

Na Figura 35 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o adulterante óleo de milho.

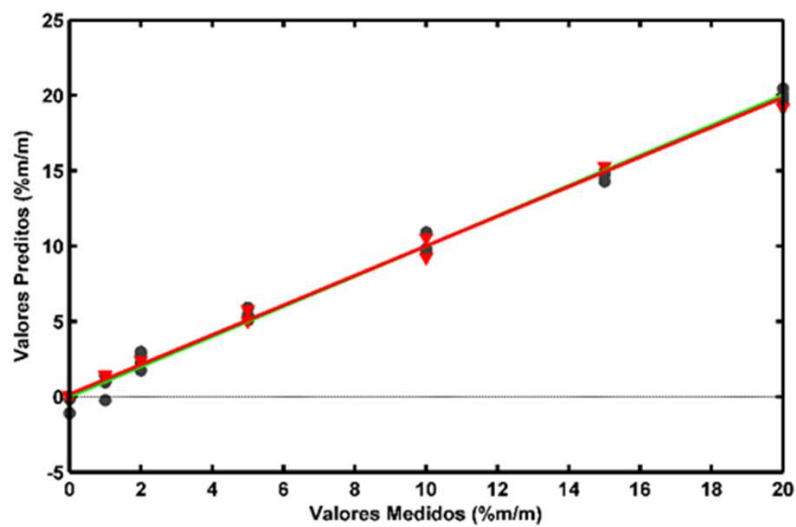
Figura 35 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de milho.



4.4.1.7. Óleo de soja

Na Figura 36 são apresentados os gráficos de valores *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o adulterante óleo de soja.

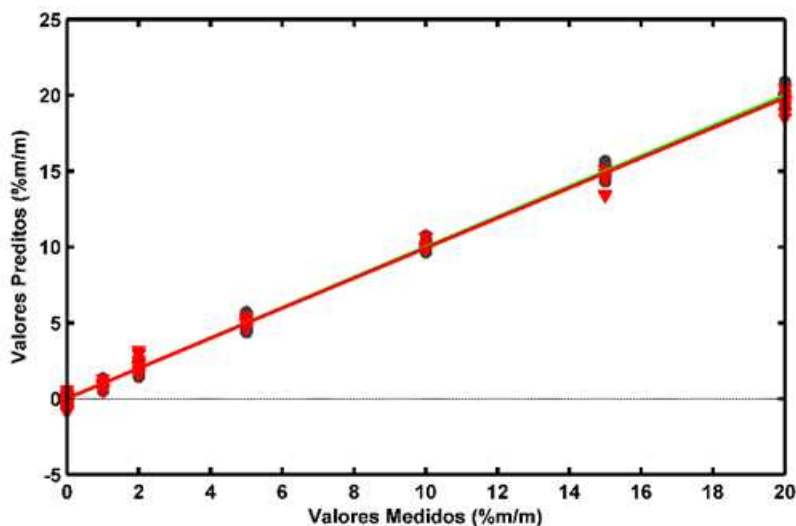
Figura 36 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o adulterante óleo de soja.



4.4.1.8. Modelo robusto

Na Figura 37 são apresentados os gráficos de valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo e pelo método OPS para o modelo robusto.

Figura 37 - Valores preditos *versus* valores medidos (% m/m) para os modelos de calibração PLS completo (vermelho, ▼) e pelo método OPS (verde, ●) para o modelo robusto.



4.4.1.9. Parâmetros estatísticos para os modelos de calibração

Nas tabelas 16 e 17 são apresentados os parâmetros estatísticos para os modelos de calibração PLS para os adulterantes e para o modelo robusto.

Tabela 16 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para os adulterantes lampante, virgem, tipo único e óleo de canola.

	Lampante		Virgem		Tipo Único		Óleo de Canola	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	2	2	2	2	2	2	2	2
nvars	3133	180	3133	280	3133	145	3133	225
RMSEC (%)	0.66	0.31	0.93	0.23	1.04	0.22	0.98	0.60
RMSECV (%)	5.46	1.09	4.83	1.17	5.11	0.75	4.64	1.04
RMSEP (%)	3.80	1.32	3.93	1.02	4.35	0.91	3.95	1.25
Rc	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Rcv	0.88	0.98	0.94	0.99	0.85	0.99	0.80	0.99
Rp	0.97	0.99	0.93	0.99	0.87	0.99	0.98	0.99

Tabela 17 - Parâmetros estatísticos para os modelos de calibração PLS completo e pelo método OPS, para os adulterantes óleo de girassol, óleo de milho, óleo de soja e para o modelo robusto.

	Óleo de Girassol		Óleo de Milho		Óleo de Soja		Robusto	
	Completo	OPS	Completo	OPS	Completo	OPS	Completo	OPS
nLv	2	2	2	2	2	2	3	3
nvars	3133	25	3133	215	3133	170	3133	350
RMSEC (%)	0.75	0.50	0.67	0.46	1.51	0.57	0.71	0.30
RMSECV (%)	4.17	0.56	4.32	0.78	4.00	0.71	1.50	0.63
RMSEP (%)	2.56	0.47	3.68	0.85	3.60	0.51	1.51	0.57
Rc	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99
Rcv	0.97	0.99	0.95	0.99	0.89	0.99	0.99	0.99
Rp	0.97	0.99	0.96	0.99	0.99	0.99	0.98	0.99

4.4.2. Discussão dos resultados

Os modelos PLS por fusão de dados apresentaram bom desempenho na calibração das amostras adulteradas. Para os modelos individuais completos dos adulterantes, os valores de RMSEC, RMSECV e RMSEP foram mais altos do que para o modelo robusto. Ainda, os nLv foram iguais a 2 para todos os modelos individuais dos adulterantes, e 3 para o modelo robusto.

O número total de variáveis para os modelos por fusão de dados foi de 3133, pois inclui todos os dados próprios dos espectros por FT-NIR (1501 variáveis) e todos os dados de intensidade dos picos presentes nos cromatogramas (1632 variáveis). A aplicação do método OPS para a seleção de variáveis melhorou consideravelmente os modelos para todos os adulterantes. Em geral, o número de variáveis utilizadas nos modelos PLS-OPS foi menor que 9% do número de variáveis originais, aumentando a capacidade preditiva dos modelos de calibração. A aplicação do método OPS nos modelos por fusão dos dados gerou reduções maiores no número de variáveis em relação aos modelos individuais por FT-NIR.

Um número menor de variáveis utilizadas pelos modelos é satisfatório, pois estas representam os conjuntos de informações espectrais e cromatográficas mais importantes para predizer a composição das amostras adulteradas, eliminando muitas variáveis não informativas.

Para os modelos dos adulterantes azeites virgem e tipo único foram selecionadas variáveis espectroscópicas na região abaixo de 6000 cm^{-1} , indicando que a composição em triacilgliceróis e compostos minoritários permitiu a diferenciação entre as amostras. Para o adulterante azeite virgem foram selecionadas variáveis de ambas as técnicas, na região abaixo de 6000 cm^{-1} para a espectroscopia e tempos de retenção variando de 5,009 a 8,241 min para a cromatografia, incluindo o pico referente ao esqualeno. Variáveis de ambas as técnicas também foram selecionadas para os modelos dos adulterantes óleo de soja, óleo de canola e óleo de milho. Já os modelos PLS-OPS dos adulterantes azeite lampante, azeite tipo único e óleo de girassol foram construídos somente com variáveis cromatográficas (5,029 a 8,263 min). Para estes adulterantes as diferenças entre os teores de esqualeno foram determinantes para a diferenciação entre as amostras.

A heterogeneidade das variáveis selecionadas está relacionada à complexidade das amostras adulteradas, visto que a composição do azeite de oliva extra virgem apresenta características de ambas as variedades. Além disso, tem-se as características associadas à composição de cada adulterante.

Os valores de RMSEP para os modelos de fusão de dados foram reduzidos após a seleção de variáveis. Para o modelo do adulterante óleo de soja a redução foi de 86% (de 3,60% para 0,51% m/m); para o adulterante tipo único, a redução foi de 79% (de 4,35% para 0,91% m/m); o modelo PLS-OPS para o adulterante azeite lampante teve uma redução de 65% (de 3,80% para 1,32% m/m) no valor do erro de predição. Os valores de RMSEC e RMSECV também foram significativamente menores e os coeficientes de correlação foram maiores, em geral, para os modelos PLS-OPS para todos os adulterantes.

Nas Figuras 30 a 37 são apresentados os gráficos dos valores preditos *versus* valores medidos que permitem avaliar a linearidade dos modelos de calibração. Para os modelos PLS-OPS para todos os adulterantes, todos os valores de coeficientes de correlação foram superiores a 0,98, mostrando uma boa linearidade em toda a faixa investigada, apresentando melhores resultados gerais para a linearidade em relação aos modelos construídos somente com os dados de FT-NIR.

Para o modelo robusto, a aplicação do método OPS reduziu o número de variáveis utilizadas para menos de 12%, com um total de 350 variáveis. Foram utilizadas variáveis cromatográficas (5,049 a 8,263 min) e espectroscópicas (região abaixo de 6000 cm^{-1}), sendo considerados para diferenciação entre as amostras principalmente os teores de esqualeno e os triacilgliceróis. Em relação aos modelos individuais, o modelo robusto apresentou uma quantidade maior de variáveis selecionadas devido à maior complexidade dos dados utilizados, pois este modelo abrange os espectros de todas as amostras adulteradas e dos adulterantes puros. Ainda, o modelo PLS-OPS robusto apresentou baixos valores de RMSEC, RMSEP e RMSECV e todos os valores de coeficientes de correlação superiores a 0,99. Isso mostra que o modelo robusto PLS-OPS construído por fusão de dados é muito confiável, permitindo a diferenciação dos sete adulterantes mesmo quando os adulterantes se encontram em baixas concentrações. Utilizar a fusão de dados tornou o modelo robusto mais sensível e com maior capacidade preditiva que a maioria dos modelos robustos individuais construídos somente com informações espectrais, incluindo informações espectroscópicas e cromatográficas de todas as amostras adulteradas e os adulterantes puros.

São raros os trabalhos encontrados na literatura que utilizam a fusão de dados de cromatografia gasosa e espectroscopia para análises de azeite de oliva. Schwolow e colaboradores (2019) construíram um modelo PLS por fusão de dados de cromatografia gasosa e espectroscopia no infravermelho mais sensíveis do que os modelos criados com dados individuais das técnicas para diferenciação geológica de azeites de oliva. O modelo construído apresentou uma taxa de erro de validação de 14%.

4.5. CONCLUSÕES

Modelos PLS-OPS por fusão de dados foram construídos para cada adulterante. Os parâmetros estatísticos obtidos para todos os métodos mostram que a aplicação do método OPS para seleção de variáveis diminuiu significativamente os erros de calibração, validação e predição, pois possibilitou a eliminação de informações redundantes, aumentando a robustez e a precisão das classificações.

Para os sete adulterantes, os modelos robustos apresentaram baixos valores de RMSEC, RMSECV e RMSEP e boa linearidade, caracterizada por coeficientes de correlação maiores que 0,98, mostrando a boa capacidade preditiva dos modelos em baixas concentrações de adulteração.

Um modelo robusto PLS-OPS por fusão de dados para calibração das amostras adulteradas foi construído com bom desempenho. O modelo permitiu a diferenciação simultânea dos sete adulterantes em baixas concentrações para um *blend* de azeites de oliva extra virgem genuínos, com coeficientes de correlação superiores a 0,99. A utilização da fusão de dados tornou o modelo robusto com maior capacidade preditiva que a maioria dos modelos robustos individuais, incluindo informações espectrais e cromatográficas de todas as amostras adulteradas e os adulterantes puros.

CONCLUSÕES GERAIS

Neste trabalho foi possível verificar a autenticidade de azeite de oliva aplicando-se a espectroscopia no infravermelho próximo, a cromatografia gasosa acoplada à espectrometria de massas sequencial, métodos quimiométricos, fusão de dados e seleção de variáveis.

A análise dos parâmetros de validação mostrou que a aplicação dos métodos OPS e GA para seleção de variáveis contribuiu para a construção de modelos mais sensíveis e simples de interpretar, promovendo a eliminação de informações redundantes, aumentando a robustez e a precisão das classificações e calibrações multivariadas.

Dados de FT-NIR foram usados na construção de modelos PLS-DA. Os modelos classificaram corretamente amostras de azeite de oliva, de variedades diferentes e adulteradas com sete diferentes adulterantes, mesmo com adulteração muito baixa, ao nível de 1,0 % m/m. A FT-NIR também foi utilizada como fonte de dados para a construção de modelos de calibração multivariada para quantificar sete adulterantes em azeite de oliva extra virgem, com baixos erros de previsão, mesmo em amostras com baixos teores de adulterante.

Um método utilizando a SPE e análise por GC-MS/MS foi validado, mostrando-se simples e eficiente para determinação de esqualeno em amostras de azeite de oliva adulteradas e amostras de fiscalização.

A metodologia de fusão de dados de espectroscopia e cromatografia é uma contribuição importante deste trabalho. A fusão de dados permitiu a construção de um modelo de calibração PLS robusto que permitiu a quantificação simultânea de sete adulterantes em azeite de oliva extra virgem, com baixos erros de calibração e previsão, mesmo em amostras com nível de adulteração de 1,0% m/m. Assim, a fusão de dados representa uma ferramenta eficaz para verificação da autenticidade de azeite de oliva, visando uma maior qualidade deste produto.

REFERÊNCIAS BIBLIOGRÁFICAS

- Andersen, C. M., & Bro, R. (2010). Variable selection in regression-a tutorial. *Journal of Chemometrics*, 24(11–12), 728–737. <https://doi.org/10.1002/cem.1360>
- Antoniassi, R., & Lago, R. (1998). Evaluation of the identity and quality characteristics of olive. *Brazilian Journal of Food Technology*, 1 (1,2), 32-43
- AOAC (1999). Squalene in oils and fats. Titrimetric method. Official method. AOAC, p. 943.04
- Aparicio, R., Morales, M. T., Aparicio-Ruiz, R., Tena, N., & García-González, D. L. (2013). Authenticity of olive oil: Mapping and comparing official methods and promising alternatives. *Food Research International*, 54(2), 2025–2038. <https://doi.org/10.1016/j.foodres.2013.07.039>
- Aresta, A., Damascelli, A., De Vietro, N., & Zambonin, C. (2020). Measurement of squalene in olive oil by fractional crystallization or headspace solid phase microextraction coupled with gas chromatography. *International Journal of Food Properties*, 23(1), 1845–1853. <https://doi.org/10.1080/10942912.2020.1833033>
- Aued-pimentel, S., Takemoto, E., Emy, E., Bonaldi, C., Lutz, I. A., Alimentos, S. De, Bromatologia, D. De, & Arnaldo, A. (2008). *Química Nova*, 31(1), 31–34.
- Azizian, H., Mossoba, M. M., Fardin-Kia, A. R., Karunathilaka, S. R., & Kramer, J. K. G. (2016). Developing FT-NIR and PLS1 methodology for predicting adulteration in representative varieties/blends of extra virgin olive oils. *Lipids*, 51(11), 1309–1321. <https://doi.org/10.1007/s11745-016-4195-0>
- Balabin, R. M., & Smirnov, S. V. (2011). Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data. *Analytica Chimica Acta*, 692(1–2), 63–72. <https://doi.org/10.1016/j.aca.2011.03.006>
- Bevilacqua, M., Bro, R., Marini, F., Rinnan, Å., Rasmussen, M. A., & Skov, T. (2017). Recent chemometrics advances for foodomics. *TrAC - Trends in Analytical Chemistry*, 96, 42–51. <https://doi.org/10.1016/j.trac.2017.08.011>
- Biancolillo, A., Boqué, R., Cocchi, M., & Marini, F. (2019). Data fusion strategies in food analysis. In *Data Handling in Science and Technology* (Vol. 31). <https://doi.org/10.1016/B978-0-444-63984-4.00010-7>
- Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: A rapid-response analytical

- tool. *TrAC - Trends in Analytical Chemistry*, 21(4), 240–250. [https://doi.org/10.1016/S0165-9936\(02\)00404-1](https://doi.org/10.1016/S0165-9936(02)00404-1)
- Bokobza, L. (1998). Near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 6(1–4), 3–17. <https://doi.org/10.1255/jnirs.116>
- Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., & Busto, O. (2015). Data fusion methodologies for food and beverage authentication and quality assessment - A review. *Analytica Chimica Acta*, 891, 1–14. <https://doi.org/10.1016/j.aca.2015.04.042>
- Boskou, D. (2006) *Olive Oil: Chemistry and Technology*. AOCS Press, 2nd edition.
- Boskou, D., Blekas, G., & Tsimidou, M. (2006). Olive oil composition. In *Olive Oil: Chemistry and Technology: Second Edition*. AOCS Press. <https://doi.org/10.1016/B978-1-893997-88-2.50008-0>
- Brasil (2012). Ministério da Agricultura, Pecuária e Abastecimento. Instrução normativa 001/2012. Disponível em <https://www.gov.br/agricultura/ptbr/assuntos/inspecao/produtosvegetal/legislacao-1/normativos-cgqv/pocs/instrucao-normativa-no-1-de-30-de-janeiro-de-2012-azeite-de-oliva/view>
- Brereton, R. G. (2007). *Applied Chemometrics for Scientists*. Chichester, Inglaterra
- Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: Taking the magic away. *Journal of Chemometrics*, 28(4), 213–225. <https://doi.org/10.1002/cem.2609>
- Broadhurst, D., Goodacre, R., Jones, A., Rowland, J. J., & Kell, D. B. (1997). Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. *Analytica Chimica Acta*, 348(1–3), 71–86. [https://doi.org/10.1016/S0003-2670\(97\)00065-2](https://doi.org/10.1016/S0003-2670(97)00065-2)
- Cayuela, J. A., & García, J. F. (2018). Nondestructive measurement of squalene in olive oil by near infrared spectroscopy. *LWT - Food Science and Technology*, 88, 103–108. <https://doi.org/10.1016/j.lwt.2017.09.047>
- Cecchi, L., Migliorini, M., Giambanelli, E., Rossetti, A., Cane, A., Mulinacci, N., & Melani, F. (2020). Authentication of the geographical origin of virgin olive oils from the main worldwide producing countries: A new combination of HS-SPME-

- GC-MS analysis of volatile compounds and chemometrics applied to 1217 samples. *Food Control*, *112*, 107156. <https://doi.org/10.1016/j.foodcont.2020.107156>
- Christy, A. A., Kasemsumran, S., Du, Y., & Ozaki, Y. (2004). The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics. *Analytical Sciences*, *20*(6), 935–940. <https://doi.org/10.2116/analsci.20.935>
- Ciafardini, G., & Zullo, B. A. (2018). Virgin olive oil yeasts: A review. *Food Microbiology*, *70*, 245–253. <https://doi.org/10.1016/j.fm.2017.10.010>
- Codex Alimentarius Commission (2003). *Standard for olive oils and olive pomace oils*. Codex Stan, v. 33, n. 8.
- Collins, C. H., Braga, G. L., & Bonato, P. S. (2006). *Fundamentos de cromatografia*.
- Cozzolino, D., Murray, I., Chree, A., & Scaife, J. R. (2005). Multivariate determination of free fatty acids and moisture in fish oils by partial least-squares regression and near-infrared spectroscopy. *LWT - Food Science and Technology*, *38*(8), 821–828. <https://doi.org/10.1016/j.lwt.2004.10.007>
- da Costa Filho, P. A., & Poppi, R. J. (1999). Genetic algorithm in chemistry. *Quimica Nova*, *22*(3), 405–411. <https://doi.org/10.1590/s0100-40421999000300019>
- da Costa, J. R. O., Dal Bosco, S. M., Ramos, R. C. de S., Machado, I. C. K., Garavaglia, J., & Villasclaras, S. S. (2020). Determination of volatile compounds responsible for sensory characteristics from Brazilian extra virgin olive oil using HS-SPME/GC-MS direct method. *Journal of Food Science*, *85*(11), 3764–3775. <https://doi.org/10.1111/1750-3841.15467>
- da Silveira, R., Vágula, J. M., de Lima Figueiredo, I., Claus, T., Galuch, M. B., Santos Junior, O. O., & Visentainer, J. V. (2017). Rapid methodology via mass spectrometry to quantify addition of soybean oil in extra virgin olive oil: A comparison with traditional methods adopted by food industry to identify fraud. *Food Research International*, *102*, 43–50. <https://doi.org/10.1016/j.foodres.2017.09.076>
- Ferreira, M. M. C. (2015). *Quimiometria: conceitos, métodos e aplicações*. Editora da UNICAMP
- García-Nicolás, M., Arroyo-Manzanares, N., Arce, L., Hernández-Córdoba, M., &

- Viñas, P. (2020). Headspace gas chromatography coupled to mass spectrometry and ion mobility spectrometry: Classification of virgin olive oils as a study case. *Foods*, 9(9). <https://doi.org/10.3390/foods9091288>
- Gelad, P., & Kowalski, B. R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, 185, 1–17. [https://doi.org/10.1016/0003-2670\(86\)80028-9](https://doi.org/10.1016/0003-2670(86)80028-9)
- Gertz, C., Matthäus, B., & Willenberg, I. (2020). Detection of Soft-Deodorized Olive Oil and Refined Vegetable Oils in Virgin Olive Oil Using Near Infrared Spectroscopy and Traditional Analytical Parameters. *European Journal of Lipid Science and Technology*, 122(6), 1-11. <https://doi.org/10.1002/ejlt.201900355>
- Goicoechea, H. C., Olivieri, A. C., Xiaobo, Z., Jiewen, Z., Povey, M. J. W., Holmes, M., Hanpin, M., Da Costa Filho, P. A., & Poppi, R. J. (2010). A new family of genetic algorithms for wavelength interval selection in multivariate analytical spectroscopy. *Journal of Chemometrics*, 22(3), 405–411. <https://doi.org/10.1002/cem.812>
- Green, H. S., Li, X., Pra, M. De, Lovejoy, K. S., Steiner, F., Acworth, I. N., & Wang, S. C. (2020). A rapid method for the detection of extra virgin olive oil adulteration using UHPLC-CAD profiling of triacylglycerols and PCA. *Food Control*, 107(April 2019), 106773. <https://doi.org/10.1016/j.foodcont.2019.106773>
- Grigoriadou, D., Androulaki, A., Psomiadou, E., & Tsimidou, M. Z. (2007). Solid phase extraction in the analysis of squalene and tocopherols in olive oil. *Food Chemistry*, 105(2), 675–680. <https://doi.org/10.1016/j.foodchem.2006.12.065>
- Harwood, J., & Aparicio, R. (2000). *Handbook of Olive Oil - Analysis and Properties*.
- Hodaifa, G., Martínez Nieto, L. M., Lozano, J. L., Sánchez, S. (2012). Changes of the wax contents in mixtures of olive oils as determined by gas chromatography with a flame ionization detector. *Journal of AOAC International*, 95 (6), 1720–1724. <https://doi.org/10.5740/jaoacint.12-011>
- Holman, R. T., & Edmondson, P. R. (1956). Near-infrared spectra of fatty acids and some related substances. *Analytical Chemistry*, 28(10), 1533–1538. <https://doi.org/10.1021/ac60118a010>
- Hourant, P., Baeten, V., Morales, M. T., Meurens, M., & Aparicio, R. (2000). Oil and fat classification by selected bands of near-infrared spectroscopy. *Applied*

- Spectroscopy*, 54(8), 1168–1174. <https://doi.org/10.1366/0003702001950733>
- Indelicato, S., Bongiorno, D., Pitonzo, R., Di Stefano, V., Calabrese, V., Indelicato, S., & Avellone, G. (2017). Triacylglycerols in edible oils: Determination, characterization, quantitation, chemometric approach and evaluation of adulterations. *Journal of Chromatography A*, 1515, 1–16. <https://doi.org/10.1016/j.chroma.2017.08.002>
- INMETRO (2016). *Guidance in validation of analytical methods*. Instituto Nacional de Metrologia, Qualidade e Tecnologia, 31. http://www.inmetro.gov.br/Sidoq/Arquivos/CGCRE/DOQ/DOQ-CGCRE-8_05.pdf
- International Olive Council (IOC) (2003). *Determination of wax content by capillary column gas chromatography*. COI/T.20/Doc. No. 18/Rev.
- International Olive Council (IOC) (2010). *Determination of the difference between actual and theoretical content of triacylglycerols with ECN 42*. COI/T.20/Doc. No. 20/Rev. 3.
- International Olive Council (IOC) (2018). *Method of analysis evaluation of the coherence of TAG composition with the fatty acid composition*. COI/T.20/Doc. No. 25/Rev. 2.
- International Olive Council (IOC) (2020). *Key-Figures On The World Market For Olive Oils*. Data adopted at the 110th session of the IOOC. Disponível em <https://www.internationaloliveoil.org/wp-content/uploads/2020/04/110-HO-2018.pdf>, acessado em Janeiro de 2020
- Izquierdo, M., Lastra-mejías, M., González-flores, E., & Cancilla, J. C. (2020). Deep thermal imaging to compute the adulteration state of extra virgin olive oil. *Computers and Electronics in Agriculture*, 171(July 2019), 105290. <https://doi.org/10.1016/j.compag.2020.105290>
- Jiang, H., & Chen, Q. (2019). Virgin olive oil using FT-NIR spectroscopy. *Molecules*, 24, 2134(1-10).
- Jiménez-Carvelo, A. M., Lozano, V. A., & Olivieri, A. C. (2019). Comparative chemometric analysis of fluorescence and near infrared spectroscopies for authenticity confirmation and geographical origin of Argentinean extra virgin olive oils. *Food Control*, 96, 22–28. <https://doi.org/10.1016/j.foodcont.2018.08.024>
- Jiménez-Carvelo, A. M., Osorio, M. T., Koidis, A., González-Casado, A., & Cuadros-

- Rodríguez, L. (2017). Chemometric classification and quantification of olive oil in blends with any edible vegetable oils using FTIR-ATR and Raman spectroscopy. *LWT - Food Science and Technology*, 86, 174–184. <https://doi.org/10.1016/j.lwt.2017.07.050>
- Li, Y., Xiong, Y., & Min, S. (2019). Data fusion strategy in quantitative analysis of spectroscopy relevant to olive oil adulteration. *Vibrational Spectroscopy*, 101(November 2018), 20–27. <https://doi.org/10.1016/j.vibspec.2018.12.009>
- Lima, T. K. De, Musso, M., & Menezes, D. B. (2020). Using Raman spectroscopy and an exponential equation approach to detect adulteration of olive oil with rapeseed and corn oil. *Food Chemistry*, 127454. <https://doi.org/10.1016/j.foodchem.2020.127454>
- MAPA (2015). *Manual de Garantia da Qualidade Analítica*. In Ministério da Agricultura Pecuária e Abastecimento. file:///C:/Users/carol/Downloads/Manual-de-Garantia-Analitica-ilovepdf-compressed-ilovepdf-compressed (3).pdf
- Marcos Lorenzo, I., Pérez Pavón, J. L., Fernández Laespada, M. E., García Pinto, C., & Moreno Cordero, B. (2002). Detection of adulterants in olive oil by headspace-mass spectrometry. *Journal of Chromatography A*, 945(1–2), 221–230. [https://doi.org/10.1016/S0021-9673\(01\)01502-3](https://doi.org/10.1016/S0021-9673(01)01502-3)
- Milani, A., Lucci, P., Sedran, M., Moret, E., Moret, S., & Conte, L. (2020). Improved method for determination of waxes in olive oils: Reduction of silica and use of a less hazardous solvent. *OCL - Oilseeds and Fats, Crops and Lipids*, 27(1985). <https://doi.org/10.1051/ocl/2020016>
- Mossoba, M. M., Azizian, H., Fardin-Kia, A. R., Karunathilaka, S. R., & Kramer, J. K. G. (2017). First application of newly developed FT-NIR spectroscopic methodology to predict authenticity of extra virgin olive oil retail products in the USA. *Lipids*, 52(5), 443–455. <https://doi.org/10.1007/s11745-017-4250-5>
- Mustorgi, E., Malegori, C., Oliveri, P., Hooshyary, M., Bounneche, H., Mondello, L., Oteri, M., & Casale, M. (2020). A chemometric strategy to evaluate the comparability of PLS models obtained from quartz cuvettes and disposable glass vials in the determination of extra virgin olive oil quality parameters by NIR spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 199(January), 103974. <https://doi.org/10.1016/j.chemolab.2020.103974>

- Oussama, A., Elabadi, F., Platikanov, S., Kzaiber, F., & Tauler, R. (2012). Detection of olive oil adulteration using FT-IR spectroscopy and PLS with variable importance of projection (VIP) scores. *Journal of the American Oil Chemists' Society*, *89*(10), 1807-1812. <https://doi.org/10.1007/s11746-012-2091-1>
- Ozcan-sinir, G. (2020). Detection of adulteration in extra virgin olive oil by selected ion flow tube mass spectrometry (SIFT-MS) and chemometrics. *Food Control*, *118*(June), 107433. <https://doi.org/10.1016/j.foodcont.2020.107433>
- Özdemir, İ. S., Dağ, Ç., Özinanç, G., Suçsoran, Ö., Ertaş, E., & Bekiroğlu, S. (2018). Quantification of sterols and fatty acids of extra virgin olive oils by FT-NIR spectroscopy and multivariate statistical analyses. *LWT - Food Science and Technology*, *91*(January), 125–132. <https://doi.org/10.1016/j.lwt.2018.01.045>
- Pacetti, D., Scortichini, S., Boarelli, M. C., & Fiorini, D. (2019). Simple and rapid method to analyse squalene in olive oils and extra virgin olive oils. *Food Control*, *102*(January), 240–244. <https://doi.org/10.1016/j.foodcont.2019.03.005>
- Pasquini, C. (2003). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, *14*(2), 198–219. <https://doi.org/10.1590/S0103-50532003000200006>
- Pereira, L., Gomes, S., Barrias, S., Fernandes, J. R., & Martins-Lopes, P. (2018). Applying high-resolution melting (HRM) technology to olive oil and wine authenticity. *Food Research International*, *103*, 170–181. <https://doi.org/10.1016/j.foodres.2017.10.026>
- Pérez-Castaño, E., Sánchez-Viñas, M., Gázquez-Evangelista, D., & Bagur-González, M. G. (2018). Quantification of extra virgin olive oil in dressing and edible oil blends using the representative TMS-4,4'-desmethylsterols gas-chromatographic-normalized fingerprint. *Food Chemistry*, *239*, 1192–1199. <https://doi.org/10.1016/j.foodchem.2017.07.069>
- Revelou, P. K., Pappa, C., Kakouri, E., Kanakis, C. D., Papadopoulos, G. K., Pappas, C. S., & Tarantilis, P. A. (2021). Discrimination of botanical origin of olive oil from selected Greek cultivars by SPME-GC-MS and ATR-FTIR spectroscopy combined with chemometrics. *Journal of the Science of Food and Agriculture*, *101*(7), 2994–3002. <https://doi.org/10.1002/jsfa.10932>
- Rohman, A., & Man, Y. B. C. (2010). Fourier transform infrared (FTIR) spectroscopy

- for analysis of extra virgin olive oil adulterated with palm oil. *Food Research International*, 43(3), 886–892. <https://doi.org/10.1016/j.foodres.2009.12.006>
- Ruiz-Samblás, C., Tres, A., Koot, A., Van Ruth, S. M., González-Casado, A., & Cuadros-Rodríguez, L. (2012). Proton transfer reaction-mass spectrometry volatile organic compound fingerprinting for monovarietal extra virgin olive oil identification. *Food Chemistry*, 134(1), 589–596. <https://doi.org/10.1016/j.foodchem.2012.02.135>
- Sagratini, G., Allegrini, M., Caprioli, G., Cristalli, G., Giardina, D., Maggi, F., Ricciutelli, M., Sirocchi, V., & Vittori, S. (2013). Simultaneous determination of squalene, α -tocopherol and β -carotene in table olives by solid phase extraction and high-performance liquid chromatography with diode array detection. *Food Analytical Methods*, 6(1), 54–60. <https://doi.org/10.1007/s12161-012-9422-6>
- Salvo, A., La Torre, G. L., Rotondo, A., Mangano, V., Casale, K. E., Pellizzeri, V., Clodoveo, M. L., Corbo, F., Cicero, N., & Dugo, G. (2017). Determination of Squalene in Organic Extra Virgin Olive Oils (EVOOs) by UPLC/PDA Using a Single-Step SPE Sample Preparation. *Food Analytical Methods*, 10(5), 1377–1385. <https://doi.org/10.1007/s12161-016-0697-x>
- Samaniego-Sánchez, C., Quesada-Granados, J. J., de la Serrana, H. L. G., & López-Martínez, M. C. (2010). β -Carotene, squalene and waxes determined by chromatographic method in picual extra virgin olive oil obtained by a new cold extraction system. *Journal of Food Composition and Analysis*, 23(7), 671–676. <https://doi.org/10.1016/j.jfca.2010.03.010>
- Sato, T., Kawano, S., & Iwamoto, M. (1991). Near infrared spectral patterns of fatty acid analysis from fats and oils. *Journal of the American Oil Chemists Society*, 68(11), 827–833. <https://doi.org/10.1007/BF02660596>
- Schwolow, S., Gerhardt, N., Rohn, S., & Weller, P. (2019). Data fusion of GC-IMS data and FT-MIR spectra for the authentication of olive oils and honeys—is it worth to go the extra mile? *Analytical and Bioanalytical Chemistry*, 411(23), 6005–6019. <https://doi.org/10.1007/s00216-019-01978-w>
- Siano, F., & Vasca, E. (2020). GC-FID Analysis to Evaluate the Possible Adulteration of Extra Virgin Olive Oil with Different Vegetable Oils. *Journal of Chemical Education*, 97(11), 4108–4116. <https://doi.org/10.1021/acs.jchemed.0c00278>

- Sinelli, N., Casale, M., Di Egidio, V., Oliveri, P., Bassi, D., Tura, D., & Casiraghi, E. (2010). Varietal discrimination of extra virgin olive oils by near and mid infrared spectroscopy. *Food Research International*, 43(8), 2126–2131. <https://doi.org/10.1016/j.foodres.2010.07.019>
- Sohng, W., Park, Y., Jang, D., Cha, K., Jung, Y. M., & Chung, H. (2020). Incorporation of two-dimensional correlation analysis into discriminant analysis as a potential tool for improving discrimination accuracy: Near-infrared spectroscopic discrimination of adulterated olive oils. *Talanta*, 212, 120748. <https://doi.org/10.1016/j.talanta.2020.120748>
- Teófilo, R. F., Martins, J. P. A., & Ferreira, M. M. C. (2009). Sorting variables by using informative vectors as a strategy for feature selection in multivariate regression. *Journal of Chemometrics*, 23(1), 32–48. <https://doi.org/10.1002/cem.1192>
- Trullols, E., Ruisánchez, I., & Rius, F. X. (2004). Validation of qualitative analytical methods. *TrAC - Trends in Analytical Chemistry*, 23(2), 137–145. [https://doi.org/10.1016/S0165-9936\(04\)00201-8](https://doi.org/10.1016/S0165-9936(04)00201-8)
- Vaclavik, L., Cajka, T., Hrbek, V., & Hajslova, J. (2009). Ambient mass spectrometry employing direct analysis in real time (DART) ion source for olive oil quality and authenticity assessment. *Analytica Chimica Acta*, 645(1–2), 56–63. <https://doi.org/10.1016/j.aca.2009.04.043>
- Wu, J., Dong, J., Dong, W., Chen, Y., & Liu, C. (2016). Rapid authentication of adulteration of olive oil by near-infrared spectroscopy using support vector machines. *Infrared Technology and Applications, and Robot Sensing and Advanced Control*, 10157(11), 1015701. <https://doi.org/10.1117/12.2244294>
- Xiaobo, Z., Jiewen, Z., Povey, M. J. W., Holmes, M., & Hanpin, M. (2010). Variables selection methods in near-infrared spectroscopy. *Analytica Chimica Acta*, 667(1–2), 14–32. <https://doi.org/10.1016/j.aca.2010.03.048>
- Yan, J., van Stuijvenberg, L., & van Ruth, S. M. (2019). Handheld Near-Infrared Spectroscopy for Distinction of Extra Virgin Olive Oil from Other Olive Oil Grades Substantiated by Compositional Data. *European Journal of Lipid Science and Technology*, 121(12), 1–11. <https://doi.org/10.1002/ejlt.201900031>
- Yun, Y. H., Wang, W. T., Deng, B. C., Lai, G. B., Liu, X. bo, Ren, D. B., Liang, Y. Z.,

Fan, W., & Xu, Q. S. (2015). Using variable combination population analysis for variable selection in multivariate calibration. *Analytica Chimica Acta*, *862*, 14–23. <https://doi.org/10.1016/j.aca.2014.12.048>

ANEXO I

Figura 1A - Amostras de azeite de oliva extra virgem adulteradas com óleo de soja nas concentrações de 1%, 2%, 5%, 10%, 15% e 20% m/m.

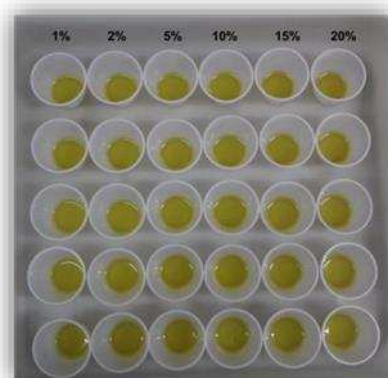


Figura 2A - Espectrofotômetro Büchi NIRFlex N-500 com transformada de Fourier.



Figura 3A - Acessório em cromo para análise em espectrofotômetro no modo de transfectância.



Figura 4A - Suporte para amostras do tipo XL para análise em espectrofotômetro.

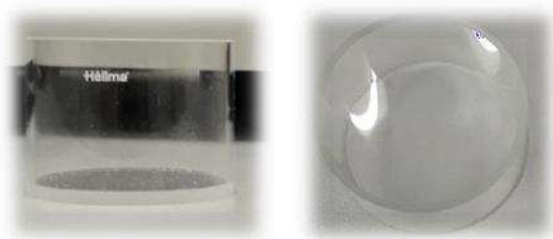


Figura 5A - Sistema para realização da extração em fase sólida com recipiente de vidro com *manifold*.

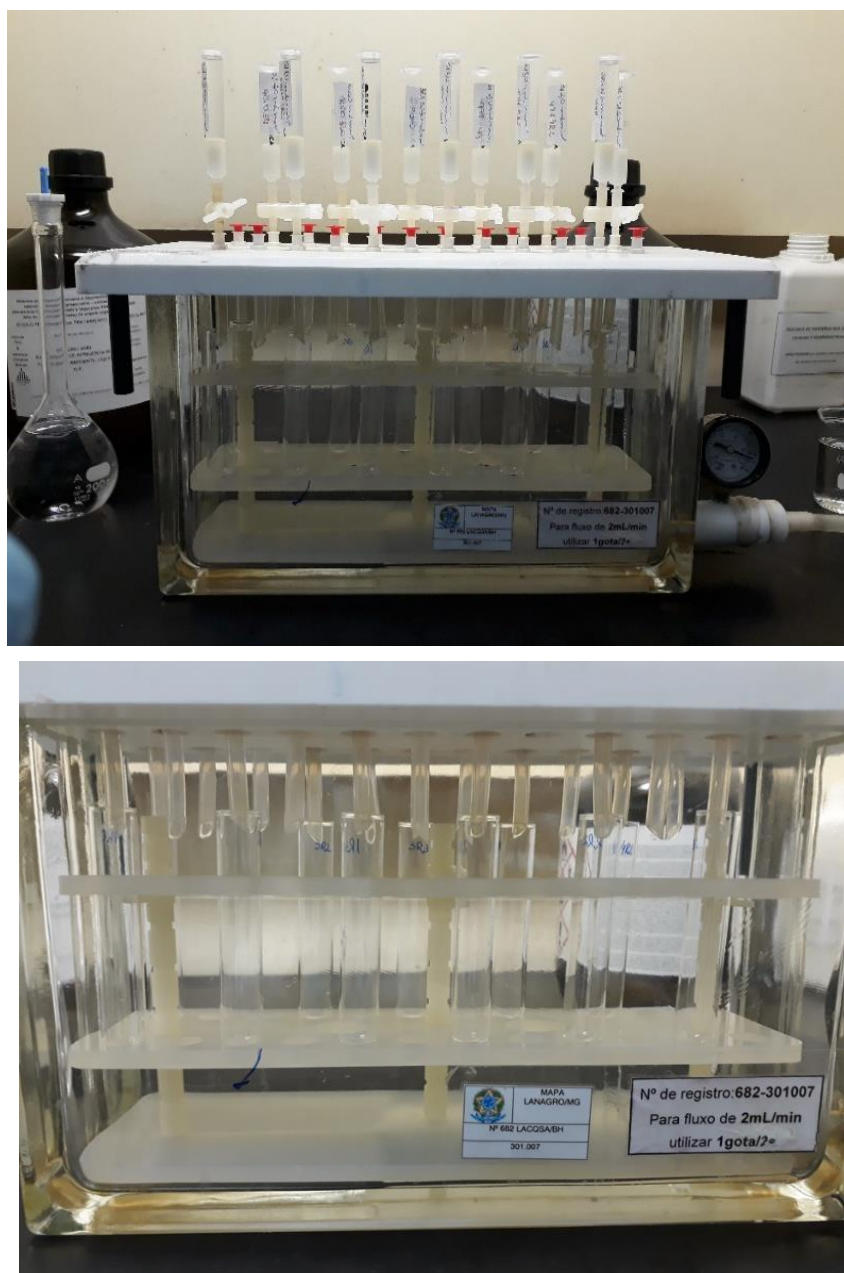


Figura 6A - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante azeite virgem. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.

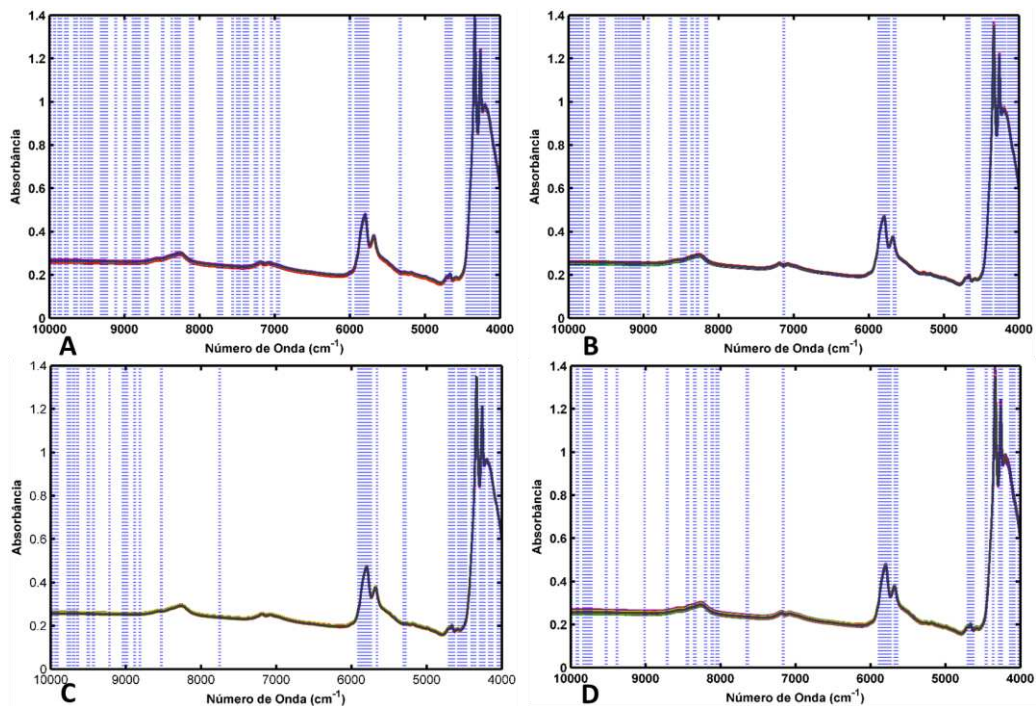


Figura 7A - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante azeite tipo único. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.

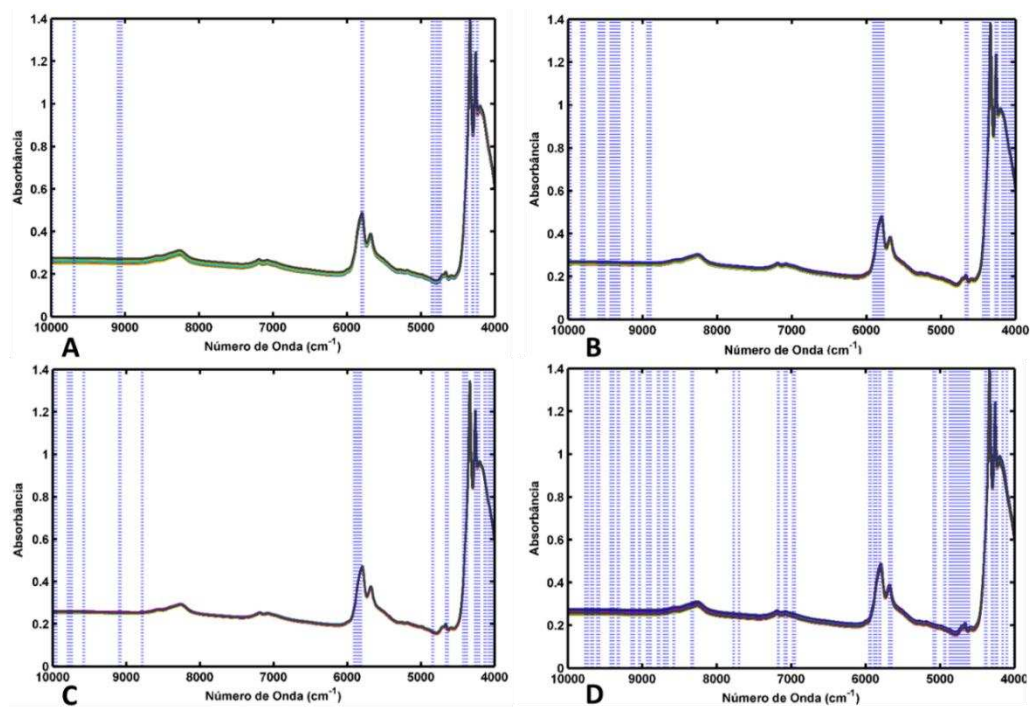


Figura 8A - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante óleo de canola. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.

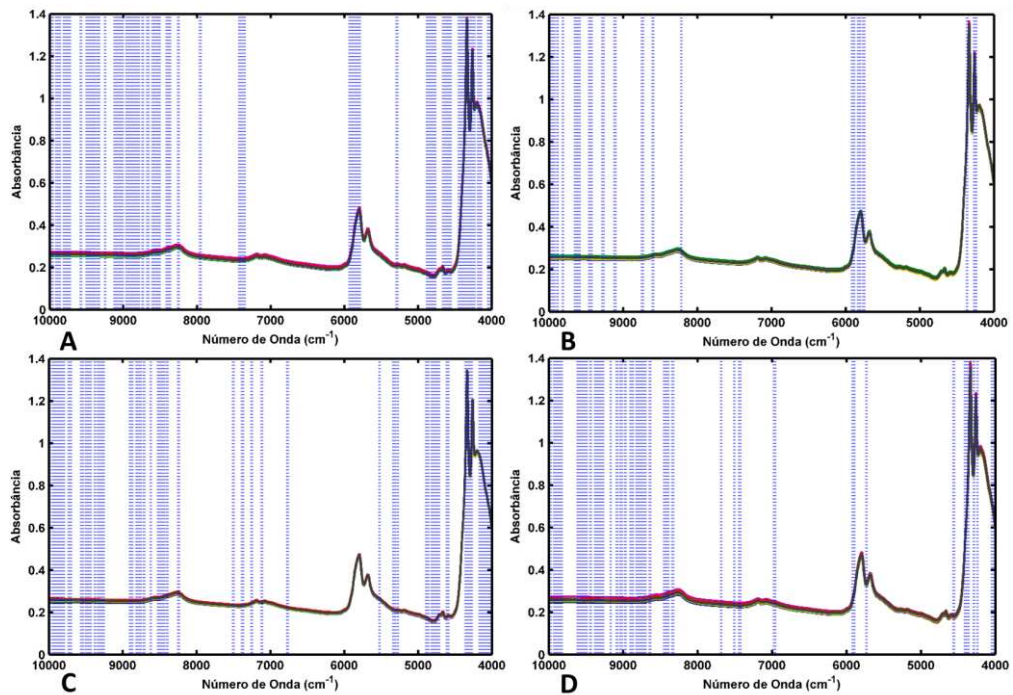


Figura 9A - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante óleo de girassol. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.

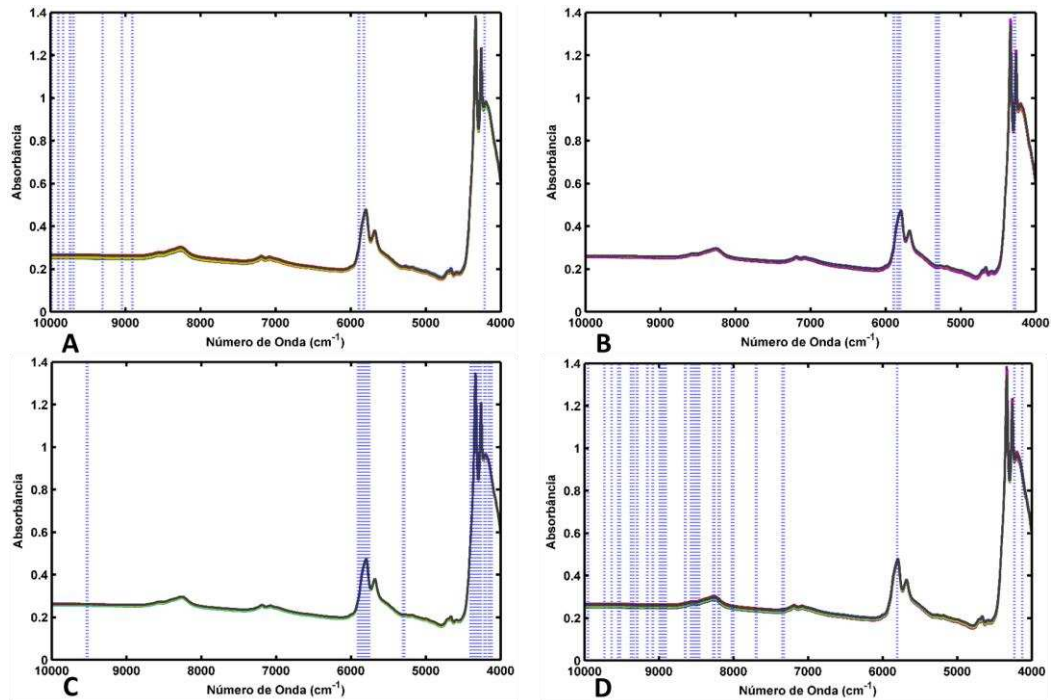


Figura 10A - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante óleo de milho. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.

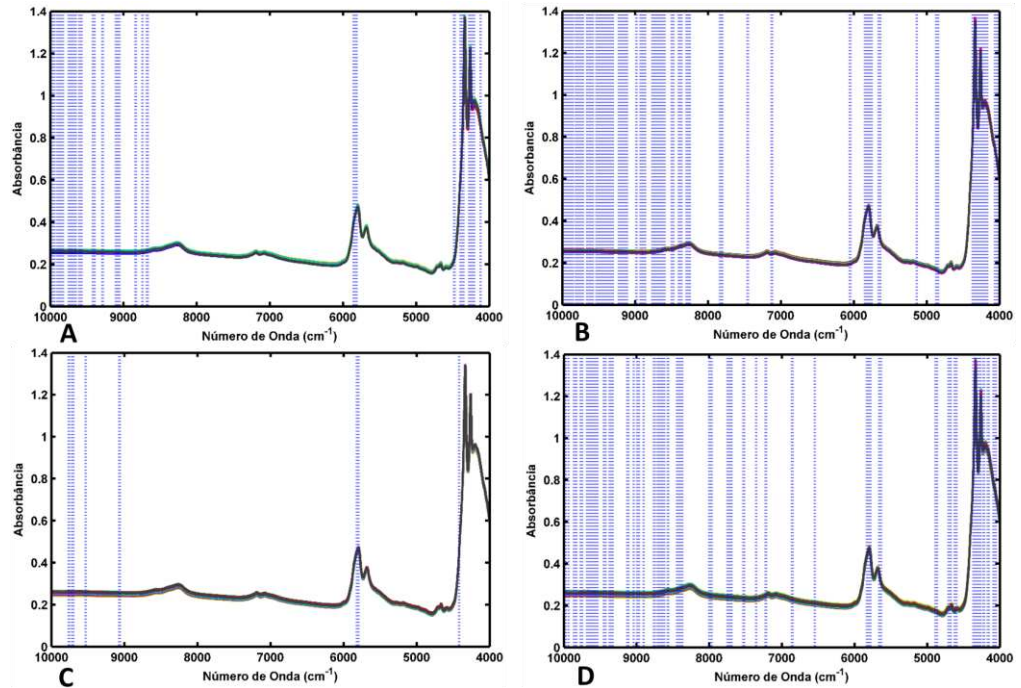
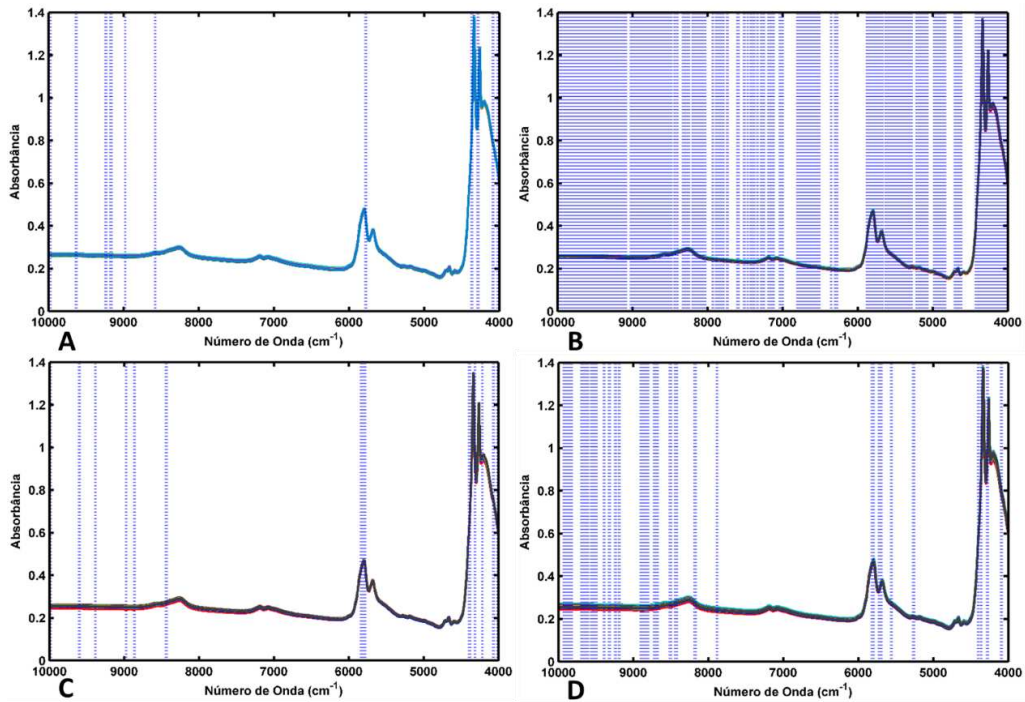


Figura 11A - Variáveis selecionadas para os modelos PLS-OPS de calibração para o adulterante óleo de soja. A: Arbequina; B: Koroneiki; C: *Blend*; D: Robusto.



ANEXO II

Artigo publicado na revista Food Chemistry (Fator de impacto: 6,306)

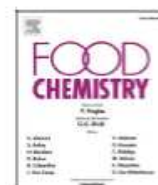
Food Chemistry 345 (2021) 128866



Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem



Building robust models for identification of adulteration in olive oil using FT-NIR, PLS-DA and variable selection

Laurence Souza Vieira, Camila Assis, Maria Eliana Lopes Ribeiro de Queiroz^{*},
Antônio Augusto Neves, André Fernando de Oliveira

Chemistry Department, Federal University of Viçosa (UFV), 36570-000 Viçosa, MG, Brazil

ARTICLE INFO

Keywords:

Olive oil
Adulteration
FT-NIR
PLS-DA
Variable selection

ABSTRACT

Being a product with a high market value, olive oil undergoes adulterations. Therefore, studies that make the verification of the authenticity of olive oil more efficient are necessary. The aim of this study was to develop a robust model using FT-NIR and PLS-DA to discriminate extra virgin olive oil samples and build individual models to differentiate adulterated extra virgin olive oil samples. The best PLS-DA-OPS classification model for olive oils showed specificity (Spe) and accuracy (Acc) values higher than 99.7% and 99.9%. For the classification of adulterants, PLS-DA-OPS models presented values of Spe at 96.0% and Acc above 95.5% for varieties. For the blend, the best PLS-DA-GA model presented Acc and Spe values greater than 98.2% and 98.8%. Reliable and robust models have been built, allowing differentiation from seven adulterants to genuine extra virgin olive oils.

<https://doi.org/10.1016/j.foodchem.2020.128866>